technopolis |group|

5th September 2011

Evaluation of innovation activities: methods and practices

Report of the Literature Review

technopolis |group|

Evaluation of innovation activities: methods and practices

Report of the Literature review

Nelly Bruno
Flora Giarracca
Andrej Horvath
Todor Kamburow
Cristina Navarrete Moreno
Lark Parker-Rhodes
Paul Simmonds

technopolis |group|,
5th September 2011

Table of Contents

1. Introduction: methodology of the literature review and sources 3
   1.1 This paper 3
   1.2 Analytical framework 4
   1.3 Data sources for the literature review 5
   1.4 ERDF evaluation reports 6
   1.5 Academic literature 8

2. Review of evaluation methods and practices by type of innovation support measure 13
   2.1 Innovation management support and dissemination 13
   2.2 Intermediary bodies and agencies 20
   2.3 Start-ups and spin-offs 28
   2.4 Science-industry cooperation 36
   5.2 Strategic research 44
   5.3 Direct financial support for innovation activities 50
   5.4 Networks and clusters 56

6. Conclusions and recommendations 66
   6.1 Overall 66
   6.2 Horizontal analysis of evaluation tools and methods 69
   6.3 Innovation management support and dissemination 75
   6.4 Intermediary bodies and agencies 75
   6.5 Start-ups and spin-offs 76
   6.6 Science-industry cooperation 78
   6.7 Strategic research 79
   6.8 Direct financial support for innovation activities 80
   6.9 Networks and clusters 82
   6.10 Possible case studies 83

Appendix A Bibliography 85

Appendix B Definition of evaluation criteria, methods and approaches used in the Literature Review 88

Appendix C List of reports used by sections 92

Appendix D Country abbreviations used in the Literature Review 98

Table of Figures

Figure 1 List of categories of innovation support measures used in the study 4
Figure 2 Main sources of materials on evaluation of innovation measures 5
Figure 3 Number of reports used for in-depth analysis in the Literature Review 6
Figure 4 Criteria used in the evaluation of Innovation management support and dissemination 14
Figure 5 Methods used in the evaluation of Innovation management support and dissemination 16
Figure 6 Criteria used in the evaluation of Intermediary Bodies 22
Figure 7 Methods used in the evaluation of Intermediary Bodies 23
Figure 8 Good practice in the evaluation of Intermediary Bodies 25
Figure 9 Criteria used in the evaluation of Start-ups and Spin-offs 30
Figure 10 Methods used in the evaluation of Start-ups and Spin-offs 32
Figure 11 Good practice in the evaluation of support to Start-ups and Spin-offs 35
Figure 12 Criteria used in the evaluation of Science-industry cooperation 39
Figure 13 Methods used in the evaluation of Science-industry cooperation 40
Figure 14 Criteria used in the evaluation of Strategic research 45
Figure 15 Methods used in the evaluation of Strategic research 47
Figure 16 Good practice in the evaluation of measures to support Strategic research 49
Figure 17 Criteria used in the evaluation of Direct financial support to innovation activities 51
Figure 18 Methods used in the evaluation of Direct financial support to innovation activities 52
Figure 19 Good practice in the evaluation of Direct financial support to innovation activities 55
Figure 20 Theoretical underpinning of clusters 57
Figure 21 Intervention logic for clusters and networks 58
Figure 22 Criteria used in the evaluation of Networks and clusters 59
Figure 23 Methods used in the evaluation of Networks and clusters 60
Figure 24 Differences in evaluation across science and innovation 67
Figure 25 Analysis of evaluation tools and methods and their advantages and limits for the evaluation of innovation support measures 69
Figure 26 List of potential candidates for case studies issued from the evaluation reports reviewed 83
Figure 27 List of country abbreviations (EU countries) 98
Figure 28 List of country abbreviations (Non-EU countries) 99

1. Introduction: methodology of the literature review and sources

1.1 This paper

This paper presents the results of a literature review carried out as part of a larger study exploring current practice in the evaluation of innovation activities, which is being undertaken on behalf of DG Regional Policy. The paper is one of several deliverables that will be appended to the interim report of the larger study, and should be read in conjunction with those other papers.

The literature review has gathered together a large number of evaluations of innovation support measures, in order to come to a view on the current state of the art across Europe, and to provide an empirical reference point for the study more generally. Specifically, the literature review has sough to:

- Identify the approaches used for evaluating different kinds of innovation activities commonly funded by the ERDF, including mixed support
- Analyse the advantages and limits of these methods for evaluating different types of innovation measure
- Examine in particular the use of selected data collection tools and analytical techniques that are in common use: questionnaire surveys, impact analysis, network analysis, as well as methods used to assess mixed support and behavioural change

For each type of innovation activity selected, we sought to obtain 5-10 published evaluations from a mixture of regional settings, and ultimately compiled a bibliographic database and repository of 58 evaluations. This portfolio comprises a good spread of innovation evaluations, however the repository of reports is not exhaustive.[1]

The study team created a log of the material scanned (as a shared file on Google Docs) to facilitate a decision on what material to include in the analysis. The papers were deposited in a hard-disk folder or repository and a bibliographic database constructed. The relevant items were then profiled using their methodological descriptions:

- Evaluation questions addressed
- Data sources and data collection methods
- Analytical techniques used

We have made use of this portfolio of evaluation reports in two ways:

- Profiling the basic study design parameters (questions, methods, analytical techniques) for each evaluation, in order to carry out some simple descriptive statistics. The profile analysis is presented in the first chapter of this report
- Reading the methodological descriptions in each evaluation report, in order to explore the rationale for choosing a particular approach and to test the extent to which those design choices are determined by the type of innovation measure (the evaluation entity). These in-depth assessments of clusters of reports have been informed by a more general reading of the academic literature and evaluation guidance

[1] It is possible that readers will be aware of excellent evaluation reports in the public domain that were not identified by either our survey of Managing Authorities or our literature review, however we trust that these missing 'gems' will not detract from the relevance and insight of the analyses we have carried out.

1.2 Analytical framework

The literature review has made use of the taxonomy of innovation measures developed in the preparatory stages of the wider study, to provide a common analytical framework for each of the methods investigated.

Figure 1 List of categories of innovation support measures used in the study

	Type of measure	Description
1	**Direct financial support for innovation activities**	Support for R&D and demonstrator projects (through loans or grants)
2	**Innovation management support and dissemination**	Support for non-R&D related aspects of innovation such access to advice and training for innovation related management or for entrepreneurship, etc.
3	**Intermediary bodies and agencies**	Support for intermediary organisations to facilitate technology transfer, including science parks and technology transfer agencies, poles and incubators.
4	**Start-ups and Spin-Offs**	Mechanisms aiming to support the creation and growth of new firms, including seed funding and venture capital.
5	**Networks & Clusters, collaboration and Technology/Knowledge Transfer**	Support aimed at the development of inter-organisational cooperation in the production and transfer of knowledge / innovation. Generally involves inter firm networks rather than individual collaborations. Can involve mobility of personnel.
6	**Science – industry cooperation**	Support for linkages or direct cooperation between science (including both HEIs and public research establishments) and industry to facilitate/promote exchange of knowledge. Can involve mobility of personnel.
7	**Support for the development of ICT**	Support for the uptake of ICT by firms and households, support for the supply and demand of ICT products and services including e-government, e-business, e-learning and e-health, broadband infrastructures
8	**Strategic research**	Promotion of research and innovation activities in thematic areas of regional and/ or regional strategic interest.

The first seven categories listed above were identified during the inception phase of the present study in discussion with the scientific officer and agreed with the steering committee. In essence, they encompass the categories of ERDF measures. They were therefore used in the questionnaire survey, when asking managing authorities to classify the evaluation studies they had commissioned. They were also used to structure the literature review. However, there were two small departures from this list of seven categories:

- Firstly, since very few of the evaluation reports addressed support for the development of ICT, we chose to exclude this category from the empirical analysis of evaluation practice;
- Secondly, given the scope of evaluation reports collected, we added a new category, which we entitled “strategic research”. This new category takes its name from the OECD’s Frascati Manual (2002), where strategic research is defined as being that part of total applied research that has evident important social or commercial relevance, and yet is not sufficiently advanced to have specific practical applications, which is to say it sits somewhere between public and private interests. It might otherwise have been named use-oriented basic research

following Donald Stokes' seminal work discussing the connection between science and technological innovation.[2]

On the one hand, the relatively important number of reports focusing on thematic/sectoral research and innovation in our repository calls for this inclusion. On the other hand, strategic research and innovation is often a focus of support co-funded by ERDF funding, under the form support to area of strategic importance for regional interest – e.g. urban, rural, coastal, maritime, touristic, agricultural and food-related innovation activities. We are therefore positive that this addition might be helpful for Managing Authorities in charge of evaluation ERDF co-funded innovation support measure.

A glossary of the evaluation criteria, methods and approaches investigated in the literature review is available in Appendix B, based on the EVALSED guide.

1.3 Data sources for the literature review

Figure 2 summarises the main sources of information and evaluation materials that we have drawn on in carrying out the literature review.

Figure 2 Main sources of materials on evaluation of innovation measures

Source	Description
EU-level evaluation materials	• EVALSED guide and sourcebooks
DG REGIO evaluation network	• Innovation papers
Inno-Appraisal database	• Structured database of evaluations undertaken in the EU27 • Summary reports • http://www.proinno-europe.eu/appraisal
ERAWATCH TrendChart database of policy measures	• Structured information on innovation policy measures including section on results and evaluations • Annual country reports and updates on innovation policy
Regional Innovation Monitor (RIM)	• Structured information on regional organisations and policy measures in favour of innovation • Forthcoming regional innovation reports for 50 regions (early 2011)
National innovation agencies	• http://www.fteval.at/cms/ • http://www.evaluationsonline.org.uk • http://www.tekes.fi • http://www.vinnova.se • etc.
Academic literature & papers	• Major journals including: – Regional Studies – Evaluation – Research Evaluation – Research Policy • Working papers/reports of specialised research institutes, think tanks and consultancy companies.

The principal data sources were the CORDIS and RIM databases, which together generated more than 1,300 leads and this was supplemented by 196 additional references (URLs) secured through the survey of managing authorities. The 1,500 leads were followed up systematically in order to build a repository of relevant

[2] Donald E. Stokes, Pasteur's Quadrant - Basic Science and Technological Innovation, Brookings Institution Press, 1997.

evaluation reports and specific guidance material. Unfortunately, the very great majority (70%+) of the leads simply linked to an organisational web page with an organisational profile or scheme description, but not published reports. Where the links did connect to a downloadable report, the very great majority of those documents were annual reports or programme descriptions of some sort and were not evaluations. The residual group of evaluation reports were then screened individually to confirm that they did indeed include a discrete review of one or other of the eight types of innovation activities in scope, and we ultimately arrived at a portfolio of some 58 relevant reports.

On balance, the search and screening process proved to be less productive than had been anticipated either in terms of the numbers of evaluations obtained or the quality of those reports. Three points stand out, which may warrant further reflection:

- A large proportion of organisations that fund innovation support measures either do not publish the evaluations they commission, or do so only very occasionally and selectively. Assuming innovation measures are being evaluated reasonably frequently, which the survey of managing authorities suggests is the case, there may be value in pressing for more open publication, of summaries at least, and possibly in both the national language and one other (English?)
- The format and presentation of many of the evaluation reports was rather poor, inasmuch as most reports do not include a specific chapter or appendix explaining the choice of methodology or any reflection on how it might be improved in future. As with the previous point, good practice would suggest that every evaluation ought to reflect on its study design and lessons learned as a means by which to support learning among funding agencies and practitioners
- The majority of ERDF evaluation reports focus on testing the coherence of investments (alignment with operational programmes) and reconciling project outputs with contracted results. Only a minority looked explicitly at the effectiveness of the specific innovation measures supported.

Figure 3 shows the number of evaluation reports used for qualitative analysis in every category of innovation support measures.

Figure 3 Number of reports used for in-depth analysis in the Literature Review

Innovation Measure	Number of reports used by categories
Direct financial support for innovation activities	7
Innovation management support and dissemination	6
Intermediary bodies and agencies	11
Start-ups and Spin-Offs	8
Networks & Clusters, collaboration and Technology/Knowledge Transfer	12
Science – industry cooperation	11
Strategic research	7
Total number of reports used	58 (of which 4 were used twice)

1.4 ERDF evaluation reports

The very great majority of the evaluation reports referred to previously was identified through our desk research, and only a handful of those 58 reports addresses ERDF-financed innovation measures directly. Most are simply published evaluations of types of innovation scheme that ERDF commonly co-finances, albeit they were commissioned and paid for by regional and national agencies independent of ERDF. The focus for the literature review was on good practice in the use of tools and

methods to evaluate such measures, rather than ERDF practice per se, which was addressed by the survey of managing authorities.

The survey, which was run towards the end of the literature review process, did identify a long list of leads:

- We followed up each of the 196 URLs provided in the survey, with some 176 linking to actual reports. Of these, 29 were duplicates of entries in our existing repository, leaving 142 new documents. A further 29 comprised general papers rather than evaluation reports, leaving a total of 118 evaluation reports. Of 118 evaluation reports, 40 had a strong focus on innovation.
- This group of reports look rather different in their focus and style in comparison with the 58 reports that form the primary focus for our empirical analysis. This second group of ERDF reports tended to be somewhat more formulaic in their style, with few if any methodological specificities.

The 40 reports were predominantly ongoing evaluations with a focus on organisation and coherence issues rather than effectiveness, impact or efficiency in the sense of value for money. This emphasis on coherence is a natural reflection of the ERDF rationale and its co-financing of and integration with numerous local measures: its investments must be shown to dovetail with the manifold efforts of local actors. A second and related critical test, is the extent to which planned investments have proved possible in practice as regional bodies and organisations have endorsed the strategic focus and come forward with high-quality project proposals and matching funds. Furthermore, ERDF evaluations tend to be commissioned by Managing Authorities (MAs), organisations that have an especial interest in specific efficiency questions related to management, coordination, communication towards potential beneficiaries and the usual organisational aspects of the ERDF funding.

While effectiveness and impact may be less of a focus of ERDF evaluations compared to other types of evaluations reviewed here, it is nonetheless a question that studies do touch upon, but typically in a rather qualitative manner and principally through a limited number of semi-structured interviews conducted with various regional stakeholders. It is also the case that the ERDF modus operandi, with its tactical support for many aspects of rather mixed portfolios of economic development, makes any kind of robust impact assessment deeply challenging.

Notwithstanding these observations, a small minority of the ERDF-evaluation reports identified do include an in-depth assessment of programme impacts. Here one could point to the Hungarian Impact evaluation of grants for SME modernisation in the framework of National Development Plan 2004- 2006[3] or to the Evaluation of the Berlin Senate's support for Innovation and Technology.[4] Both of these evaluations are the subject of an in-depth case study, and their methodologies will be presented at some length in the appendices to the final report.

In line with the focus of the evaluation questions, most of these ERDF innovation-related evaluations are based on the following methods:

- Desk research to compile administrative and other secondary data (including strategic / policy papers, past evaluations of similar operational programmes, monitoring reports / data and any relevant contextual statistics for the region in question (e.g. time-series data on growth in economic output) to answer the core questions about relevance and programming efficiency;

[3] PPH Consulting, Evaluation of NDP I „investment support granted for small and medium enterprises (ECOP 2.1.1), 2010

[4] PWC, Evaluierung der Berliner Innovations- und Technologieförderung der Senatsverwaltung für Wirtschaft, Technologie und Frauen, 2010

- Semi-structured interviews and possibly workshops with beneficiaries and stakeholders to again look at the core evaluation questions;
- Several of the 40 included questionnaire surveys directed to beneficiaries and also selected case studies to describe and show case particularly successful projects.

1.5 Academic literature

There is a very large academic literature that discusses the relationship between research and innovation and economic development, and expounds on the role of governments in helping to address various classic market failures that mean private enterprise will tend to under-invest in these areas when considered from a public (societal) rather than private perspective.

Within this broad literature, there is a body of work concerned with the measurement of the effects of research and innovation on the economy. This work can focus on measuring micro-economic impacts (the effects of individual and firm-level responses) to the research and innovation policy, or on macro-economic impacts (the analysis of the behaviour of the economy at the aggregate level in response) to the research and innovation policy, or more usually, it can focus on both.

1.5.1 Microeconomic Impact Analysis

Two types of Microeconomic Impact analyses are explored by Bach et al. (1992),[5] based on the previous work of Grilliches.[6,7] Their work identifies two well-recognised, but important techniques for measuring the impact a policy has that is intended to stimulate technological innovation, both concerning the measurement of **consumer welfare**. According to the neoclassical theory of the firm, and assuming perfect competition as a suitable approximation of the market being studied, technological innovations should be fully captured by the pricing mechanism and should manifest themselves either as a consumer surplus, or as a producer surplus. Consumer surplus is basically derived from the consumer demand curve and represents the difference between what the consumer is willing to pay for a product and what they have to pay. The producer surplus is basically the inverse of this concept. It is derived from the producers supply curve and represents the difference between the marginal cost of production and the actual price demanded (profit).

Still following Grilliches, Bach et al. (1992), identify three principle challenges to the evaluator surrounding the difficulty of measuring these effects:

- 'The complexity of the relationship between the suppliers at each step of the production process'[8] – this can make it hard for the evaluator to identify where the technological spillover has taken place in each instance, and multiple steps can make it hard to follow the diffusion of the technology down the value-chain
- 'The ability of price indexes used by evaluators to reflect the change in the quality of the product'[9] – this can be because as general prices change over time to due to inflation, the exact real change in price can be obscured and can depend heavily on the index chosen to calculate the real price

---

[5] L. Bach, P. Cohendet, G. Lambert and M.J. Ledoux, *Measuring and managing Spinoffs: the case of the spinoffs generated by ESA programs*, in 'Space Economics', Volume 144, Progress in Astronautics and Aeronautics, 1992.

[6] Griliches. Z., *Issues in Assessing the contribution of R&D to productivity Growth*, Bell Journal of Economics, Vol. 10, No. 1, 1979, pp. 92-116

[7] Griliches. Z., *The search for R&D Spillovers*, Working paper, Harvard University, Cambridge, MA, 1990.

[8] L. Bach, et al, ibid. p. 180

[9] L. Bach, et al, ibid. p. 180

- 'The competitive structure of the industry determining the distribution between buyers' and suppliers' surplus'10 – depending on whether market tends toward perfect competition or monopoly, or whether there are any other market disturbances such as price controls, the consumers will have more or less power in the market, and the evaluator should concentrate their efforts on capturing the effect on the producer or consumer surplus.

1.5.2 Macroeconomic Impact Analysis

As mentioned above, Macroeconomic impact analysis focuses on the wider returns to the economy on the public investment in technological innovation. Georghiou and Rossener,[11] discuss the work of Cozzens et al.,[12] wrote a report for the White House Office of Science and Technology Policy where they reviewed the methods available for evaluating fundamental science. They found that one of the major ways that Macroeconomic impact has been analysed by evaluators is through 'econometric methods employing, as measures of performance, productivity growth, increase in national income, or improvements in social welfare.' This method will tend to take key indicators or the success of the program or scheme, such as the number of patents that have been produced, the amount of extra employment created or even the change in consumer or producer surplus. This is then regressed using the method of ordinary least squares, or one of its many variants, and regressed against an indicator of the health of the economy, such as GDP, unemployment statistics or the turnover of similar firms in the region.

The major problems that can occur with this kind of analysis are:

- Can reasonable indicators of the outputs of the policy be found? For example, is there good data on the number of patents that have been produced? Can the evaluator trace whether these patents have been taken up by industry? And if so what has been done with them?
- Is there a strong relationship between the indicator of the health of the economy and the effect of the policy? There need to be a strong relationship between the indicator chosen by the evaluator and the outputs of the policy, otherwise the effect of the policy can be obscured by other factors. For example, if GDP is chosen as the dependent variable, there may be a strong 'crowding out' effect as there are many countervailing tendencies also being captured. This problem is in some ways rectified by the use of independent variables, which are designed to 'hold still' (ceteris paribus) other factors that may be having an effect.
- Is the relationship linear, or can it be otherwise captured by a suitable regression? – Innovation is not linear: there are often many feedback loops and diversions, which can make tracing the effect of a new technology in the economy less tractable

Another econometric method used estimate the value of knowledge spillovers in the country is the NEMSIS model (New Econometric Model of Evaluation by Sectoral Interdependency and Supply), which has been funded by the 5th and 6th Framework Programmes to model all structural policies that involve long term effects, including RTD. This model attempts to isolate the influence of the stock of knowledge in a number of economic sectors on the productivity of a particular sector. This is achieved by using the OECD's 'Technology Concordance' data, which maps International Patent Classification data into various economic sectoral classifications. The influence of the

---

[10] L. Bach, et al, ibid. p. 180

[11] Georghiou, L. and Rossener, D., *Evaluating technology programs: tools and methods*, Research Policy, 29, 2000, pp. 657-678

[12] Cozzens, S. Popper, S., Bonomo, J., Koizumi, K. Flanagan, A., Methods for evaluating fundamental science, Washington D.C 1994

economic sector under consideration is then analysed by its 'technological proximity' to the other sectors. 'Technological proximity' is measured by the extent of overlap of the technological classification of their patents.

It is widely accepted in the literature that evaluating the effects of research and innovation programmes require at a minimum an adjustment of the 'standard' tool-box of evaluation methods and indicators. For instance, Fleischhauer (2005) argued that there are inherent difficulties in the evaluation of innovation-related activities and identified three specific issues:

- Innovation is complex and uncertain; there is no "guarantee" that the public resources invested in a project will generate innovation.
- Innovation can only be appraised in the long run, but policy-makers and society ask for evaluations to report on short-term efficiency.
- Innovation causes complex and multiple effects that do not evolve linearly. Their evaluation is delicate and makes a linear impact analysis impossible.

The European Commission has invested in the development of numerous evaluation guidelines, including several that address innovation explicitly. The MEANS guidelines (1999)[13] are perhaps the most ambitious and noted that innovation should be seen as a 'dynamic, interactive and non-linear process' implying that trying to evaluate impact in a simplistic 'cause' (innovation projects in enterprises, etc.) and 'effect' (e.g. higher sales from innovative products) is likely to prove insufficient. They point to practical solutions to the methodological challenges, which range from the use of multiple data collection methods – to facilitate some basic triangulation – and the active use of monitoring data and other ongoing, longitudinal data collection methods, to better capture evolutionary perspective. And larger samples – even whole populations – to avoid the risk of missing important developments and technological breakthroughs (skewedness).

Use of multiple approaches both within the microeconomic analysis and outside it, can help ameliorate the kind of dilemma that can be involved in interpreting evaluation results highlighted by Klette et al (2000)[14]:

> *"... we face the paradoxical situation that if an evaluation study finds little difference between the supported firms and the non-supported firms, it could either be because the R&D programme was unsuccessful and generated little innovation, or because the R&D programme was highly successful in generating new innovations which created large spillovers to non-supported firms".*

The risky and unpredictable nature of innovation leads Perrin (2002)[15] to argue that this implies adopting a 'key exceptions' approach rather than checking the mean results of projects against a pre-set quantified target. Equally 'simplistic models of impact' that assume a direct-cause effect relationship such as a 'return on investment' of R&D and innovation funding fail to understand that 'innovation never occurs alone' but always within a context of structured relationships, networks, infrastructures and in a wider social and economic context. Hence, the need to take into account the context of a specific measure: a grant to encourage regional enterprises to undertake collaborative R&D and technology transfer from academic or public research labs will be ineffective if the enterprises do not have qualified staff able to work with their counterparts in the research labs or if the incentive system in the labs does not

[13] MEANS (1999) Guide to Methods for Evaluating Structural Policies, 2000. Other touchstone references include the ASIF report, "RTD Evaluation Toolbox- Assessing the Socio-Economic Impact of RTD Policies", Fahrenkrog, Gustavo, Wolfgang Polt, Jaime Rojo, Alexander Tubke, Klaus Zinöcker (2002), European Commission.

[14] MEANS (1999) Guide to Methods for Evaluating Structural Policies, 2000

[15] Perrin, B., 'How to — and How Not to — Evaluate Innovation', Evaluation, Vol.8, No. 1, 13-28, 2002

'reward' commercialisation of research but rather number of academic papers published.

As Boekholt et al (2001)[16] and Arnold et al (2009)[17] note, many of the available techniques allow us to demonstrate with a fair degree of confidence that there are effects, and sometimes to say that these effects may be quite large in comparison with the state's investments in the intervention being evaluated. A key concern of many policymakers is not to prove direct effects of single measures (or projects) but rather to understand the relative effects of different types of intervention, since they want to optimise the allocation of scarce financial resources. No single technique is on its own sufficient to obtain a robust evaluation result. R&D evaluators therefore tend to take care to use several methods in combination and to look for convergence among the results they provide.

In practical terms, one has seen efforts to improve our ability to estimate wider economic effects – realised through spillovers for example – in the work of the Value for Money review of the Science Foundation Ireland (with an estimation of the impacts of the reviewed programmes on human capital development in Ireland) or in the Estonian State's enterprise supports on the competitiveness of the economy.

Elsewhere, it has become commonplace for budget holders and evaluation practitioners to broaden their search for benefit types beyond the economic to encompass behavioural, social and environmental effects. In essence, making a better job of detailing the spectrum of benefit types. And this broadening of the investigative envelope has been mirrored by the diffusion of theory-based evaluation in the past decade[18] and the use of logic models[19] to cope with the complexities of these kinds of innovation measures, which seek to make progress often on several fronts and seek to catalyse and connect large, dynamic socio-economic systems. Which is to say, policy makers are increasingly using a theory of change to itemise the kinds of programme effects being sought in terms of changed attitudes, skills, resources, relationships, etc (the intermediate effects, which should all things being equal lead to increased levels of innovation and economic growth).

Almost all of this work concerns the benefits of direct financial support (by governments and other public bodies) to research and innovation carried out by individual business enterprises.

There has been less academic study as regards the link between other types of innovation support measures and the particular tools and methods that might be best used to describe, count and explain a particular outcome.

There is a very substantial literature on the different phenomena and measures of interest, however the links to evaluation tools is rarely addressed and where it is a focus for reflection, the work tends to draw heavily on selected empirical cases. It is not well theorised.

---

[16] Boekholt P., Lankhuizen M., Arnold E., Clark J., Kuusisto J., de Laat B., Simmonds P., Cozzens S., Kingsley G., Johnston R., 'An international review of methods to measure relative effectiveness of technology policy instruments', 2001

[17] Arnold E., Malkin D., Good B., Clark J., Ruiz Yaniz M., 'Evaluating the National Innovation Strategy for Competitiveness'. Report to the Chilean National Innovation Council for Competitiveness, 2009 -

[18] A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009. American Journal of Evaluation, June 1, 2011 32: 199-226

[19] Logic Models: A Tool for Telling Your Program's Performance Story, John A. McLaughlin, Gretchen B. Jordan, Evaluation and Program Planning, Volume 22, Number 1, February 1999. A theory-based logic model for innovation policy and evaluation, Jordan, Gretchen B., Research Evaluation, Volume 19, Number 4, October 2010, pp. 263-273(11). Evaluation of ST&I programs: a methodological approach to the Brazilian Small Business Program and some comparisons with the SBIR program, Salles-Filho, Sergio; Bonacelli, Maria Beatriz; Carneiro, Ana Maria; De Castro, Paula F Drummond; Santos, Fernando Oliveira, Research Evaluation, Volume 20, Number 2, June 2011 , pp. 157-169(13)

Science industry cooperation is an excellent case in point, where hundreds of books, articles and theses have been written extolling the importance of the relationships between these two communities. Few have gone on to explain how such a set of relationships might be measured and how public programmes might be evaluated. The minority of papers that consider evaluation have tended to assume budget holders and evaluation practitioners will use standard social scientific research methods (e.g. questionnaire surveys and interviews) to gather data (indicators) about the nature and extent of those interactions, their input and importance within given functions, their evolution and persistence over time, etc.[20]

Some researchers have begun to analyze university-industry cooperation through patent data analysis. Margherita and Andrea (2006) used patent data integrated with information collected through interviews and measured the extent and intensity of the ties of academic with industrial researchers, and apply social network analysis to reconstruct the network of collaborations.21 Patent data are quite attractive inasmuch as they are recorded systematically and unambiguously (applications or patents granted) and each record comprises a list of named owners and inventors as well as the key pieces of prior art that have been build upon. National, European and US patent offices also maintain good online access to patents and to basic patent statistics, which can provide useful contextual data (underlying frequencies and trends in a given technology field and country for example). It can however, be quite laborious – and involves specialist software – to build up the kind of before and after databases one needs to understand clearly what changes have occurred in the 'productivity' of the relationships between pre-specified parties and to anchor that in a counterfactual analysis of trends in the specific technology field more generally.

[20] University-Industry Relations: A Review of Major Issues, by Eliezer Geisler and Albert H. Rubenstein, Chapter 3 in Cooperative research and development: the industry, university, government relationship, edited by Albert N. Link, Gregory Tassey, Springer, 1989. Assessing the impact of university interactions on an R&D organization, Antonio J. Bailetti and John R. Callahan, R&D Management, Volume 22, Issue 2, pages 145–156, April 1992. A theoretical framework for the evaluation of university-industry relationships, Andrea Bonaccorsi and Andrea Piccaluga R&D Management, Volume 24, Issue 3, pages 229–247, July 1994.

[21] Margherita, B., Andrea, L. (2006). University–industry interactions in applied research: The case of microelectronics, Research Policy, 35(10), 1616-1630.

2. Review of evaluation methods and practices by type of innovation support measure

2.1 Innovation management support and dissemination

2.1.1 Introduction

Six of the evaluation reports in our repository address measures that are directly concerned with innovation management support, which comprises all non-R&D related aspects of innovation, from the provision of information and advice to prospective innovators, to the delivery of training in innovation management and coaching and mentoring support to individual entrepreneurs.

In all cases, the six evaluations were specified and commissioned by the public authorities – regional or national agencies – responsible for the design and oversight of the innovation measures in question (e.g. Scottish Enterprise, Danish Agency for Science, technology and Innovation, etc). The evaluations were carried out by private consultancies with the exception of the economic impact study of the UK Business Link Local Service, which was carried out by a consortium of universities, albeit through a commercial (consultancy) contract.

The budget and duration of the evaluations were not disclosed, however almost all of the studies addressed a broad range of evaluation questions and made use of a mixture of data collection methods and analytical techniques, suggesting they were all substantial exercises in their own right.

2.1.2 Rationale and objectives

Policy makers around the globe believe that businesses under invest in research and innovation due to the classic market failures of information asymmetry, uncertainty over the likely outcomes (risk and reward), and an inability to secure an exclusive flow of benefits. There is also an issue with the innovation capabilities of many enterprises, their absorptive capacity, which exacerbates these information failures and risk aversion. As a result, aggregate innovation rates are constrained and so too are the socio-economic benefits that might be expected to follow a more dynamic approach.

Innovation management and dissemination measures are intended to alleviate these impediments in some degree providing wide-ranging information and support that engenders increased innovative activity amongst large numbers of smaller businesses in particular and complements the much more selective and intensive modes of assistance and in particular direct financial support for R&D pursued elsewhere within the regional innovation policy mix.

They aim at fostering innovation awareness and improved capacities amongst business enterprises through information and advice possibly more substantive support or training subsequent to some form of needs assessment. Topically, these measures cover the spectrum of innovation management issues, from the location of types of financial assistance to the identification of prospective innovation partners with specific qualities or skills and a mutual interest in the field. Given theses measures tend to provide information in the first instance to any and every business, most tend to be delivered by specialist intermediaries whether that is large, geographically extensive technology consultancies or networks of regional agents offering standard information services and advice locally.

Innovation management and dissemination support measures are rather often included within broader innovation support programmes and support structures (i.e.

such as incubators, regional development agencies, etc). Most commonly, they complement direct financial support for R&D, with a view to encourage enterprises to apply for grants, providing awareness, knowledge and professional advice on the innovation system and opportunities, and the innovation grants and related procedures. As a matter of fact, evaluations focusing specifically on this kind of measure are somewhat rare, for the scope of evaluation reports focus usually on a broader range of measures or on the organisations that provide support. The six measures evaluated in our repository can be classified into two categories, according to their rationale:

- Provision of business advice/ training/ innovation management services by specialists in innovation
- Campaigns designed to raise the awareness of enterprises on entrepreneurship activities

The remainder of this section tries to analyse more in-depth how the effects on innovation culture and innovation knowledge are questioned and assessed in the different reports evaluated and how these measures are taken into account into the global outcomes and impact on innovation capacities and innovation performance.

2.1.3 The evaluation record

The evaluation questions

Figure 4 presents the scope of evaluation questions assessed in each of the six evaluations reviewed. It also provides indications on the focus of evaluation (type of scheme evaluated) and on the periodicity of the evaluation (ex-post or mid-term).

Figure 4 Criteria used in the evaluation of Innovation management support and dissemination

Name of the measure evaluated	Definition of the scheme	Type of scheme	Periodicity of the evaluation	Relevance	Efficiency	Effectiveness	Sustainability
Scottish Enterprise Facilitator pilot (UK)	Pilot initiative to test the extent to which the local community can provide mentoring, coaching and business support to local entrepreneurs and businesses	Provision of business advice/ services/training	Ex-post	√	√	√	√
Danish Regional Innovation Agents (DK)	A regional network of innovation specialists working with SMEs to assess their innovation potential and to help them take first steps in developing innovation plans	Provision of business advice/ services/training	Mid-term	√	√	√	√
Business Link Local Service (UK)	Government's support services (advice and coaching) for SMEs	Provision of business advice/ services/training	Mid-term	√	√	√	√
Network of Invention Advisers of the Foundation for Finnish Inventions (FI)	Assistance in the initial phases of the development of inventions by promotion of inventions made by private individuals and small enterprises	Provision of business advice/ services/training	Mid-term	√	√	√	√

Make your mark (UK)	Entrepreneurship awareness campaign to create a more enterprising youth culture amongst 14-30 year olds	Entrepreneurship awareness campaign	Ex-post			√	
Campaign uni:invent (AT)	Entrepreneurship awareness campaign	Entrepreneurship awareness campaign	Mid-term	√		√	√

4 out of the 6 reports are mid-term exercises, and 2 are ex-post assessment since they do not relate to a simple programme but to a range of different measures.

Measuring effectiveness and assessing future development/ sustainability issues is the prime focus of all these reports. As might be expected, effectiveness issues are reviewed in three ways:

- What are the behavioural changes amongst beneficiaries that could be attributed to a raising innovation awareness and innovation knowledge and to increasing cooperation between enterprises and between enterprises other actors (e.g. UK Business Link Local Service; Strategic Review of the Scottish Enterprise Facilitator pilot)
- What are the micro-economic effects on beneficiaries in terms of enterprise creation/ job creation/ patents, productivity etc (e.g. Strategic Review of the Scottish Enterprise Facilitator pilot)
- What are the macro-economic effects on the regional/national economy (e.g. UK Business Link Local Service)

Questions that relate to the efficiency of the measures/structures implemented are also developed by all the reports. To some extent, that might be explained by the fact that intermediary bodies and agencies implement many of the related support measures. Being distinct from these bodies, the funders might be willing to look into the programme efficiency and functioning to see how the money is spent when evaluating the different measures.

Mid-term evaluations also tend to spot relevance issues mainly with the view to identify any change of needs and any changing patterns that might affect the continuation of the programme. The Scottish Review of the Enterprise Facilitator pilot is the most oriented towards relevance and future development issues with a view to offer recommendations whether the pilots should be extended given the changed economic development landscape and how through the assessment of development scenarios.

The evaluation of the Network of Invention Advisers of the Foundation for Finnish Inventions has specific patterns since it focuses mainly on relevance issues – i.e. is it appropriate that public funding for private individuals and small business development is channelled through the Foundation?

2.1.3.1 The evaluation methods implemented

Figure 5 presents the evaluation methods used in the six evaluation reports.

Figure 5 Methods used in the evaluation of Innovation management support and dissemination

	Qualitative methods			Quantitative methods				Approaches			
Name of the measure evaluated	**Individual interviews**	**Case studies**	**Peer review**	**Questionnaire survey**	**Use of administrative/ secondary data**	**Micro-/ macro-economic model**	**Bibliometrics**	**Baseline**	**Counterfactual**	**Benchmarking**	**Indicators**
Scottish Enterprise Facilitator pilot (UK)	√				√					√	√
Danish Regional Innovation Agents (DK)	√			√				√			
Business Link Local Service (UK)	√	√		√		√		√	√		√
Network of Invention Advisers of the Foundation for Finnish Inventions (FI)	√			√	√				√		
Make your mark (UK)				√	√			√			√
Campaign uni:invent (AT)	√			√	√						

The following methodological patterns are underlined in the eight evaluations reports:

- All of them mix qualitative and quantitative analysis with most of the time the implementation of several evaluation tools
- 5 of 6 include a survey of beneficiaries, mainly to canvass the opinions of participants on the functioning of the programme and on the effects and impacts of the programme evaluated. In some cases, the survey also aims at gathering quantitative data used afterwards in the economic modelling - e.g. in the UK Business Link evaluation, the survey was specifically designed to support an econometric approach in order to collect GVA data, business growth (employment and sales) and sales per employee.
- 5 of 6 make use of stakeholder and beneficiary consultations, mainly in the form of interviews. The review of the Network of Invention Advisers of the Foundation for Finnish Inventions includes a workshop with staff involved in the scheme in order to brainstorm around division of work and cooperation with other stakeholders and the impacts of the programme. In most cases, interviews are used along with beneficiaries survey and help to gain a more in-depth analysis on the opinion of stakeholders on the programme and on the cause and effects leading to the outcomes and impacts. Interviews also address most of the time relevance/strategy, efficiency/organisation and future development issues.
- 4 of 6 involve the use of administrative and secondary data, mainly through a review of monitoring data and documents for each measure. Monitoring data are most of the time used in order to document the outputs and, when available, the outcomes of the schemes. When not available or incomplete, these data might be complemented through a survey of beneficiaries. A few also include the results of previous evaluation/ studies. Furthermore, desk study is used in some of the reports to assess the relevance/appropriateness of the different measures. For instance, the Strategic Review of the Scottish Enterprise facilitator documents the genesis and subsequent development of the Sirolli model for mentoring and business support by local entrepreneurs, a model that lay at the basis of the pilot implemented in Scotland.

- 3 of 6 include a baseline approach, mainly to assess the outcomes/ impacts of the schemes over time and after a period of reference.
- 3 of 6 involve a set of indicators. In the evaluation reports reviewed, indicators are used to document the performance of the measures and most of the time a set of limited indicators focusing on micro-economic impacts on the beneficiaries of the programme is used, comprising data such as the growth in employment, creation of new enterprises, number of patents and the gross added value of the project. Indicators are documented either through secondary data issued from monitoring documents, or through surveys results.

Each of the other methods was used in a minority of the six evaluations, and specifically, in descending order of use:

- 2 of 6 include a counterfactual approach, most of them through surveys of beneficiaries and non-beneficiaries through a counterfactual impact evaluation assessing. Whereas some only use a survey question to assess for the counterfactual, one has developed a whole micro-economic model (the UK Business Link evaluation). The counterfactual approach is central to most of the evaluation, since several of the reports put a strong focus on the added value of the measure, compared either to a situation with no measure or to a situation with alternative schemes.
- 1 of 6 includes case studies, mainly based on interviews of selected beneficiaries in order to complete surveys carried out on a large number of organisations.
- 1 of 6 involves benchmarking the performance of the measure in question. For example, the review of the Scottish Enterprise Facilitator pilot assesses the performance of the Sirolli model[22] that sits at the heart of the scheme in part by comparing its measured performance with more mainstream business support models (diagnostics) implemented usually in the UK. Such approaches are based mainly on interviews and desk study comprising evaluations of similar measures.

As might be expected given the scope of the measures that target mainly non-R&D support towards enterprises, none of the evaluations include peer review or bibliometrics.

2.1.4 The advantages and limits of the methods used

The sample of evaluation reports reveal three features that might be described as advantages, which are:

- Service providers are pretty systematic in carrying out pre-support interviews to profile prospective clients and most also administer post-support exit polls to gather feedback on knowledge or skills acquired, changed attitudes and service quality. The monitoring data – on activities and outputs – are usually comprehensive and closely aligned with programme objectives. The quality of the monitoring data reveals itself in the evaluation reports, which devote rather more space and thought to operational efficiency and behavioural changes than do most of the studies addressing other innovation measures. The comprehensive monitoring data is especially helpful in determining the extent to which a service has attracted the right people and delivered the anticipated insights and improved confidence

[22] The Sirolli Model was developed by the Sirolli Institute in the US and has been widely implemented, and involves the creation of a local community of innovation experts and mentors that provide the first line of support to client businesses who discuss their business ambitions with a dedicated Enterprise Facilitator. He or she carries out a reflective process and attempts to identify opportunities for business development or innovation, and looks in the first instance to the enterprise community to help the client business to realise its ambitions

- Questionnaire surveys can also work rather well, as the services are delivered to very large numbers of businesses and one can secure responses in the high hundreds or even thousands, even on a sample basis (permitting more detailed analyses and cross-tabulations)
- Several of the studies given reasonably standard types of advice and support and most people will make changes / see benefits in short order (or do nothing).

On the downside, the interventions are typically pretty small – a piece of advice, an introduction or a few hours training – which clients may be delighted to receive but find it hard to be specific about the benefits.

- Many of the studies struggle to get much beyond counting the numbers of assisted businesses that believe they have derived a great deal – proportionately – from the information or advice provided
- Questionnaire surveys can struggle somewhat with low response rates and a risk of exaggeration where in truth the actual benefits were rather soft / diffuse and may not have sufficient weight for anyone to seriously attribute a measurable, firm-scale benefit. In practice, several of these innovation management measures provide support at two or even three levels of intensity, and the economic impact assessments have focused on the recipients of the more substantive support. Help might progress through a basic enquiry and referral, on to a more substantive training exercise or in-company diagnostic and ultimately, for a proportion of clients, might lead to a further award to support an in-company innovation project (e.g. subsidised access to an approved innovation coach or mentor)
- The evaluations focus most on economic impacts and devote little space to testing the extent to which the measures actually correct for the assumed information failures and generate behavioural changes (i.e. that could be done for instance through micro-economic modelling) in terms of innovation readiness or innovation awareness. This might be explained by the fact that many of the evaluation reports were commissioned by clients in Northern Europe, where there appears to be a relatively greater concern with economic impact and return on investment (value for money).

The large populations and directly attributable benefits do lend themselves to quantification and even counterfactual analyses (using other large and mixed business populations as the comparator) are widely influenced by the nature of the innovation measure(s) evaluated. Most of the reports also adjust the directly observed economic benefits for displacement (i.e. gains by beneficiaries are matched in some degree by losses to their non-participating competitors), substitution, leakage (losses to the country or region as a result of IP/ innovations being moved around within an international conglomerate), and multiplier effects (purchases in local economy by employees of beneficiaries and sales to the region via supply chains). It is interesting to notice that the UK Business link's evaluation makes use of an economic model to assess the impacts of the scheme. The value for money of the scheme is derived from the difference between the attributable costs of the programme extracted from accounting data and the Gross Value Added estimates derived from a survey of beneficiaries and grossed up to the whole of the assisted population. Furthermore, the evaluation implements the same kind of model to answer efficiency issues, i.e. to differentiate business performance according to the different types of Business Link organisations implemented.

2.1.5 Identification of good practice

We would single out for further elaboration the following reports, that present interesting methodological design and combination of methods: **The Economic Impact Study of Business Link Local Service (UK)**, which is quite a complete evaluation in terms of methods used, detailed hereafter:

- The evaluation methods used are diverse and complete each other in the analysis of the data:
  - An extensive telephone survey of approximately 3,500 assisted businesses and a similarly sized control group
  - Face-to-face interviews with 34 firms with a focus on those who received intensive assistance. This provided more detailed information on the more organisational and strategic impact of support
  - Visits / semi-structured interviews with 18 Business Link Organisations, which amounts to around half of all delivery organisations nationally, and the subsequent development of a detailed typology of alternative brokerage models
  - Descriptive statistics
  - Micro- and Macro-economic modelling
- Approaches used for the analysis and based on the previous mentioned tools were particularly suitable to his type of evaluation of innovation management support measures:
  - A national estimate of the value for money (VFM) of the Business Link network, essentially based on estimates of employment growth attributable to the assistance, through econometric modelling of the impact of assistance on performance. Business growth (employment and sales) and sales per employee indicators are used as key proxies for performance measures in the models
  - It is amongst the few evaluations that apply a Theory-Based Impact Evaluation of the impacts, i.e. that tries to look into the black box and assess how the programme is working, based on a reconstruction of the intervention logic of the programme. This was mainly based on firms' views of the quality and impact of service received and its impact on business strategy.
  - A Counterfactual Impact Evaluation, though limited to a question to participating firms under the form of "What would have happened without the programme"
  - An analysis comparing the effectiveness of alternative brokerage models of assistance on business performance since the Business Links Offices are organized in different ways all over the UK. The analysis was based on the impact perceived by firms (survey and interviews) as well as on the econometrically modelled impact of BLO assistance on business growth.
  - A regional spatial analysis (i.e. comparison across the 9 English regions) relating to the type of businesses being assisted by BL in each region to provide an overview of the similarities and differences between intensively-assisted, other-assisted and non-assisted firms in the different Government Office Regions (GORs) across England; completed by am rural perspective to provide headline data on the operation of Business Link in rural areas in England

The Business Link evaluation is an example of a substantive study that is carefully reported and has much to recommend it as a source of methodological insight, albeit it has several important limitations. The first limitation is the reliance upon a single question in the beneficiary questionnaire to determine the counterfactual, which is a less robust approach to establishing the net effects of a particular measure as compared with the kind of randomized control group methodology used in the evaluation of the Dutch Innovation Vouchers. Likewise, the large-scale survey is the principal means by which the study attempts to dimension behavioural effects, however the semi-quantitative approach, quite naturally, lacks qualitative insight and reduces the formative value of the study findings on this point.

2.2 Intermediary bodies and agencies

2.2.1 Introduction

- This category was defined as "Support for intermediary organisations to facilitate technology transfer" which in particular includes Science and Technology parks and Technology Transfer Agencies/Offices, and Business Incubators/Innovation Centres.
- Only 3 out of the 120 or so evaluation reports in our primary repository fitted this definition as two of them were evaluations of Technology transfer programmes and one evaluated support of Science and Technology parks. We have extended our sample by eight more studies to have a more balanced selection for the analysis.
- The selected evaluation reports assessed three Incubator/Innovation centre support programmes (Swedish National Incubator Programme, BITS Incubator Programme in Australia and Incubator Support Programme in New Zealand), three Science and Technology park support programmes (Science and Technology parks in Poland, Kent Science parks and West of Scotland Science Park) and five Technology transfer programmes (Canadian Initiative for International Technology Transfer, Regional Office Technology Transfer Programme and Knowledge transfer programmes funded through the science budget in the UK, Support for Technology Transfer in Germany, and TechnoKontakte in Austria).
- The authors of these evaluations are in more than a half of the cases consultancies[23] (or consortiums including consultancy firms), two of them were carried out by research institutes with a department specialising in public policy, one by a ministry and two reports did not specify the authors.
- The budget and duration of the evaluations tend not to be disclosed but a high number and breadth of evaluation questions and evaluations tools used suggest that some of them were substantially large studies. Those evaluation reports that disclosed information about the duration of the studies suggest that they ranged from eight months to a year. Most of the reports in the sample are mid-term evaluations from the period 2003-2009 with one report from the early 1990s.

2.2.2 Rationale and objectives

The schemes that were evaluated in these reports had a variety of innovation-related functions.

Two of the incubator support programmes within our sample had the following aims.

1. The Swedish National Incubator Programme:
   - Increase the national capacity to take care of research based business ideas by establishing an infrastructure of "world-class" incubators i.e. a structural capital that Sweden can capitalise on moving forward;
   - Increase the national capability and skill in managing and supporting early-stage commercialisation of research based business ideas (i.e. establish incubation management as a "profession" and incubation as a managed process);
   - Generate high-quality start-ups to the financing and investment market
2. Incubator Support Programme:
   - The promotion of incubation best practice;

[23] In the case of Mid-term evaluation Of the Swedish National Incubator Programme, the evaluation was in fact done by an expert panel supported by a secretariat provided by the consultancy, Inno Germany AG.

- Encouraging networking among incubator managers and with organisations that are interested in incubation and incubated businesses; and
- Enhancing networking between incubators and universities/Research institutes.

Incubators themselves tend to provide early stage finance (grants, seed and venture capital funding), assist the incubated firms with business planning, financial advice, marketing, networking, mentoring, legal and accounting services, secretarial and other services.

The schemes that support Science and Technology parks do so to provide a location for the development of businesses focusing on emerging and future technologies.

The last sub-category is formed by evaluations of programmes for support of Technology Transfer that have a variety of focus areas. Some aim at fostering technology transfer between business enterprises and academic teams, some do so to increase technology transfer to small and medium enterprises, and some are even thematically bound. They are mainly delivered by Technology Transfer Agencies / Offices, through networks and other tools for stimulation of such activity.

When evaluating any one of these three sub-measures, one might expect two different kinds of effects:

- Direct outcomes and impacts linked to growth of firms through innovation, the knowledge and technology transfer;
- Indirect outcomes and impacts linked to the resulting increase of innovation projects amongst the beneficiaries participating in the network, being placed in an incubator / science and technology park or receiving funding from the technology transfer office.

The extent of transfer of knowledge and technology is often difficult to assess beyond the traditional output indicators because knowledge has a tacit element and the beneficiaries are not always willing to share their experiences with technology transfer, especially in instances when the technology is not in commercialisation development stage. Nevertheless these cases can be identified by questionnaire surveys and followed up further by in depth interviews.

Studies evaluating the activities of intermediary bodies (that facilitate technology transfer) can either be standalone evaluations or can form part of a broader innovation support programme that integrated activities of such bodies. In such broader evaluations, one might expect to test the global effect of the programme on beneficiaries (possibly through an analysis of the net effect of the measure and/ or through cost/benefit analysis), as well as the wider macro-economic effects at the regional/national levels. In both cases (standalone or boarder evaluations) can the performance and design of the support be benchmarked to similar institutions within or outside the region.

The following sections present and attempt to analyse in-depth how the innovation support activities of intermediary bodies facilitating technology transfer were assessed in the different reports within our sample.

2.2.3 The evaluation record

2.2.3.1 The purpose of the evaluations

The purpose of the evaluations in the mapping was to a large extent focusing on assessment of performance and impact of the schemes. They also provided feedback to the programme designers and implementers regarding the ability of the scheme to meet its objectives. In some cases this was done through a comparison to other similar schemes (i.e. benchmarking). This exercise was to identify the key design factors impacting on the success of the schemes and especially the impact of the programmes on society.

2.2.3.2 The evaluation questions

Figure 6 presents the scope of evaluation questions assessed in each of the eleven evaluations reviewed. It also provides indications on the scheme type.

Figure 6 Criteria used in the evaluation of Intermediary Bodies

Name of the measure evaluated	Type of measure	Type of evaluation	Relevance	Efficiency	Effectiveness	Sustainability
Incubator Support Programme (NZ)	Incubator/ Innovation centre	Mid-term		√	√	√
BITS Incubator programme and the Intelligent Island Incubator (AU)	Incubator/ Innovation centre	Ongoing		√		
Swedish national Incubator Programme (SE)	Incubator/ Innovation centre	Mid-term		√	√	√
West of Scotland Science Park (UK)	Sci & Tech park	Ongoing		√	√	√
Kent Science Park (UK)	Sci & Tech park	Ongoing				√
Science and Technology Parks (PL)	Sci & Tech park	Ongoing	√	√	√	√
Canadian Initiative for International Technology Transfer (CA)	Tech transfer	Mid-term	√	√	√	√
Regional Office Technology Transfer Programme (UK)	Tech transfer	Ongoing	√	√	√	√
Knowledge transfer programmes funded through the science budget (UK)	Tech transfer	Mid-term	√			√
Support for Technology Transfer (DE)	Tech transfer	Mid-term	√		√	√
TechnoKontakte (AT)	Tech transfer	Mid-term			√	√

Six out of the eleven reports are mid-term evaluations. The remaining 5 reports did not specify their type however most of them assessed a scheme's performance over a period of time (retrospectively) and was meant to answer questions about its future development, hence we have labelled these 'ongoing' evaluations.

The question whether the programme is meeting its formal objectives is addressed in almost all evaluation reports with exception of the report on the Kent Science Park. In some studies the answers to these questions are based on secondary research but generally a survey of or interviews with beneficiaries would serve as input to this section of the report.

What's common to all of the reviewed evaluations is a list of recommendations for improvement of the schemes, based on the study findings.

Consideration of the management and operational efficiency of the scheme was addressed by all incubator programmes, 2 out of 3 science park evaluations and less than a half of technology transfer programmes. On the other hand fundamental question of relevance of the scheme is addressed almost by all evaluations of technology transfer programmes, while evaluations of incubator programmes and science parks tend to omit this question overall.

Question of effectiveness was in the evaluations deal with from two different angles:

- How effective was the scheme in meeting its objectives
- What were the factors that impeded its performance, especially in terms of programme design (Canadian Initiative for International Technology Transfer (CIITT))

Secondly the scope of the effect of the scheme differed from case to case. Some of the evaluations focused primarily on the effect on the beneficiaries, other evaluations compared these effects with an "untreated" control group. Only rarely the link was established between the scheme and its effect on the economy.

Eight of eleven evaluations also included an analysis of the impacts of the programme, in fact the only ones that did not were the German Support for Technology Transfer programme, Canadian Initiative for International Technology Transfer and BITS Incubator programme and the Intelligent Island Incubator.

Evaluation of Technology Transfer scheme in Germany focused mainly on the perceived relevance and effectiveness of the scheme, based on a survey of innovation driven SMEs. On the other hand BITS Incubator programme and the Intelligent Island Incubator was heavily oriented towards future development of the programmes and the detailed recommendations included remarks on introduction of competitive process where incubators demonstrate the value added they can bring to incubatees and that basic incubator operating costs and funds earmarked for investment in incubatees should be separately identified.

2.2.3.3 The evaluation methods implemented

Figure 7 presents the evaluation methods used in the eleven evaluation reports.

Figure 7 Methods used in the evaluation of Intermediary Bodies

	Qualitative methods			Quantitative methods				Approaches		
Name of the measure evaluated	**Individual interviews**	**Case studies**	**Peer review**	**Questionnaire survey**	**Use of administrative/ secondary data**	**Micro-/ macro-economic model**	**Bibliometrics**	**Baseline**	**Counterfactual**	**Benchmarking**
Incubator Support Programme (NZ)	√			√	√			√	√	√
BITS Incubator programme and the Intelligent Island Incubator (AU)	√			√	√			√	√	√
Swedish National Incubator Programme (SE)	√	√	√		√					√
West of Scotland Science Park (UK)	√			√	√	√		√	√	√
Kent Science Park (UK)				√	√	√		√		
Science and Technology Parks (PL)				√	√			√	√	√
Canadian Initiative for International Technology Transfer (CA)	√	√		√					√	
Regional Office Technology Transfer Programme (UK)				√						√

Knowledge transfer programmes funded through the science budget (UK)		√		√	√			√	√	
Support for Technology Transfer (DE)	√									
TechnoKontakte (AT)	√			√						

The following methodological patterns are evident across the evaluation reports:

- Majority of the studies use a mix of qualitative and quantitative analysis with most of the time the implementation of several evaluation tools. This is especially true for reports evaluating incubator programmes in which Incubator Support Programme and BITS Incubator Programme applied more quantitative methods and the evaluation of the Swedish National Incubator Programme made use of an expert panel, case studies and interviews. In this case questionnaire survey was not conducted during within the study as the beneficiaries were surveyed within another evaluation. Evaluations of science parks and knowledge transfer visits in several instances included also site visits (e.g. Knowledge transfer programmes funded through the science budget).
- Nine out of eleven reports used questionnaire surveys and the reminding two used secondary data from surveys of beneficiaries/innovation driven SMEs. In some cases were the beneficiaries surveyed by structured interviews, and in other cases did the interviews provide additional information for assessment of counterfactual of the scheme. For example in Canadian Initiative for International Technology Transfer have the evaluators conducted a series of interviews with non-participating companies and used the data to assess the counterfactual of the support measure. Interviews with stakeholders have been used widely to determine future development of the scheme.
- The majority of surveys of beneficiaries collect opinions on performance of the programmes, its effects on the beneficiaries themselves. In two cases, the survey also aims at gathering quantitative data used afterwards in the micro-economic modelling – in the evaluation of the West Scotland Science Park and in the evaluation of Kent Science Parks, the surveys were specifically designed to support an econometric approach in order to collect GVA data, business growth (employment and sales) and sales per employee. In the latter case was this data used to model 3 different expansion scenarios through a multi-criteria analysis.
- All evaluations of incubators and science parks within our sample included a descriptive statistical analysis of the beneficiaries participating in the schemes – most of them presenting a sectoral breakdown of the businesses within the incubators / science and technology parks. Five of these schemes also assessed performance of the scheme against baseline targets. Similarly, the Interim evaluation of knowledge transfer programmes funded through the science budget contained such analysis.
- Over a half of the reports include a counterfactual approach, most of them, as mentioned, through surveys/interviews of beneficiaries and non beneficiaries and/ or through an economic analysis assessing the net effect of the measure(s) evaluated and the deadweight. The latter is true only for one report - evaluation of the West Scotland Science Park.
- One of the micro-economic model includes gross and net turnover, employment, and GVA contributions of the firms based in the parks. Another report is based on a multi-criteria analysis and estimates direct and indirect economic impacts for 3 scenarios based on combination of survey and analysis of secondary data of the geographic area. Broadly speaking, 4 of 8 reports make use of statistical analysis, using data from beneficiaries survey and/ or from administrative data.

- Six of the eleven evaluations benchmarked measured performance against a similar programme / intermediary. All three incubator evaluations used international comparators. The remaining three evaluations compared the schemes with the performance of other analogous initiatives within the country in question (e.g. with the Evaluation of the West of Scotland Science Park, the evaluators chose to compare this park with the performance of all other science parks within the portfolio of Scottish Enterprise).

As might be expected given the scope of the measures, none of the evaluations include bibliometrics. Some of the evaluations nevertheless posed IPR-related questions to their beneficiaries in the surveys.

2.2.4 The advantages and limits of the methods used

Most of the reports focus on the following aspects of the support scheme delivered through intermediary bodies:

- What are the direct and indirect economic impacts of the intermediary bodies and agencies?
- How does the scheme affect the survival/growth of enterprises and the use of the facilities provided?
- What is the value for money of the scheme?

Broadly speaking, these evaluations focus mainly on the economic outcomes and impacts of the programmes, calling for economic impact analysis. The sample of evaluation offers however a good mix of quantitative and qualitative data allowing for a straightforward approach and an in-depth cross-analysis.

Where the economic impacts of the scheme are questioned, a descriptive statistical analysis is used, based either on a counterfactual, or a baseline, or a benchmarking approach comparing the evaluated programme with the performances of similar programmes abroad. This often builds on desk research, the use of administrative data and the results of large-scale surveys among beneficiaries. Micro-economic and multicriteria statistical analysis would be expected in this respect but they were only rarely used in the sample of evaluations collected.

Where behavioural changes occurring as a result of the scheme are addressed, the sample of evaluation mainly makes use of individual interviews and large-scale survey to beneficiaries.

Future developments are interestingly approached in one evaluation through a multi-criteria analysis that estimates direct and indirect economic impacts for 3 scenarios based on combination of survey and analysis of secondary data of the geographic area.

2.2.5 Identification of good practice

As previously mentioned, most of the evaluations reviewed in this section involve a broad mix of qualitative and quantitative methods.

Figure 8 cross-tabulates each of the classic policy evaluation questions with the one of the eleven evaluation reports that we deem to be the best example of good practice.

Figure 8 Good practice in the evaluation of Intermediary Bodies

Objective of the evaluation	Evaluation Report	Good practice in the use of evaluation methods
Assessing the relevance/ appropriateness and future development of the measure	Interim evaluation of knowledge transfer programmes funded through the science budget (UK)	The study used a range of qualitative research methods (desk research, beneficiary interviews, site visits lasting 1-3 days) to answer the research questions.

Objective of the evaluation	Evaluation Report	Good practice in the use of evaluation methods
Efficiency, performance and impact assessment	Mid-term evaluation of the Swedish National Incubator Programme (SE)	The research was undertaken by an expert panel, through use of interviews with the incubator managers, universities and other stakeholders. Efficiency was compared to other European incubator programmes based on a survey.
Assessing the performance of incubators, effectiveness and efficiency of the scheme with respect to original objectives and targets	Evaluation of BITS Incubator programme and the Intelligent Island Incubator (AU)	The evaluation used baseline indicators to assess the performance effectiveness and efficiency of the BITS and III programmes. In one section there was a table ranking benchmarked incubators in terms of efficiency, effectiveness and utility.
Economic impact and future development	West of Scotland Science Park (WSSP) (UK)	The study assessed the economic impact at a local and national level. A range of information was collected using the monitoring system, stakeholder and questionnaire surveys and regional statistics to provide insights into deadweight, displacement, substitution, leakage, and multiplier effects. The answers were used to calculate the economic impact – or net effect - of WSSP.
Economic impact and future development	Kent Science Parks – Economic Impact Assessment (UK)	The good practice in this evaluation was use of scenarios to inform decisions on the possible future expansion of KSP. The direct and indirect economic impacts were estimated for 3 scenarios based on combination of survey and analysis of local data. The questionnaire asked respondents regarding origin of the incubated firms, salary, growth, cooperation with research and views of the science park.
The use of a counterfactual impact evaluation	Incubator Support Programme Evaluation (NZ)	The good practice in this evaluation was the rather varied methods used for cross-analysis, especially the use of two control groups (companies that were accepted for incubation but chose not to enter an incubator and incubated companies that left before graduation without the mutual agreement of the incubator and company) to assess to which extent the changes observed can be attributed to the programme reviewed.

Two evaluation reports are singled out for further elaboration as they present interesting methodological design and combination of methods:

3. **Incubator Support Programme Evaluation**, which is a high quality evaluation in terms of evaluation methods and approaches used, detailed hereafter:

- Evaluation tools used:
  - A file review of policy documents and New Zealand Trade and Enterprise (NZTE) records (NZTE is the national economic development agency of the NZ government, responsible for oversight of the incubator programme)
  - A literature review of incubation and international evaluations of incubator programmes

  - A survey questionnaire of exited companies from incubators supported under the programme. The design of the survey was informed by an open discussion with current and exited tenant companies
  - The use of industry benchmarks
  - On-site interviews with managers of incubators, both supported and not supported by the programme
  - Interviews with other stakeholders to the New Zealand incubation environment including the programme manager at NZTE, other NZTE decision makers, founding partners of incubators, commercialisation offices of universities, Crown Research Institutes (CRIs), and Incubators NZ (the industry association)

- Approaches used for the analysis and based on the previous mentioned tools:
  - To help ascertain the added value of the programme and to what extent it can be attributed the changes observed, the evaluators endeavoured to use two control groups: (1) companies that were accepted for incubation but chose not to enter an incubator and (2) incubated companies that left before graduation without the mutual agreement of the incubator and company. However, numbers for both control groups were statistically too small. Instead, we relied upon industry benchmarks on the growth and survival rates of New Zealand firms

4. The **Mid-term evaluation of the Swedish National Incubator Programme** is a good example of a study based mostly on qualitative research methods and secondary sources, analysed by an international expert panel

- Evaluation tools used:
  - Panel: The panel consisted of international experts with a proven track-record in incubation. The panel members and its chairman were appointed by VINNOVA (i.e. the Swedish Governmental Agency for Innovation Systems, managed under the Swedish Ministry of Enterprise, Energy and Communication)
  - In-depth interviews with incubator managers, representatives of universities, businesses and other stakeholder organisations. Additional interviews with regional stakeholders dealing with innovation policy and especially commercialisation of early stage R&D results
  - Collection and exploitation of statistic material available within Innovationsbron with regards to the incubators involved in the NIP and their performance
  - Collection and exploitation of material on the NIP provided by Innovationsbron and Vinnova
- Approaches used for the analysis and based on the previous mentioned tools:
  - Benchmarking: Comparison with European incubation programmes
  - Different case studies based on interviews with regional stakeholders dealing with innovation policy and especially with the commercialisation of early stage R&D results
  - Assessment of impact on incubator manager as an occupation, regional innovation environment, early stage commercialisation of R&D in the region

2.3 Start-ups and spin-offs

2.3.1 Introduction

'Start-ups' are young but established companies at an early stage of development, which are likely to be expending most of their effort on commercialisation, and may have little or no market presence or sales income. 'Spin-offs' is a term used to describe enterprises that split off from their parent organisation, usually a university or centre of excellence, taking with it licenses, intellectual property and / or technology developed in the parent institution.

In the market place, these firms will often be funded by early stage venture capital. Early stage venture capital involves the provision of finance in exchange for an equity stake in the young firm. This kind of investment goes far beyond a loan because it gives the venture capitalist an incentive to be involved in the actual running of the company and to help bring it to market.

Start-ups and spin-offs have a much lower success-rate than established firms and are accordingly much riskier to invest in. The reason that they are still an attractive prospect is because they tend to have skewed returns, with a minority of firms yielding very high profits. A typical early stage venture capital investor would expect to have a large portfolio of early stage firms, where only one or two will yield profits that will cover the investments in the rest of the portfolio. Success-stories of 'high growth' companies backed by early stage venture capitalists are easy to find, especially in the US. They include companies like Intel, Apple, eBay, Google and Genentech. There are also many important, although possibly less visible, success-stories in the field of medicine.[24] These high-growth firms are concentrated in a very few sectors that generate extremely high rates of return such as biotechnology and healthcare, information and communications technology and increasingly clean technologies.[25] These high-growth firms are believed to have a disproportionately large impact on the economy, despite the very small number of investments that are made each year. For example, in the US, only 3,000 firms get venture capital investment each year and of these only 500 of them are start-ups.[26]

2.3.2 Rationale and objectives

It is now generally accepted that young firms face financial constraints when they are in their early stages of development. This is usually explained as being because investors are not willing to shoulder the high risk that is involved in investing in start-ups and spin-offs. The higher risk can be attributed to the lack of collateral these companies have, the lack of track record they have due to being new companies and the high number of them that fail on average. Another explanation for the shortage of capital available for early-stage businesses is that they can require small amounts of capital (> €300,000), which makes the transactions costs surrounding managing the funds higher and can make the transactions less worthwhile for investors. In other words, these opportunities go below the investor's 'radar.' This is perceived to be a market failure if the product is high-tech or green and might have substantial 'spillover' effects for the rest of the economy or society.

---

[24] Early-stage Investment by the Estonian Development Fund: An appraisal of activities 2007-2009 and scenarios for future development, Paul Nightingale (SPRU, University of Sussex) Alasdair Reid (Technopolis Group), 2010.

[25] Ibid.

[26] From funding gaps to thin markets: UK Government support for early-stage venture capital, Paul Nightingale (University of Sussex), Gordon Murray (Exeter University), Marc Cowling (Institute for Employment Studies), Charles Baden-Fuller (City University), Colin Mason (University of Strathclyde), Josh Siepel (University of Sussex), Mike Hopkins (University of Sussex) and Charles Dannreuther (Leeds University), BVCA, NESTA, 2009.

Because the market for early stage venture capital is so delicate, almost every major economy in the world has attempted to find some way to support it. Most of these attempts are based on the experience given by the US. This is because the US has over 50 years of experience in developing policies for supporting early stage businesses, and because venture capital funds in the US have been able to produce such disproportionate profits on a scale, which has not, been successfully replicated in other countries.

There are various ways that governments can support the creation of start-ups and spin-offs. They can provide grants to university 'spin-offs' in order to facilitate the commercialisation of licenses, technology or intellectual property that has been developed in the parent institution. This helps to transfer knowledge out of the knowledge-producing institution and to encourage the formation of new businesses.

The government can also fund established 'start-ups' through venture capital measures. It is now generally agreed that the best way to achieve this is through a mixture of public and private funding, often referred to as 'hybrid' venture capital. This type of funding instrument has several advantages. Firstly, it lowers the risk of exposure to the private investor and thus encourages them into the market. Secondly, the private investor is tasked with selecting the appropriate firms to invest in, which is an advantage given the government's notorious inability to make good investment decisions. Thirdly, it means that the private investor's exposure is still high enough that they have an incentive to manage the new firms and transfer their experience and expertise.

The way that governments finance their share of the investment varies from fund to fund. The investment can either be in the form of a grant, such as in the Scottish Enterprise Proof of Concept Programme, which funds the development of the product (but the product should be at the stage where it has already been patented). The fund also puts the new enterprises in touch with experienced venture capitalists who continue to nurture them. Another way is for the government to finance the firm either through debt or equity or both (mezzanine funding). Another way is for the government to guarantee some proportion of the finance made by the private investor in order to lower their risk.

2.3.3 The evaluation record

2.3.3.1 Results of the mapping

We have selected eight evaluations of funds that provide seed and early stage capital for entrepreneurs:

- The Scottish Enterprise **Proof of Concept Programme** is typically awarded after curiosity-driven or strategic research has been developed to the point where a patent has been filed. It is not just another form of research funding, but designed to attract exactly the sort of project that bring super-profits that venture capital is looking for and putting the projects in contact with venture capitalists who will nurture the projects until they reach the market. The Proof of Concept Programme has been enormously successful in finding successful projects and is now used as a best practise model for many public funds around the world
- **EXIST-SEED** funds academics to turn university-based research into their own spin-off enterprises
- The evaluation of ERDF Supported Venture Capital and Loan Funds is an evaluation of 9 different funds including **Sigma Sustainable Energies Fund**, which is a managed fund that invests public and private funding in a mixture of debt and equity packages. The fund invests in companies based in the East of Scotland involved in novel renewable and sustainable technologies; the **Prince's Scottish Youth Business Trust** provides seedcorn finance and professional support to young people generally aged between 18-25 and **Genomia**, which provides seed-funding to support emerging technologies and to help bring them to

market in the area of life science research from institutes such as the Institute for Animal Health, Moredun Research Institute, Roslin Institute, Rowett Research Institute, and the Scottish Agriculture College

- The early assessment of BIS Equity Fund Initiatives is an evaluation of four fund initiatives, all of which provide equity funding to early stage and established businesses, including the **Enterprise Capital funds**, which provide equity finance to early stage and established businesses including hi-technology businesses and the **Finance South East Accelerator Fund**, which provides debt and mezzanine finance to smaller-scale early stage and established businesses
- The Evaluation of **Community Development Finance Institutions** look at not-for-profit intermediary organisations that 'on-sell' public venture and equity and loan funds to private businesses and recover their operating costs from the government schemes. They are the independent 'operational/delivery' arm of many of the Government's various investment funds and schemes
- The report 'From funding gaps to thin markets: UK Government support for early-stage venture capital ' is a piece of independent research for NESTA which looks at the relative performance of all **early stage venture capital funding in the UK** and situates them in policy context of the UK
- The external evaluation of the pilot scheme **CREA** is an evaluation of the European scheme that funds organisations engaged in early stage venture capital
- The **Finnish Industry Investment Ltd** is a publicly-owned Finnish bank set up to provide corner-stone funding for early stage venture capital enterprises, where the majority of the funding is expected to come from the private sector

2.3.3.2 The questions addressed by the evaluations

Figure 9 Criteria used in the evaluation of Start-ups and Spin-offs

Name of the measure evaluated	Type of measure	Relevance	Efficiency	Effective-ness	Sustainability
Scottish Enterprise Proof of Concept Programme Evaluation – Rounds I to VI: Final Report (UK)	Grants for Proof of concept and commercialisation			√	√
Effects of EXIST from the perspective of grantees (DE)	Grants to fund university spin-offs		√	√	
Evaluation of ERDF Supported Venture Capital and Loan Funds (UK)	Venture Capital and Loans	√	√	√	√
Early assessment of the ~ Impact of BIS Equity Fund Initiatives (UK)	Venture Capital and Loans	√		√	
Evaluation of Community Development Finance Institutions (CDFIs) (UK)	Venture Capital and Loans	√	√	√	√
From funding gaps to thin markets: UK Government support for early-stage venture capital (UK)	Venture Capital	√	√	√	
External Evaluation of the Pilot Scheme CREA Concerning Support for Venture Capital Companies Financing SMEs in the Seed and Start-up Phase: Final Report (EU)	Venture Capital	√	√	√	√

Finnish Industry Investment Ltd: An International Evaluation (FI)	Venture Capital	√	√	√	√

Within our sample, six out of the eight evaluations explored the issue of **relevance** to some extent, and only two did not mention it at all. The fact that so many of these evaluations were interested in uncovering whether the intervention was appropriate may reflect how close to market the interventions were. For example, it could be because the funds were intended to stimulate the supply of private venture capital, and if successful the need for the funds should subside. In several cases, such as the evaluation of ERDF supported Venture Capital and Loan Funds it was because the funds being evaluated existed in a landscape where many similar or overlapping early stage and established business support funds existed, so the evaluator needed to map existing provision of finance for SMEs (by stage of development, type of finance, size of firm, etc.) to determine whether the scheme or schemes under evaluation partly or fully addressed identified 'gaps'. In the two cases where relevance was not an issue, one was the Scottish Enterprise Proof of Concept Programme, which was highly successful at attracting profitable opportunities and putting them in touch with venture capitalists, so may not have felt the need to examine whether the scheme was still relevant. In the case of the EXIST-SEED programme, it may have been assumed that the need to give funding toward university spin-offs is ongoing and thus does not need to be reviewed.

Six out of the eight evaluations tackled questions of **efficiency** to some degree, although what was meant by the term varied. The schemes providing financing to spin-off and early stage companies in most cases looked at the efficiency (or economic rate of return) of the scheme as a whole, in terms of the extent to which the investments achieved a net positive return in terms of calculated economic impacts. Only two of the evaluations considered the operational (or management) efficiency by calculating the administrative costs of running the scheme and comparing this to the scale of the investments being made.

All of the evaluations considered **effectiveness** within the scope of their assessment, which could be seen as the primary objective of all the evaluations. However, it can be noted that in most cases the specific objectives of the programme being evaluated were not clearly spelled out, so in many cases schemes were considered to have been effective if they generated the kinds of impacts expected. In this regard the distinction between effectiveness and impacts were somewhat blurred. None of the evaluations reviewed here were of programmes that could be seen to have SMART objectives (Specific, Measurable, Achievable, Realistic, Time-bound). This could be a feature of early-stage capital investments because they are known to give notoriously skewed returns, with most projects failing and some giving very high returns. As a result of returns being hard to predict, SMART objectives may be difficult to set.

Five of the evaluations considered future **sustainability** as part of the scope of their work. This was where the schemes provided financing to early-stage business (equity, loans, etc.) and the question of sustainability revolved around whether the funds that had been set up were, or could be expected to become, self-sustaining. That is, the evaluations were asked to determine whether the income being generated by the funds (through equity income and loan repayments) was sufficient for the schemes to continue to operate in the absence of continued public support.

Almost all of the evaluations also sought to identify and in many cases quantify the **impacts** that the programmes have generated. However, once again there was a great deal of variability in the extent to which impact was explored, even if all of them are based on a counterfactual impact evaluation. Three of the evaluations used full gross to net calculations that sought to determine the net (additional) effects by allowing for deadweight and displacement, etc., while one evaluation simply asked the beneficiaries of the funds whether they would have gone ahead with the funding anyway. The evaluation 'From funding gaps to thin markets: UK Government support for early-stage venture capital', was a meta analysis (desk study) that used

econometric techniques to quantify the impact of venture capital support, comparing the commercial performance (impact of financial support on subsequent employment growth) for almost 800 client businesses (using micro-economic data provided by six venture capital funds) with the equivalent performance for almost 8,000 unsupported businesses (a matched control sample)

In terms of other issues addressed by the evaluations, five looked at the **operational management** of the schemes, and four investigated the extent to which the schemes have achieved a high level of **customer satisfaction**.

2.3.3.3 The evaluation methods employed

Figure 10 Methods used in the evaluation of Start-ups and Spin-offs

	Qualitative methods			Quantitative methods				Approaches			
Name of the measure evaluated	**Individual interviews**	**Case studies**	**Peer review**	**Questionnaire survey**	**Use of administrative/ secondary data**	**Micro-/ macro-economic model**	**Bibliometrics**	**Baseline**	**Counterfactual**	**Benchmarking**	**Indicators**
Scottish Enterprise Proof of Concept Programme Evaluation – Rounds I to VI: Final Report (UK)	√			√							
Effects of EXIST from the perspective of grantees (DE)				√							
Evaluation of ERDF Supported Venture Capital and Loan Funds (UK)	√			√	√			√	√	√	√
Early assessment of the ~ Impact of BIS Equity Fund Initiatives (UK)	√	√		√							
Evaluation of Community Development Finance Institutions (CDFIs) (UK)	√	√		√	√			√	√	√	√
From funding gaps to thin markets: UK Government support for early-stage venture capital (UK)						√		√	√		
External Evaluation of the Pilot Scheme CREA Concerning Support for Venture Capital Companies Financing SMEs in the Seed and Start-up Phase: Final Report (EU)	√			√					√		
Finnish Industry Investment Ltd: An International Evaluation (FI)	√			√							

Given the broad range of intervention schemes as well the broad range of beneficiaries, the tools used for analysis vary considerably:

- Seven of the eight evaluations used **questionnaire surveys** and in fact this was the principal method employed to gather the data and information with which to evaluate the schemes. In most cases the surveys were used in addition to the documentation and reports supplied by the funds, and used to establish the

answer to the majority of the evaluation questions, such as how successful the projects funded had been, how much revenue they had accrued, whether the beneficiaries could have applied for funding elsewhere or would have pursued their projects without funding, etc. In only one case was questionnaire surveys not used to establish the majority of the information needed for the evaluation, and that was in 'From funding gaps to thin markets: UK Government support for early-stage venture capital', which was an econometric study.

- Six of the eight evaluations used **individual interviews** as part of their methodology; this was always done in conjunction with a questionnaire survey that formed the bulk of the evidence. Individual interviews were used to add depth to their understanding of the situation and to speak to key stakeholders with an interest in the operation and outcomes of the schemes under evaluation. For example, key players in the private investment sector were spoken to in order to validate and extend understanding of gaps in the equity markets that the schemes were designed to address.
- **Descriptive statistics** were used in only two of the eight evaluations. These were used to profile the range of beneficiaries supported and the nature and scale of the assistance provided. There did not appear to any particular pattern to when such a method was appropriate, implying that they could have been employed in any or all of the evaluations. Indeed, it would appear that the nature and extent of the questions addressed by the study and the ranges of methodologies used were driven by the scale and ambition-level of the evaluations carried out, rather than by the nature of the schemes under evaluation.
- **Case studies** were used in two of the evaluations, but these were used simply to add 'colour' by providing qualitative descriptions of some of the 'projects' or businesses supported, rather than forming part of the analysis.
- One of the evaluation reports sought to quantify the commercial impact of venture capital through consideration of the performance of 780 firms supported by six government-backed VC-funds. The study made use of **econometric analyses** to compare the performance of assisted firms with non-assisted firms in the wider economy and concluded that all schemes had produced a positive impact on aggregate client performance (albeit quite modest) when compared with the control group. This evaluation was an attempt to understand whether the rationale for intervention that underlines early stage venture capital was in fact correct. To test this assumption a range of recent, similar early stage venture capital funds were examined for profitability.
- None of the evaluations used **peer review** or **bibliometrics**, which is appropriate given the nature of the support schemes under consideration here.

2.3.4 The advantages and limits of the methods used

None of the evaluations analysed here made explicit either the budget of the evaluation or the time frame. However, the costs did appear to vary considerably depending on the approach taken. For example, while almost all the evaluations used questionnaire surveys, it is clear that the costs associated with these will have depended greatly on the size of the samples, the extent of the data collected, and whether a control group was used. There were also significant variations in the extent to which quantitative data were collected through the surveys, and in the extent to which complex analyses were performed on the data.

Questionnaire surveys were the most common tool used for data collection and showed the greatest flexibility in being able to collect data suitable for the questions being addressed by the evaluation. The questions usually surrounded asking the beneficiaries what they had used the funds for and whether the objectives of the grant or financing had been realised. Where the counterfactual was approached, respondents were asked whether they had sought alternative forms of assistance and whether they would have undertaken their project in the absence of funding, in order

to control for 'deadweight'. In some cases the questions also sought to control for displacement effects by identifying the proportion of additional turnover realised at the expense of other firms in the region. In a small number of cases control groups have also been used to control for background variables (e.g. more general changes in turnover and employment) within the populations of businesses targeted by the support. Clearly the studies collecting more extensive bodies of quantitative and qualitative data from beneficiaries were able to answer a broader set of questions and in a more comprehensive and robust way.

The only real alternative to questionnaire surveys as a method of data collection was in one of the reports (From funding gaps to thin markets: UK Government support for early-stage venture capital), which used the annual **financial statements and balance sheets** for the SMEs funded by a range of early stage venture capital schemes. The data started before the SMEs received funding. Control groups were selected on a basis of ten control-group SMEs to one treated sample SME in order to increase the sample size to 7,741. The control-group SMEs were matched to the treated firms by sector (using Standard Industrial Classification codes) and by matching age and employment levels. **Econometric analysis** was then used to analyse the data. This method was interesting for several reasons: firstly it allowed the companies to be assessed in the period before they received funding, to test for whether there had been a selection bias. This revealed that the funding had had a real effect on the growth of the firms. Secondly, it allowed the evaluators to conclude that because the returns to funding had been modest that the market failure could not just be a short supply of venture capital. If there had just been a short supply of venture capital, then the entrepreneurial firms should have displayed the expected super-profits used to justify the funding. Instead, they argued that there was a 'thin market', which means that there are entrepreneurial growth firms, but spread very thinly, just as there are investors willing to shoulder great risk, but these investors can be hard to find. The evaluators argued that the 'market failure' is not that there is a shortage in the supply or demand for venture capital, but that because the markets are 'thin', they find it hard to find each other. This would also explain why the Scottish Enterprise Proof of Concept Programme was successful: it linked entrepreneurial growth firms to potential investors.

Descriptive statistics was put to good effect in some of the evaluations. In these cases the analyses were used to profile the range of programme investments and types of beneficiaries, often helping to provide a set of background variables that were used in the analysis of impact data. This permitted the evaluations to compare and contrast the relative costs and benefits realised within different target groups and through different forms of support. By 'benchmarking' a range of schemes against each other a small number of the evaluations were able to reach conclusions concerning the effectiveness and efficiency of alternative policy instruments, and to base their recommendations around these.

Many of the evaluations made use of **individual interviews.** In most cases interviews were employed to gain input and feedback from a wider group of stakeholders than just the direct beneficiaries. For example, by speaking to experts from the private investment sector, the evaluators could better understand the degree to which the public support schemes complement or compete with private initiatives, the extent to which real market failures were being addressed, and whether improved targeting of the support was necessary in light of changes within the market.

A minority of the schemes used **case studies** to help to exemplify the nature of the support provided, the beneficiaries being targeted and the kinds of impacts being generated. Such cases studies were usually carried out in small numbers and were used to add 'colour' to the reports by focusing on 'successful' or 'best practice' projects. We did not identify any examples where case studies were used to highlight specific issues or problems.

Two of the evaluations made good use of **benchmarking**. These evaluations reviewed a range of schemes and benchmarked the performance of each scheme against the

others, rather than comparing performance with other interventions in other regions or against international best practice. Generally, the latter would not be possible because these funds tend to be unique in terms of the exact financial instrument used, the type of beneficiary and the terms of which the support is offered. What these reviews do achieve is a good sense of what financial risk capital funds landscape looks like and where each scheme fits into this landscape. This enables the evaluation to identify areas where there is poor coverage and where public interventions could be most legitimately and most effectively deployed.

2.3.5 Identification of good practice

Figure 11 summarises our findings, i.e. displays a few examples of best practice for the evaluation of schemes to support start-ups and spin-offs.

Figure 11 Good practice in the evaluation of support to Start-ups and Spin-offs

Objective of the evaluation	Example	Best practice in the use of evaluation method
Assessing the relevance of the financial risk funds	Evaluation of ERDF Supported Venture Capital and Loan Funds (UK)	Benchmarking (comparison of funding schemes with others in the same evaluation)
Assessing the efficiency of the financial risk funds	Evaluation of Community Development Finance Institutions (CDFIs) (UK)	Calculated the internal rate of return, comparing the success rate to the operating costs
Assessing the impacts of the financial risk funds	Scottish Enterprise Proof of Concept Programme Evaluation – Rounds I to VI: Final Report (UK)	Use of a large-scale questionnaire survey (telephone) to quantify commercial benefits (in terms of income and employment) and estimate the extent to which those benefits were wholly or in part attributable to the support and would not have been realised in the absence of the support. The study also included selected interviews and case studies in order to estimate any displacement effects
Assessing the future development of the financial risk funds	From funding gaps to thin markets: UK Government support for early-stage venture capital (UK)	Use of econometric techniques to estimate the net effect of six government-backed VC funds on the growth of client businesses compared with a large control group

By way of conclusion, we found that all of these evaluations had some good points to report on. However, some of them had points that we have chosen to highlight as examples of best practice:

- The **Evaluation of ERDF Supported Venture Capital and Loan Funds (UK)** had a strong method for analysing the relevance of the schemes covered by the study. It did this by mapping the landscape of venture capital and loan funds (both public and private) and locating each of the ERDF supported funds within that landscape. The benchmarking encompassed factors such as the nature and scale of the financing provided and the types of beneficiary (size, age, stage of development), and allowed the evaluators to identify how the ERDF supported schemes related to other pre-existing or new interventions on the market. The evaluators successfully identified gaps in provision in some areas and overlaps in others, thereby improving the potential for subsequent schemes to be more precisely focused on established (i.e. proven) areas of market failure. In this way the 'relevance' of future interventions could be better assured and the 'net' benefits increased.
- The **Evaluation of Community Development Finance Institutions (CDFIs) (UK)** made good use of the UK's Impact Evaluation Framework (IEF)

methodology for calculating the net effect of the programme, and compared the results to the operating costs of the schemes in order to calculate the relative efficiency of a range of different types of support targeted towards different sets of beneficiary. In general the analyses showed that schemes offering smaller amounts of finance to each beneficiary had much higher operating costs than those providing larger investments, leading to the conclusion that the smaller schemes were much less likely to become self-sustaining in the future.

- The **Scottish Enterprise Proof of Concept Programme Evaluation (UK)** employed a good practice example for calculating the net impacts of the support scheme on the target businesses. The evaluation followed the UK's Impact Evaluation Framework (IEF) methodology, allowing the evaluators to provide a robust assessment of the (net) additional economic impact of the scheme by allowing for deadweight, displacement effects and spillovers. Application of the methodology permitted the evaluators to calculate the 'rate of return' for each € (or £ in this case) invested in the support scheme, and used the (positive) outcome to confirm the scheme's effectiveness.
- The **Impact evaluation of grants for SME modernization in the framework of National Development Plan 2004 – 2006 (UK)** made use of econometrics to assess the effectiveness that the fund had on the growth potential of the beneficiaries. It did this by using the accounts collected from the Ministry for Justice and Law enforcement, which meant it, had reliable data for both the beneficiaries and a control group for the full term the programme (three years). This enabled the evaluators to use a difference-in-difference technique, which did not require the use of control variables and kept the analysis very simple. The results showed that the majority of beneficiaries had already been growing before they received the funds, suggesting some selection bias and that the control group was very likely to make similar improvements and gains in the absence of support, suggesting a high level of deadweight.
- The report **'From funding gaps to thin markets: UK Government support for early-stage venture capital' (UK)** made use of a strong economic conceptual model and tested this using econometrics. The conceptual model was that the reason entrepreneurial growth firms struggle to find funding is not because there is a shortage in the supply of venture capital or a lack of 'good ideas' that lie behind growth firms (a lack of demand for venture capital), but that instead there was a sufficient supply of both but they were spread 'thinly' across the country making it hard to find each other. This theory was tested econometrically against a control group, which confirmed that increasing the supply of capital did have a positive effect on the treated firms, but that the firms were not yielding the expected super-profits that were needed to justify the funding. This confirmed their theory that the problem was more complex than anticipated by policy-makers and enabled them to suggest that future-funding programmes should be more focused on increasing networks between entrepreneurial growth firms and venture capitalists, and less on simply increasing the supply of capital.

2.4 Science-industry cooperation

2.4.1 Introduction

Eleven of the 120 or so evaluation reports in our repository address measures that are uniquely concerned with the creation or improvement of science-industry cooperation. Improving science-industry cooperation is often a subsidiary objective shared by many more of the 120 evaluation reports identified, however the methodological descriptions tend not to disaggregate their choice of methods and analytical

techniques in such a way as to permit one to infer any relationship between the study design and secondary objectives.[27]

Given their scope, these evaluations were mostly carried out on behalf of national agencies, however in most cases the reports indicate that the schemes have benefited to some degree from ERDF funding (except for Norwegian and USA schemes). However, the Estonian evaluation of Competence Centres is the only report to more formally deal with the co-investment from the structural funds. Appendix 1 lists the eleven evaluation reports considered here.

2.4.2 Rationale and objectives

The very great majority of EU member states and regions have chosen to implement measures to strengthen the links between science and industry, in order to improve the innovativeness of their business communities and increase the social return from public investments in science. These university-industry measures come in many guises – reflecting the diversity of types of interactions through which universities and businesses derive value – however our review identified 10 published evaluations, which split roughly into two sub-groups:

- **Collaborative research centres**, located at universities in the great majority of cases with a focus on a highly particular and strategically important area of applied research and close involvement of leading businesses in governing bodies and strategy definition.[28] These kinds of centres create numerous forms of interaction and exchange, but arguably have their most critical impact through shaping research and researcher education (user-orientation) and securing or extending the global intellectual networks of all parties. Two types of centres are to be distinguished:
  - Centres of competence (CCs), combining researchers in a HEIs (and/or sometime a research institute) and a consortium of industrials on areas of high innovation potential, with a view drive university's resources and strategy29
  - Centres of excellence (CoEs), operating often within universities and intended to build a critical mass of competitive world-standard research with a view to foster high-level of international visibility.30 Those are focused primarily on

[27] Science-industry cooperation is also often assessed while running evaluation of broader programme and agencies – e.g. the Value for Money Review of Science Foundation Ireland that review the Centres for Science, Engineering and Technology (CSET). However, measures in favour of science-industry cooperation are often assessed together with other measures in such reports and the level of detail offered does not allow one to disentangle the study design and link a methodological strategy with a particular innovation objective.

[28] Measures to foster longer-term cooperation have been particularly widespread in the portfolio of innovation public policies over recent years. In this respect, the US Engineering Research Centers programme is one of the oldest CC programmes - initiated in 1985 - and it has influenced the programme design in other countries. Programmes for competence centres/centres of excellence are long-term but often set for a limited period of time (e.g. maximum six years for the Finnish centres of excellence with funding received in two instalments at the beginning and after three years; five years in the Irish competence centres). The size of such centres varies according to the focus of the centre (e.g. Finnish Centres of Excellence are composed of research units with about 20 to 200 staff). They are either created as a distinct legal entity (Austrian model) or integrated in universities (Swedish Model). They are implemented most of the time under the form of national programmes, most of the time with a regional dimension (e.g. in Austria the Bundeslanders (provinces) provide additional funding for the COMET centres of competences and can take part to the decision-making process).

[29] Arnold, E., Deuten, J., & Van Giessel, J. (2004). An international review of Competence Centre Programmes. Technopolis Group.

[30] Action for "centres of excellence" with a European dimension, online: ec.europa.eu/research/era/pdf/centres.pdf

research excellence and are less industry-driven than CCs, although they often entail objectives related to science-industry cooperation.

In most cases, they will secure core funding from government with business co-financing of selected people or work. Most will be awarded longer-term grants of between five and 10 years' duration, with values in the millions of Euros. Most will be allowed to compete to secure a second term, subject to their successful and demonstrable historical performance and the continuing relevance of their strategic focus (judged against competing offers, new and old). Most of these centres will cease to be eligible for support after two terms, as a matter of principle.

- **Collaborative and knowledge exchange research projects**, with a smaller timescale and scope
  - Collaborative research projects involve one or more businesses will work together with one or more public research institutions on a specific research and development project the outcome of which has intrinsic commercial value to the private sector partners. These measures are typically co-financed by public grants of 3-5 year's duration that essentially cover the cost of the university or public research institute, while private interests pay their own costs of participation. Public investments will typically be in the many hundreds of thousands of Euros
  - Knowledge exchange projects comprise a miscellany of measures to support a specific innovation project, ranging from industrial placements to the co-financing of the private procurement of technical support services. These measures are typically very much smaller in scale. Schemes might provide a grant to offset the cost of finding and employing a recent STEM graduate to join the staff of a SME for 12-24 months, materially expanding its internal (absorptive) capacity, in order to drive forward a business development project under the supervision of both company and university. Elsewhere innovation voucher schemes have sprung up to provide SMEs with small credits – a few thousands of Euros – that can be used with any eligible knowledge partner within a 12-24 month period to purchase access to specialist facilities or know how. The smallness of the awards and the thematic and time-based flexibility makes such grants attractive to a large fraction of that group of national or regional SMEs that are innovation-active but have limited knowledge of or contact with the public sector research base

The rationale behind innovation measures in support of science-industry cooperation is based on a series of various expected effects and underlying hypotheses, as follows:

- Overcoming the widespread information and behavioural barriers to the cooperation between the public and private sectors, which greatly limit the frequency and intensity of productive interaction between the communities
- Developing stronger channels to facilitate the flow of knowledge and technology from public research organisations to public and private enterprises in a position to deploy that know how and IP in a commercial setting increasing social returns
- Conducting problem-focused research (as opposed to purely disciplinary academic research), in the expectation that this might expand the total academic effort devoted to user-oriented research and thereby accelerate technological breakthroughs in key areas
- Developing regional capacity (a critical mass of research excellence) in emerging areas adjudged to have strategic potential as an innovation platform for resident businesses, with strengthening national or regional competitiveness globally

Some of the above-mentioned objectives are shared with other innovation support measures studied in this literature review – e.g. support to Networks and Clusters.

2.4.3 The evaluation record

2.4.3.1 The evaluation questions

Figure 12 presents the scope of evaluation questions assessed in each of the eleven evaluations reviewed. It also splits the set into two sub-groups, with the first nine measures highlighted being university-based centres of excellence (CoEs) of one kind or another, and the remaining three being measures to support research and innovation projects.

Figure 12 Criteria used in the evaluation of Science-industry cooperation

Name of the measure evaluated	Type	Relevance	Efficiency	Effectiveness	Sustainability
K plus and Kind/net competence centres (AT)	CC	√		√	√
Christian Doppler Research Association (CDG) (AT)	CoE	√		√	√
Competence Centre Programme (EE)	CC	√	√	√	√
Centre of expertise Programme (FI)	CC	√	√	√	√
Finnish Programme for Centres of Excellence in Research (FI)	CoE	√	√	√	√
Berzelii centres (SE)	CoE			√	√
Institute Excellence Centres Programme (SE)	CoE			√	√
Engineering Research Centres (US)	CC			√	√
Innovation Consortium Scheme (DK)	Collaborative project			√	
Knowledge Transfer Partnerships (UK)	Knowledge transfer project		√	√	√
Dutch Innovation voucher (NL)	Knowledge transfer project			√	

While the eleven reports comprise a mixture of ex post and interim evaluations, the simple analysis in Figure 12 shows a clear pattern concerning the evaluation questions being addressed to the centres of excellence and competence centres. All of them look into effectiveness and future development issues:

- Evaluations of COEs and CCs focus on the impacts in terms of expanded research capacity, research excellence and innovation results within the Centres;
- While evaluations of collaborative and knowledge transfer projects put the focus rather on the impacts of the measures on participating enterprises (e.g. Dutch Innovation Voucher) and sometimes on the wider regional business community. (e.g. US ERC)

Efficiency and relevance questions are particularly prominent in those programmes launched in the recent past (e.g. the Mid-term evaluation of the Estonian Centres of competences) with a view to assess if the governance and functioning of the programme is efficient.

As far as the science-industry collaborative and knowledge transfer projects are concerned, the small number of measures (three) and their variety militates against any firm conclusions, however effectiveness is, again, the main focus.

Overall, only three out of eleven reports are ex-post evaluations. The three evaluations concerned mainly focus on the impacts of the programmes, either for the beneficiaries or for the society as a whole.

2.4.3.2 The evaluation methods implemented

Figure 13 presents the evaluation methods used in every report.

Figure 13 Methods used in the evaluation of Science-industry cooperation

	Qualitative methods			Quantitative methods				Approaches			
Name of the measure evaluated	**Individual interviews**	**Case studies**	**Peer review**	**Questionnaire survey**	**Use of administrative/ secondary data**	**Micro-/ macro-economic model**	**Bibliometrics**	**Baseline**			
K plus and Kind/net competence centres (AT)	√	√	√								
Christian Doppler Research Association (CDG) (AT)	√	√	√								
Competence Centre Programme (EE)	√	√	√		√					√	
Centre of expertise Programme (FI)	√	√		√	√					√	
Finnish Programme for Centres of Excellence in Research (FI)	√	√			√						
Berzelii centres (SE)	√			√							
Institute Excellence Centres Programme (SE)	√	√	√	√	√					√	√
Engineering Research Centres (US)	√	√		√	√					√	
Innovation Consortium Scheme (DK)		√			√	√		√	√		
Knowledge Transfer Partnerships (UK)	√				√				√		
Dutch Innovation voucher (NL)	√		√	√				√			

Figure 13 displays the following methodological patterns:

- 11 of 11 include a individual interviews, addressing questions about relevance in the main but also programme efficiency and added value (within the wider landscape) and programme effectiveness (to a lesser degree). The Danish IC scheme evaluation was the exception to this otherwise universal deployment, with its sharp focus on what had been achieved and the application of secondary methods (baseline, counterfactual analysis, econometrics) to data from the monitoring system (inputs, outputs). The focus on a rigorous quantification of

wider benefits appears to have been a choice of the commissioning agency driven by wider public management expectations, rather than a reflection of any intrinsic feature of the innovation measure in question.

- 8 of 11 include case studies, and especially the evaluations of the centres of excellence. However, these were almost always deployed as rounded, mini-evaluations – focusing on an individual centre – which involved multiple data collection methods and addressed most if not all of the evaluation questions. In two instances, the case studies were more narrowly directed to illustrating programme effects.
- 7 of the 11 include composition analyses to document the investments of the measure in question, as regards the volume and direction of funding and the composition of supported activities and resulting outputs. For the most part, this analysis is used as background, scene-setting to help the reader better understand the subsequent and more critical evaluation chapters, as well as the programme logic (e.g. Danish Innovation Consortium), as regards the location and composition of demand for support, and also as a reference for programme effectiveness, again, cross-checking the balance of investments and outputs with the type and mix of outcomes being sought
- 5 of the 11 include peer review, although in the case of the Dutch Innovation Voucher scheme, it was a high-level group of academics and industrialists with broad knowledge of national and international measures to improve science industry cooperation, and not the classic panel of international peers. The centres of excellence did follow convention more closely, with panels of internationally renowned domain experts being invited to judge the quality of the centres and their wider achievements in comparison with a virtual international benchmark. In several instances, the panels included eminent industrialists – representatives of the research users – and in this regard their hybridisation differs somewhat from the more traditional academic model. Almost all of the evaluations supported the peer review process with wide-ranging desk studies and self-assessments to facilitate the process of international scrutiny on the one hand and to help cope with the particularity and added value of these relatively small interventions within much larger academic settings

Each of the other methods was used in a minority of the 11 evaluations, and specifically, in descending order of frequency of use:

- 5 of 11 used a questionnaire survey, mainly to canvass the opinions of participants on the effects, efficiency and relevance of the programme evaluated. Questionnaire surveys are used whenever a large number of participants are involved in the programme. The evaluation of the Austrian K plus and K ind/net competence centres use questionnaire surveys mainly to gather data on a set of indicators on the functioning and results of the structures, while the Finnish evaluation of CoEs target implemented a series of four surveys addressed participants and non-participants in the programme in order to collect their view on the relevance and future development on the concept of CoEs. Questionnaire surveys are always used in complement to in-depth interviews.
- 4 of 11 include benchmarking analysis, mainly operated through a desk study of similar programmes abroad (i.e. evaluation of the Finish CoEs) and/or with interviews with managers of similar programmes abroad (e.g. Evaluation of the Austrian Christian Doppler Research Association). Benchmarking analysis is mainly use to compare and assess the efficiency and the structure of the programme (e.g. Estonian evaluation of centres of competences, Evaluation of the Austrian Christian Doppler Research Association). The evaluation of Finnish CoEs however use the benchmark analysis to assess the potential societal impacts driven by centres of competences, as a way to assess the relevance and possible further development of the programme.

- A few reports focus on descriptive statistical analysis in order to assess the results of the evaluated measures on participants. This is for example the case of the Danish evaluation of Innovation consortia, which develops a whole micro-economic model based on a total of 405 firm observations in the monitoring data of the programme over the period 1995-2003.
- Descriptive statistics are used when evaluation focus on the impacts of the measure on the economic/employment growth in participating firms. The analysis is based either on a counterfactual analysis and/or baseline analysis.
  - The counterfactual is approached in two reports through a survey question to the participating enterprises ('What would have happen without your participation to the programme?'. as seen in the UK evaluation of the KTP) or through a more thorough statistical analysis of the economic/employment growth in a control group made of similar but non-participating enterprises (e.g. Danish evaluation of Innovation consortia).
  - Likewise, the baseline approach is used in two reports and is based on the analysis of enterprises' statistical data over time.

5.1.1 The advantages and limits of the methods used

None of the evaluations disclose the budget for the exercise, however several do indicate the duration of the study, which is typically 6-9 months.

On the one hand, the eleven evaluation reports highlight the following strengths in the evaluation design of specific innovation measures for CoEs and CCs:

- The peer review and visits – especially where it involves users and producers – appears to be a good, practicable means by which to judge the efficiency and achievements of collaborative research structures, that are rather complex and long-lived entities that ought to produce manifold benefits across may years and in many locations through hard-to-observe knowledge spillovers. That said, such a process is quite demanding on the commissioning agency and subjects, and will have a minimum cost of perhaps at least €50K (driven by numbers of peers).
- The international benchmarks seem to be a similarly powerful and practicable means by which to gauge the relevance and future development of these structures, particularly for the programmes hat are rather recent at the time of the evaluation (e.g. Evaluation of the Estonian centres of competence).
- The use of statistical analysis of beneficiaries, coupled with the implementation of a survey to beneficiaries is the best way to assess the impact of measures on participating enterprises. The Danish evaluation of Innovation consortia, using a combination of baseline, counterfactual and economic modelling, is a good example of how one can look at the economic benefits – and wider effects – of an innovation support measure (funding smaller, commercially focused projects)

On the other hand, one cannot but notice a few limitations in the methods used:

- The obvious limitation with this pool of evaluations was the relatively light treatment of the fundamental question of science-industry collaboration:
  - When it comes to collaborative research centres, the assessment through peer review clearly focus on efficiency questions and scientific results at the time of the evaluation, which means very few of them take into account the individual impacts for participating enterprises, even if participating enterprises are most of the time involved in the evaluation process. For example, in the impact evaluation of Finnish CoEs, business partners were interviewed as part of the case studies, but the assessment of results on companies is limited to a development around the direct utilisation of research findings in companies. Few reports focus on the extent to which a particular measure had had a transformative effect on participants and the wider business community and the use of counterfactual or baseline approach is rather limited

  - On the contrary, when it comes to collaborative and knowledge exchange projects, evaluation reports focus almost exclusively on the effects for enterprises and the regional economic environment (e.g. Analysis of Firm Growth Effects of the Danish Innovation Consortium Scheme; analysis of the impact of the Dutch innovation voucher on innovative inputs and innovative output of companies; evaluation of the US ERC).
  - There are however two exceptions to this rule: the evaluation of the UK Knowledge Transfer Partnerships and the evaluation of the Christian Doppler Research Association, that both treat on an equal footing effect for the academic and business communities. Science-industry collaboration is studied through both perspectives, through individual surveys addressed to enterprises and academic partners.

- A surprise with the centres of excellence/centres of competences – was the absence of any bibliometric analysis and social network analyses. The former provides a very good source of data to help gauge trends in visibility and research quality, internationally normalised, and also trends in co-publications with industry. Network analysis, by definition, permits one to map certain types of relationship among communities and across geographies, while also tracing the evolution of those relationships through time
- Last not but not least, the assessment of the impacts of the CoEs/CCs - as longer-term collaboration - do not deal with the societal impacts of the programme. The Impact Evaluation of the Finnish Programmes for Centres of Excellence raises the question on how to assess the societal impacts of a programme for CoEs, but it does not assess the broader socio-economic impacts of the programme, focusing instead on the impacts of the different business and academics participants.

5.1.2 Identification of good practice

We would single out for further elaboration, each of the following:

- **Mid Term evaluation of the Swedish Institute Excellence Centres Programme (SE)** including a peer review of the eight individual centres through involvement of "generalist experts" looking into several centres next to the specialist experts only involved in the review of one centre. Each individual peer reviewed case study assess the long-term strategy and progress of the Centre Scientific and technical achievements and their impact, -up of a concentrated research environment and the leadership and management of the Centre. The evaluation report also includes an overall assessment of the whole programme
- **Impact Evaluation of the Finnish Programmes for Centres of Excellence (FI)** in Research 2000–2005 and 2002–2007: this very interesting evaluation tries to assess the impact and future development of the CoEs programme on the different types of participants (academics, businesses and host organisation) through a descriptive statistical analysis on monitoring data of individual centres, coupled with a series of interviews (with foreigners having worked at a Finnish CoEs; with researchers that were involved or not in CoEs; with host organisations)
- **Knowledge Transfer Partnerships Strategic Review (UK)**: Assessing the combined effects of knowledge transfer projects for academic and business participants with a (limited) counterfactual approach. It includes a combination of 202 interviews with businesses, academics and associates and a web survey of businesses, academics and associates

5.2 Strategic research

5.2.1 Introduction

Europe's regions have been increasing their support for user-oriented or applied research undertaken within universities and public research institutes – historically the domain of national research councils – where that research is directed to a programme of work deemed to be of strategic importance. The investments may be made through targeted research programmes or centre awards, however they address subjects of especial importance nationally and internationally, possibly cross-cutting issues such as climate change, where the regional agencies believe there may be an opportunity to expand local capacity and capability to a level sufficient to create an international comparative advantage in science and technology.

5.2.2 Rationale and objectives

Such investments are seen as a means by which to develop regional niches, which should in turn provide a platform for commercial success in for example the green economy as a result of the expanded and improved flow of trained people, intellectual property and know how.

In several regions, these measures are being used to brigade competing and fragmented institutions, in order to build coherence and attain critical mass rapidly. For example, we have selected an evaluation of Genome Canada, the principal funding stream for genomics in Canada. This programme has an objective not just to fund research done in genomics, but also enable large-scale projects to be conducted that would not have been possible otherwise because they are being funded from a national pot. It aims to bring isolated pockets of research in genomics into contact with each other by encouraging projects that create new networks, and it aims to enable researchers to have access to important resources by creating new technology platforms.

5.2.3 The evaluation record

5.2.3.1 Results of the mapping

We have selected seven evaluations in total, which are all in the area of strategic research:

- Genome Canada is a national programme in genomics and proteomics research set up to develop and implement a national strategy which would coordinate all of the work done in this field at a national level. The programme addresses all of the key strategic areas including health, agriculture, environment, forestry and fisheries. It has set up six Genomics centres across the country and funds 50% of large-scale projects and 100% of science and technology platforms.
- The Functional Genomics Programme in Norway (FUGE) has been set up to strengthen Norwegian research in functional genomics, which is research designed to 'to reveal the biological function of genes, and how sets of genes and their products work together in biological systems.' The backbone of FUGE is 11 technology platforms, which are coordinated at the national level.
- The Austrian Genome Research Programme (GEN-AU) was a nine year long programme, which funded a variety of cooperative projects, network projects, pilot projects and projects addressing accompanying research in the social sciences. The projects varied in the number of partners involved the running time and the funding volume.
- The Quantifying and Understanding the Earth System Programme (QUEST) is a six-year long programme. Its mission is 'quantifying Earth system processes and feedbacks for better informed assessments of alternative futures of the global environment.' The programme has the scientific goal of creating better predictive models of the Earth System and its interactions, it also has an operational goal of

increasing collaboration across disciplines and institutions and a policy goal of effectively communicating the results to national stakeholders and politicians.

- BBSRC-supported 'genomics' research supports two programmes, one is The Investigating Gene Function (IGF) initiative, which funds consortia to develop genomics technologies and associated resources for communities working on organisms key to the BBSRC remit, the other is 'Genomics' research funded through responsive mode, which is investigator-driven research from across the BBSRC remit which incorporates 'genomics' technologies and approaches.
- Science Foundation Ireland (SFI) is a national body established in 2000 with the objective of ensuring that Irelands research capabilities in certain specialist science and engineering fields (specifically ICT, Biotechnology and related fields) were subject leaders and of the highest international standards. The agency was given official legislative status in 2003, with the objective to 'promote, develop, and assist the carrying out of oriented basic research in strategic areas of scientific endeavour that concerns the future development and competitiveness of industry and enterprise in the state'
- The Flemish Strategic Basic Research financing channel is a succession of programmes, the first of which was Strategic Technologies for Welfare and Welfare (STWW), which funded scientific or technological research in universities in economic and socially relevant areas. The second programme Strategic Basic Research (SBO) was the successor of STWW and Generic Basic Research at the universities (GBOU) followed suit.

5.2.3.2 The questions addressed by the evaluations

Below are the evaluation criteria used in the evaluations under review.

Figure 14 Criteria used in the evaluation of Strategic research

Name of the measure evaluated	Type of measure	Relevance	Efficiency	Effectiveness	Sustainability
Evaluation of Genome Canada: Final report (CA)	Grants to fund theme-related projects	√		√	
Evaluation of the Functional Genomics Programme in Norway (FUGE) (NO)	Grants to fund theme-related projects	√		√	√
Austrian Genome Research Programme (GEN-AU): Mid Term Programme Management Evaluation (AT)	Grants to fund theme-related projects	√	√	√	√
Mid-term evaluation of the Quantifying and Understanding the Earth System Programme (QUEST) (UK)	Grants to fund theme-related projects			√	√
Evaluation of BBSRC 'genomics' research (UK)	Grants to fund theme-related projects	√		√	√
Value for Money Review of Science Foundation Ireland (IE)	Organisation supporting theme-related projects and science-industry cooperation	√		√	√
Flemish Strategic Basic Research financing channel (SBO programme and its predecessors STWW, GBOU) (BE)	Grants to fund theme-related projects			√	√

Five of the seven evaluations explored the issue of **relevance**. These included all four of the research funding programmes for genomics and the value for money review of the Science Foundation Ireland. This could be because the research funding

programmes for genomics each aimed to bring together research being done at university and institutional level into a single cohesive national strategy. For this reason, it may have been deemed necessary to examine the role and space the programmes occupied in the genome funding landscape in order to ask whether the programme overlapped with related funding programmes and whether there was a need for the programme based on these assumptions. Likewise, the Science Foundation Ireland aims at funding research with a strategic significance in terms of national social and economic value (e.g. biology and ICT), and as such it was deemed necessary to review the effectiveness and impacts of the research supported by the Agency.

All of the evaluations asked whether the programmes they were assessing were **effective** in meeting their objectives, and in some ways, this could be seen to be the point of the evaluations. However, none of the programmes were SMART (Specific, Measurable, Achievable, Realistic and Time-bound) and in this regard, there was blurring in the distinction between effectiveness and impacts.

Only the evaluation of the Austrian Genome Research Programme asked the question of whether the programme they were evaluating was **efficient**. It did this by calculating a ratio between the programme's outputs in terms of papers produced against the costs of running the programme. The evidence was collected via programme documents plus a survey, which asked respondents to quantify the benefits they felt the programme had produced. It is possible that the reason not many of the evaluations have not included a question on efficiency is because the outputs of strategic research are hard to measure, as the example of the Austrian Genome Research Programme shows.

All of the evaluations except one addressed the issue of the **future direction** of the programme. This could be because one of the key functions of an evaluation of strategic research is to examine how well the programme is being run and to advise on more successful ways of funding the research. Indeed, four of the seven evaluations involved a peer review committee, whose primary purpose is to give expert advise on the running of the programme. In most of the evaluations, the recommendations formed a central part of the report structure.

Additionally, all of the evaluations except one attempted to identify the **impact** the programme had either in terms of establishing effective mechanisms of liaison between the science and user community or in terms of the transformative effect it has had on the quality of the research within the nation, or in terms of economic benefits it has had on the industry in question. Only one of the evaluations did not attempt to identify the wider impact of the programme and this report had been tasked with finding out how to facilitate answering this question in the future (e.g. by keeping good records of completed programmes, having regular evaluations, doing occasional case studies.)

There was some variability in the extent to which the evaluations went into depth in calculating impacts. Some simply asked the respondents whether the programme had established good links with industry and government representatives, for example the Flemish Strategic Basic Research financing channel asked its respondents how many university spin-offs had been created, while four of them attempted to calculate **net effect** or **deadweight losses**.

In terms of other issues, five of the evaluations assessed the **managerial and operational** aspects of the programmes. This was occasionally done in considerable depth because the mandate of the programme was to coordinate research within the respective sector it represented.

5.2.3.3 The evaluation methods employed

Below are the evaluation methods used in the evaluations under review.

Figure 15 Methods used in the evaluation of Strategic research

	Qualitative methods			Quantitative methods				Approaches			
Name of the measure evaluated	**Individual interviews**	**Case studies**	**Peer review**	**5.2.4 Questio**	**Use of administrative/ secondary data**	**Micr-/ macro-economic model**	**Bibliometrics**	**Baseline**	**Counterfactual**	**Benchmarking**	**Indicators**
Evaluation of Genome Canada: Final report (CA)	√			√	√		√			√	
Evaluation of the Functional Genomics Programme in Norway (FUGE) (NO)	√		√								
Austrian Genome Research Programme (GEN-AU): Mid Term Programme Management Evaluation (AT)	√	√		√			√			√	
Mid-term evaluation of the Quantifying and Understanding the Earth System Programme (QUEST) (UK)			√	√							
Evaluation of BBSRC 'genomics' research (UK)			√	√	√					√	
Value for Money Review of Science Foundation Ireland (IE)	√		√	√	√	√	√			√	√
Flemish Strategic Basic Research financing channel (SBO programme and its predecessors STWW, GBOU) (BE)	√			√	√			√			

Despite an ostensible similarity between the programmes being evaluated, there was a surprising amount of difference in the evaluation tools used:

- All of the evaluations except one used **questionnaire surveys** to collect the data they based their findings on. In three cases this was in the form of a survey analysed by the evaluation committee and in three cases it was in the form of preparatory documents gathered by a third party and presented to a peer review committee as part of a set of evidence for them to analyse. The advantage of using questionnaire surveys is that they are highly flexible and can be used to answer almost any of the questions the evaluator might want to ask.
- Four of the evaluations used **individual interviews**, usually in the form of interviews. In three of the evaluations, these were used to add depth and colour to the bulk of the evidence that had been collected via the questionnaire surveys. However, the only evaluation that did not use questionnaire surveys was the Functional Genomics Programme in Norway (FUGE), which was evaluated using a peer review committee that spoke to key stakeholders in each of the technology platforms funded under the programme instead.
- Four of the seven evaluations used **peer review** as a central part of their assessment process. The peer review committees were each comprised of key representatives from the research community, the end-users of the technology and in some cases policy-makers. In three of the evaluations the final report was

written by the peer review, and in one case it added to the evidence reviewed by an independent consultancy.

- Three of the seven evaluations made use of **a descriptive statistical analysis**. The information to support this analysis was assembled from programme documents and was used to understand the composition of programme expenditure, activity and outputs.
- The only evaluation where **cases studies** were used was in the Austrian Genome Research Programme GEN-AU, where four case studies were done in considerable depth. Unfortunately, due to the confidentiality of the material most of the evidence was not made explicit, but the findings were published along with the recommendations.
- Two of the seven evaluations in some way made use of **bibliometrics**. The Value for Money Review of Science Foundation Ireland used a **bibliometric** assessment of the Foundation's researchers and the publications they had produced in order to ascertain the quality of the research outputs. In particular by comparing the Foundation's outputs with other Irish publications in similar disciplines and by comparing them to international publications in the same discipline. On the other hand, the Evaluation of Genome Canada did not conduct its own bibliometrics, but it did make use of the bibliometric findings of a separate report that had analysed the same programme. The evaluation of Austrian Genome Research Programme (GEN-AU), conducted its own **Social Network Analysis** (SNA), a technique used to analyse social relationships. For example, measuring whether collaboration has increased between researchers and end-users as a result of participating in one of the projects. The evidence found to support the Social Network Analysis was collected via an on-line survey, and formed part of a range of methods used by this evaluation as part of a drive to use 'mixed-methods.'
- None of the evaluations made use of any **econometrics** techniques. This was appropriate given the type of schemes under evaluation and that most of the outputs were not quantifiable econometrically.

5.2.5 The advantages and limits of the methods used

Questionnaire surveys were the primary way of collecting data in the majority of the evaluations. This method has the advantage of being very flexible and allowing the evaluator to collect data on a range of issues and questions, without creating excessive costs and can usually be done within six to nine moths. For example, in the case of the Evaluation of Genome Canada, questionnaire surveys were put to good use in examining the *relevance* of the programme. They did this by asking the respondents whether they felt that Genome Canada had been instrumental in coordinating genomics research across Canada according to a coherent national strategy, and whether they felt that since it had been set up there remained a strong rationale to keep it going. On the other hand, in the evaluation of the Austrian Genome Research Programme (GEN-AU) questionnaire surveys were used to collect data on the *effectiveness* of the programme to encourage networking. This was achieved by asking the respondents to supply data on who they had collaborated with and was then further analysed using Social Network Analysis.

Individual interviews were used in four of the evaluations we analysed. In three of the evaluations it was used to add colour and depth to the bulk of the evidence which had been collected through the questionnaire surveys. In one of the evaluations a 'hearing' was organised composed of key stakeholders in each of the technology platforms funded under the programme. The Austrian Genome Research Programme GEN-AU, did carry out extensive individual interviews, interviewing 60 people in person. The drawback to this kind of technique is that it is expensive relative to questionnaire surveys, especially if the interviewees are in different locations.

The primary way the information was analysed was using a **peer review** committee. Of the four evaluations that used this method, three were written by the peer review

committee itself and one used peer review as a major component of the evaluation. The peer review committees were usually selected by the steering group and included international experts in their field and end-users of the research. The main advantage of a peer review committee is that it has experience of the research area, but is distanced from the programme itself and can thus contribute unbiased expert knowledge. For this reason, the committees were invited to make recommendations on the future direction of the programme in all the evaluations they contributed to. A disadvantage of bringing a peer review committee into an evaluation, especially one that is already being handled by another consultancy, is that it can be relatively difficult to organise and expensive to run. This is because although the peer review committee will normally contribute to the evaluation for an honorary sum, the flights and lodgings can still be expensive. Also, because each peer review committee member is generally required to meet together for a period of time such as at least a week, peer reviews usually have to arranged considerably in advance of time.

Benchmarking is another method that was successfully used in more than half of the evaluations. This was carried out in two different ways in our sample of evaluations. If the evaluation was written by a consultancy then the benchmarking usually formed part of the desk research and would involved comparing the programme in question to similar programmes internationally. Where the evaluation was handled by an international peer review committee, then the committee were already experts in their fields in their respective countries and could contribute by discussing how the programme in question compared to how the discipline is run at home.

Social Network Analysis (SNA) was carried out in the evaluation of Austrian Genome Research Programme GEN-AU. Social Network Analysis is a technique used to analyse social relationships such as measuring whether collaboration has increased between researchers in project groups or between researchers and end-users as a result of participating in one of the projects. The way the evidence for the SNA was collected was by adding a question to the questionnaire survey about who the respondents had collaborated with and the relationships then drawn onto a map, ready for analysis. This technique was useful for understanding how effective the programme had been in achieving its objective of encouraging collaboration between researchers and industry. It also gave an indication of the quality and intensity of the relationships and allowed the evaluators to identify which the strategic and peripheral actors were in the projects, enabling the evaluators to make suggestions for improving the operation and management of the programme.

5.2.6 Identification of good practice

By way of conclusion, we found that all of these evaluations had some good points to report on.

Figure 16 Good practice in the evaluation of measures to support Strategic research

Objective of the evaluation	Example	Best practice in the use of evaluation method
Assessing the impact and relevance of the funding programme	Austrian Genome Research Programme (GEN-AU): Mid Term Programme Management Evaluation (AT)	Strong emphasis on a 'mix-methods' approach
Assessing the efficiency of the funding programme	Value for Money Review of Science Foundation Ireland (IE)	Use of bibliometrics
Assessing the future development of the funding programme	Mid-term evaluation of the Quantifying and Understanding the Earth System Programme (QUEST) (UK)	Use of Peer Review

Some of them had points that we have chosen highlight as examples of best practise:

- The **evaluation of the Austrian Genome Research Programme (GEN-AU)** (AT) used a mix of methodologies in order to arrive at its assessment of the impact and relevance of the programme. First a logic model was drawn in order to see what the objectives of the programme were and where these fit into the genomic research landscape. Then an analysis of the programme documents on the effectiveness of the programmes. 70 individual interviews were held in order to enrich and add colour to these findings, and to find out what the stakeholders felt about the programme. Then using an on-line survey, evidence was collected to do a Social Network Analysis, which told the evaluators how successful the programme had been in terms of creating new networks and diffusing the technology developed back to the end-users.
- The **Value for Money Review of Science Foundation Ireland (SFI**) (IE) used bibliometrics to assess the effectiveness of the programme on improving the quality of the research it funded and assessing whether the SFI funding had impacted on author productivity. This was achieved by comparing the research funded by SFI to other publications in similar disciplines published in Ireland and by comparing it to publications published internationally within the same discipline.
- **Austrian Genome Research Programme (GEN-AU) (AT):** Mid Term Programme Management Evaluation used a Peer Review Committee to write proposals and recommendations on the future development of the programme. The peer review committee were comprised of UK and international members with expertise across QUEST'S remit, but who were not closely involved with the programme. It also included one member from an organisation that uses research of this type. The committee were presented with documentary and survey evidence and used this to identify the key risks and issues the programme contended with. Based on this rigorous analysis the committee could make recommendations about how to improve the programme in the future.

5.3 Direct financial support for innovation activities

5.3.1 Introduction

A total of seven evaluations out of the 120 or so evaluation reports in our repository address measures that are primarily concerned with the direct financial support to business research and development (R&D), which we define as the support for business R&D and demonstrator projects (through loans or grants). Support is granted through many and varied application processes, that are either competitive (when issued through calls for proposals with competition between different projects) or non-competitive.

All of these evaluations were conducted on behalf of public agencies, national ministries of technology agencies and regional agencies (IWT, the Flemish Agency for Science and Innovation) in one case. The evaluation budgets are not disclosed. A great majority of the evaluations (five) were carried out by independent consultants or organisations. Two were carried out by national organisations: the Netherlands Bureau for Economic Policy Analysis and the National Audit Office in Estonia.

5.3.2 Rationale and objectives

The policy logic of direct financial support for business R&D is one of market failure, which is to say that businesses on average will tend to invest less in research and development than is optimal for society, due to uncertainty and an adjudged imbalance between risk and reward.

Publicly co-financing proprietary innovation projects ought to cause those assisted businesses to do materially more development work than would be the case otherwise, producing more innovation (in turn this should result in increased sales / profits for assisted businesses, increased productivity gains for their customers and maybe a consumer surplus too, etc). Offering grants, credits, loans directly to selected businesses (open to all comers, but selectively investing in the best proposals and teams) may be preferable to a tax credit in that it is targeted (on more promising opportunities), causes participants to work harder in pursuit of their project goals and such selective assistance engenders rivalry between the assisted and unassisted and should cause a re-calibration of acceptable levels of investment for a sector overall.

5.3.3 The evaluation record

The evaluation questions

Figure 4 presents the scope of evaluation questions assessed in each of the seven evaluations reviewed. It also provides indications on the focus of evaluation (type of scheme evaluated) and on the nature of the evaluation (ex-post, ongoing or mid-term).

Our sample of seven reports comprises two mid-term, four ex-post and one ongoing evaluations. All of them sought to test in particular the effectiveness of the initiatives and all except one mid-term evaluation analyse the impact of the schemes. A minority (3) of the evaluations analysed the relevance (or appropriateness) and the sustainability of the measures. The operational efficiency was analysed in a majority of evaluations (4 out of 7) and counterfactual aspects were also explored in four out of the seven reports under review, in particular through counterfactual approaches.

The evaluation of the Austrian AWS technology programmes took a portfolio perspective; building on several evaluations already done for most of the programmes in question.

Figure 17 Criteria used in the evaluation of Direct financial support to innovation activities

Name of the measure evaluated	Focus of evaluation	Type Interim/ ex-post	Relevance	Efficiency	Effectiveness	Sustainability
IWT grants for R&D projects of companies in Flanders (BE)	R&D grants for companies	Ex-post			√	
FFF (Industrial Research Promotion Fund, Austria) General Programmes (AT)	Industrial research Promotion Fund: industrial R&D projects	Mid-term			√	√
Dutch Innovation voucher 2004 and 2005 (NL)	Innovation voucher for SMEs collaborating with PROs	Ex-post			√	
Norwegian DEMO 2000 programme (NO)	Technology development programme	Ex-post		√	√	
Austrian AWS Technology Programme (AT)	Portfolio of technology programmes	Mid-term	√	√	√	√
Estonian State's enterprise supports on the competitiveness of the economy (EE)	Co-financed grants for R&D of SMEs, from micro finance to larger amounts	Ongoing[31]	√	√	√	√

[31] The evaluation is more a kind of global review of all support measures intended to foster enterprises' competitiveness in Estonia – but not a formal evaluation of a specific programme/policy. As a

Evaluation of Grant for Research and Development & SMART (UK)	Technical development of established SMEs	Ex-post	√	√	√	

5.3.3.1 The evaluation methods implemented

Figure 18 presents the evaluation methods used in every report.

Figure 18 Methods used in the evaluation of Direct financial support to innovation activities

	Qualitative methods			Quantitative methods				Approaches			
Name of the measure evaluated	**Individual interviews**	**Case studies**	**Peer review**	**Questionnaire survey**	**Use of administrative/**	**Micro-/ macro-economic model**	**Bibliometrics**	**Baseline**	**Counterfactual**	**Benchmarking**	**Indicators**
IWT grants for R&D projects of companies in Flanders (BE)	√	√		√	√			√	√		
FFF (Industrial Research Promotion Fund, Austria) General Programmes (AT)	√			√	√	√		√	√		
Dutch Innovation voucher 2004 and 2005 (NL)				√		√			√		
Norwegian DEMO 2000 programme (NO)	√		√	√	√			√			√
Austrian AWS Technology Programme (AT)					√			√		√	√
Estonian State's enterprise supports on the competitiveness of the economy (EE)	√			√		√			√		√
Evaluation of Grant for Research and Development & SMART (UK)	√			√	√			√	√	√	√

Figure 18 displays the following methodological patterns:

- All evaluations except one include a questionnaire survey, mainly to canvass the opinions of participants on the effects and impacts of the programme evaluated. In particular, all evaluations trying to develop a counterfactual approach are using surveys of beneficiaries and non-beneficiaries to collect information on the net effect of the schemes. The only evaluation that does not involve a questionnaire survey is the evaluation of the Austrian AWS technology programme which is based on previous evaluations of the different component programmes.
- 5 of 7 evaluations make use of counterfactual methods. All of these evaluations use a survey of beneficiaries/non beneficiaries (applicant and non-applicant) for this purpose (IWT R&D projects, FFF, Dutch Innovation Vouchers, SMART). Interestingly, the evaluation of the Dutch innovation vouchers is based on the fact that the vouchers have been assigned randomly by means of a lottery. Reserves are nonetheless made because of the possible selective non-response in the survey (half of the SMEs with a 2004 voucher did not respond to the interview). Therefore, effects for the total population may differ.

consequence, the evaluation is not linked to a particular periodicity of the policies involved and can be assessed more as an ongoing evaluation.

- 5 of 7 include a desk research to document the investments of the measure in question, as regards the volume and direction of funding and the composition of supported activities and resulting outputs. This analysis is mainly used as background, scene-setting to help the reader better understand the subsequent and more critical evaluation chapters, as well as the programme logic as regards the location and composition of demand for support, and also as a reference for programme effectiveness, again, cross-checking the balance of investments and outputs with the type and mix of outcomes being sought.
- 5 of 7 include a individual interviews taking in most cases the form of interviews. This is often used following a preliminary analysis through a survey or a desk research. This addresses questions about relevance in the main but also programme efficiency and added value (within the wider landscape) and programme effectiveness (to a lesser degree).

Each of the other methods was used in a minority of the seven evaluations, and specifically, in descending order of frequency of use:

- 3 of 7 involve econometrics. In two cases these were light descriptive statistical analysis to allow for a counterfactual analysis and measuring deadweight or displacement effects (e.g. Estonian evaluation, Dutch Innovation Voucher). In the case of the FFF evaluation, the work is relying mainly on a survey and simple econometrics to estimate to net effects.
- Two evaluations out of 7 involve benchmarking exercises mainly operated through a desk study of similar programmes and/or with interviews with managers of similar programmes. As regards the evaluation of the AWS technology programmes, the benchmarking takes place mainly within the different programmes of the portfolio and within the wider Austrian system for supporting high-tech companies. The Estonian evaluation provides a benchmark with the research of the impacts of enterprise supports made in Scotland
- Only one evaluation involves case studies. The evaluation of the IWT R&D support to companies includes three case studies, mainly based on interviews. The three companies presented experienced different effects as a result of the IWT intervention
- One evaluation report states that its methodology included a 'peer review' (The DEMO programme), however on closer reading this element of the work seems to be more an in-depth analysis of two projects based on interviews of project leaders and representatives from sponsoring oil companies rather than an international academic peer review as usually performed for the evaluation of basic research activities
- No evaluation makes use of bibliometrics

5.3.4 The advantages and limits of the methods used

Looked at as a set, the seven evaluations of direct financial support have tended to focus on the two classic policy evaluation questions of effectiveness and impact and to deploy a core set of data collection methods:

- Desk research to gather secondary data
- Composition analyses
- Stakeholder interviews
- Questionnaire surveys
- Counterfactual analysis

This combination of methods addresses the question of programme effectiveness wherein evaluators can judge the nature and extent of the achievements (from monitoring data and surveys) against the publicly stated ambitions for the policy or

programme – achievement of objectives, added value of the support in comparison with privately funded activities – through both objective data (composition analyses using data on recorded expenditure, activities and outputs) and more subjective data (opinions on effectiveness and impact as compared with other similar measures, through the surveys and consultations). The surveys and interviews also allow for collecting data from non-beneficiaries and form control group to assess the net impact of the initiatives.

These data collection methods are similarly used to gather data to help answer the primary evaluation question of financial efficiency – number and value of outputs produced for each million Euros of public investment. The surveys and consultations also support a more formative analysis of operational efficiency, inviting participants and other stakeholders to offer advice on affordable and practicable improvements to the rules or delivery arrangements that might reasonably be expected to produce more benefits for the same investment

The advantages of this study design are threefold:

- The approach is straightforward and easily implemented using programme records and monitoring data and a small (low cost) commission (external evaluators to run independent surveys);
- There are numerous published studies available that showcase the basic analyses and also disclose the survey questions and checklists, so there is good provenance and the work can be replicated with some confidence;
- The combination of methods provides some degree of cross-referencing and triangulation, increasing confidence in the findings, as a result of using data from different sources (e.g. programme records, participants and wider stakeholders) to encircle each core evaluation question

Its limitations are threefold:

- The majority of these seven evaluations is somewhat self-referential, with only a minority having included work to obtain comparable data for other analogous measures (benchmarking and comparative analyses). None of the evaluations sought to determine the extent to which targeted financial support was a more efficient or effective means by which to trigger increased innovation and economic growth as compared with a general tax facility that reduces the marginal cost of research for all businesses
- Several evaluations make use of counterfactual methods, which bear in themselves a lot of limitations as regards the collection of data and the interpretation of results which are based on a set of assumptions on the behaviours and performances of supported and unsupported firms
- Only a minority of the evaluations sought to control for external factors, in an effort to more robustly understand the added value of a given measure within the much broader landscape of changing competition, demand and other macro-economic conditions
- None of the evaluations made a good job of detailing and quantifying the wider effects of these measures, whether that is the impact of selective assistance on the behaviour of unassisted businesses (through competition and peer-group pressure) or the knowledge spillovers that should accrue from the step change in innovativeness. Related to this, none of the evaluations addressed the longer-term impacts of the policy or programmes in question

In order to measure the impact of support on business R&D, evaluations tend to use different type of proxies (BERD, patents, increase in sales, etc.). In most evaluations, it is however recognised that many effects of the support might occur or be visible only in the longer term.

5.3.5 Identification of good practice

Most of the evaluations used a broad mix of data collection methods in order to seek convergence across multiple streams of partial / compromised data. Having stated these basic considerations for the evaluation of support to business R&D, we consider that the most interesting evaluations using methodologically noteworthy approaches are the following.

Figure 19 Good practice in the evaluation of Direct financial support to innovation activities

Objective of the evaluation	Evaluation Report	Good practice in the use of evaluation methods
Measuring wider effects and spillovers	Norwegian DEMO 2000 programme (NO)	Computation of indicators of value creation at a socio-economic level
Attributing the observed scheme to the scheme evaluated (net effect)	IWT grants for R&D projects of companies in Flanders (BE)	Focus on assessing behavioural changes (*differences in firm behaviour resulting from an intervention)* based on survey and interviews of beneficiaries and non-beneficiaries:
Measuring net impact of scheme on the target business	Evaluation of Grant for Research and Development & SMART (UK)	Large-scale questionnaire survey, calculate impacts including estimating the net effect, displacement, deadweight, etc.

For the evaluation of the **DEMO 2000 programme** (NO) the evaluators have attempted to compute indicators of value creation at a socioeconomic level, but with a lot of reserves. Several sources of information were used. The most important were: programme documents and archives; Web-based surveys (project leaders, representatives from sponsoring oil companies and other co-operating companies and research- as well as other institutions); Peer-review; Interviews. The evaluators made no attempt to measure e.g. employment effects or effects on tax income. In addition, these indicators may to some degree overlap as different technologies and projects are needed to exploit the full potential of an asset. Such effects are not accounted for in this study.

The **evaluation commissioned by IWT in Flanders (BE)** on R&D support to companies made a good job of exploring the added value of the measure in terms of behavioural change, which is to say the extent to which the financial assistance had produced changes in attitudes and behaviour that endure beyond the life of the grant. As quantitative techniques have limited analytical power to explain the full effects of public support, more recent work (OECD 2006)[32] has underlined the importance of taking into account behavourial changes: '*the differences in firm behaviour resulting from an intervention*'. The idea here is that an evaluation should explore the effects beyond the direct impact (e.g. firm has increased sales from a new product commercialised thanks to funding for an industrial R&D project) in terms of the way the funding has generated permanent changes in the process and practice and capabilities of a firm to undertake innovation (e.g. has the firm developed new methods or tools for identifying innovation projects, have they improved methods for managing the innovation process, etc.).

The evaluation of the **Grant for Research and Development & SMART (UK)** employed a good practice example for calculating the net impacts of the support scheme on the target businesses. The evaluation followed the UK's Impact Evaluation Framework (IEF) methodology, allowing the evaluators to provide a robust assessment of the (net) additional economic impact of the scheme by allowing for deadweight, displacement effects and spillovers. Application of the methodology permitted the evaluators to calculate the 'rate of return' for each € (or £ in this case)

[32] OECD, 'Government R&D Funding and Company Behaviour, measuring behavioural additionality', 2006

invested in the support scheme, and used the (positive) outcome to confirm the scheme's effectiveness.

5.4 Networks and clusters

5.4.1 Introduction

The concept of clusters became a target for regional initiatives in the 1990s following Michael Porter's *The Competitive Advantage of Nations*. The main argument of cluster development was that firms and supporting organisations that operate in close proximity are often more competitive than isolated firms. It was argued that proximity improves innovativeness by facilitating the creation and dissemination of knowledge and skills. This is due to competition and co-operation. Co-operation not only in the form of formal alliances, but also enterprises benefiting from tacit knowledge being exchanged between firms along the value chain, or through other form of social interactions.

The cluster concept, and cluster policies and programmes have evolved considerably over time. Starting with Marshall's concept of "industrial districts", where geographically concentrated clusters can be explained by specialised labour, specialised intermediate inputs and knowledge spillovers, regional cluster policy has evolved using concepts such as 'learning regions', 'innovative milieus' and 'regional innovation systems', which stress learning as a key factor or regional competitiveness. As a consequence policy interventions have shifted from simply influencing the inputs of business activity to be more focused on the relationships between industries, which underpin competitiveness. Thus, more attention has been given over time to supporting co-operation in favour of business innovation and the role of factors and structures at the regional level that promote innovation activities

Today, numerous definitions of clusters and networks can be found in the literature review. One of the most suitable definition of clusters, and more particularly of innovation clusters, has been established by the Community Framework for State aid for research and development and innovation[33]:

> *'Innovation cluster means groupings of independent undertakings — innovative start-ups, small, medium and large undertakings as well as research organisations— operating in a particular sector and region and designed to stimulate innovative activity by promoting intensive interactions, sharing of facilities and exchange of knowledge and expertise and by contributing effectively to technology transfer, networking and information dissemination among the undertakings in the cluster.'*

For the purpose of this study, the previous definition needs to be complemented with the triple-helix[34] model definition as a process where the three institutional spheres, which formerly operated at arms' length, are increasingly working together and where linkages do emerge at various stages of the innovation process.

5.4.2 Rationale and objectives

A broad literature and empirical research on innovation systems has developed over the last two decades since the seminal articles of Freeman (1982)[35] which provide a

---

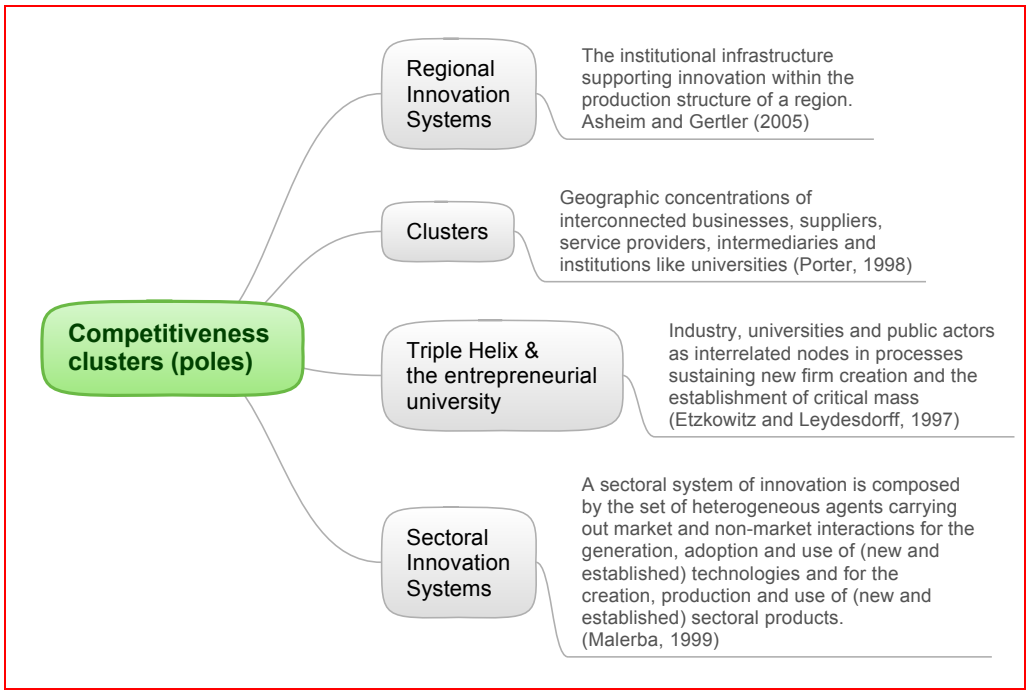
33 European Commission. (2006). Community Framework for State Aid for Research and Development and Innovation (2006/C 323/01).

34 The triple helix system was introduced by Professor Henry Etzkowitz who studied the importance of joining these three different actors (government, business and university) in the economic activities to improve the regional development continuously.

35 See Freeman C. (1982), Technological Infrastructure and International Competitiveness, paper for the OECD Expert Group on Science, Technology, and Competitiveness. Freeman, C. Technological

theoretical and conceptual under-pinning for public policies in favour of clusters or networks or clusters) Figure 20 summarises four main strands of research that contribute to the rationale behind cluster policy.

Figure 20 Theoretical underpinning of clusters



The traditional market failure justification for government intervention in favour of R&D and innovation has been increasingly challenged by innovation system[36] theory that underlines the importance interactions between agents and of policy tackling 'bottlenecks', or system failures, in innovation systems rather than 'isolated innovation events' through subsidies to single agents (companies, etc.). A systemic approach emphasises the importance of the microeconomic business environment and of linking business, universities/research and public actors in what has been called a "triple helix" of innovation[37].

The literature underpinning the logic of intervention for clusters (and associated instruments such as competitive poles and competence centres) underlines a range of potential effects ranging from agglomeration forces through improved knowledge exchange, to technological ('smart') specialisation and improved management of value chains. Equally, the literature has increasingly underlined that policies that focus exclusively on strengthening regional linkages are not optimal and it is important that involvement in such initiatives encourage firms to connect 'regional buzz' to national and international networks by encouraging the growth of national and international pipe lines[38]. One of the arguments, indeed, for linking up regional and national

Infrastructure and International Competitiveness paper . OECD Group on Science, Technology and Competitiveness, 1982.

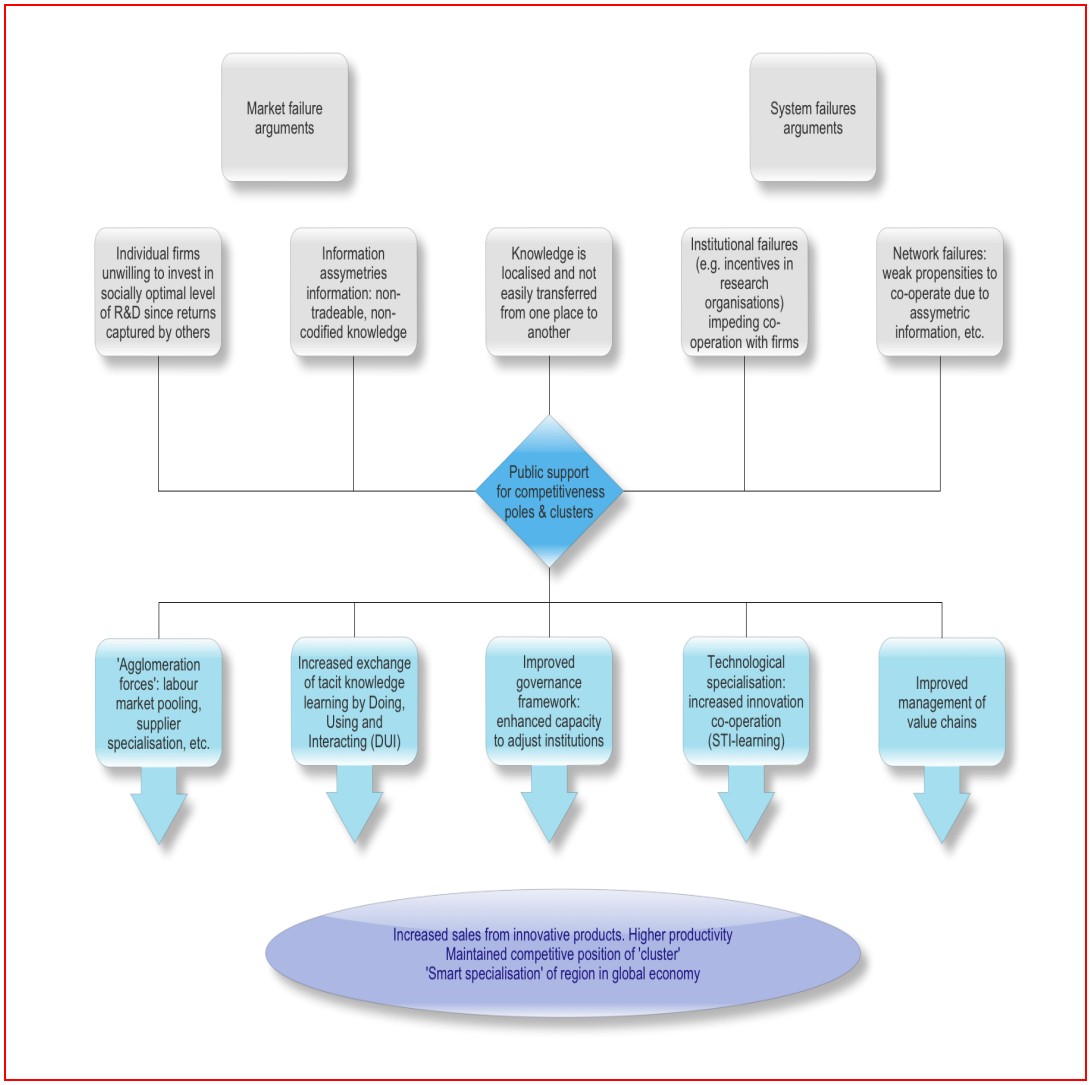
[36] Smith, K. (2000). *Innovation as a systemic phenomenon: rethinking the role of policy*. Entreprise and Innovation Management Studies, 1 (1) 73-102.

[37] Etzkowitz, H., & Leydesdorff, L. (1997). *Universities and hte global knowledge economy: a triple helix of university-industry goverment relations*. Pinter, London

[38] Huang, Q., McDonald, F., Tsagdis, D., & Tuselmann, H. (2007). *Is there evidence to support porter-type cluster policies?* Regional studies vol 41 n.1

business people with their academic counterparts is that the latter often are active in international research networks and can act as bridges to a broader knowledge base.

Figure 21 Intervention logic for clusters and networks



Unlike other innovation policy instruments, clusters and networks have been systematically evaluated and their impact analysed. For instance, a recent compilation of macro-economic benefits from success stories on cluster initiatives in the European Union by the IRE subgroup (2008) shows two types of benefits from cluster initiatives:

- Those related to knowledge spillovers, including the creation of formal and informal linkages and networks between firms, research institutions, public agents and other local organisations
- Those related to the increase in the attractiveness of the hosting regions, including productivity rises, competitiveness enhancement, and in the long-term, economic growth and employment

Mid-term and final evaluations have shown as immediate impacts evolutions of the cluster populations, notably increased participation of SMEs but also re-orientation of

university activities towards economically more relevant research as well as boosting industrial doctoral studies[39]. Impacts related to knowledge spillovers include increases in co-operation processes between research institutes and the private sector; increase of partnerships in the private sector[40]; and improvement of the linkages between public and private research institutes and groups. Public research laboratories had benefited substantially (Conseil régional de Bretagne, 2008) and became central in pushing for collaborative projects, between the public and private sectors.

5.4.3 The evaluation record

5.4.3.1 The evaluation questions

Out of the already mentioned approximate 120 evaluations reports in our repository, the team has identified **12 relevant evaluation** examples focusing on the assessment of cluster initiatives. Figure 22 shows at a glance the scope of the evaluation questions that have been assessed in each of these evaluations as well as the periodicity of the evaluation (mid-term or ex-post evaluation) and the focus of the evaluation.

Figure 22 Criteria used in the evaluation of Networks and clusters

Name of the measure evaluated	Focus of evaluation	Periodicity of the evaluation	Relevance	Efficiency	Effectiveness	Sustainability
Collaboration programme (IE)	Evaluation of the three principal programmes supported under the RTDI for collaboration component of Ireland's Operational Plan for industry. Good focus on how the components schemes complement each other	Ex post	√	√	√	√
The Centre of Expertise Programme (OSKE) (FI)	Assessment of the implementation and configuration processes of the Programme through the rest of the programming period (2007-2013)	Mid-term	√	√	√	√
Walloon Competitiveness poles (BE)	Review the effectiveness of the five priorities of the Walloon economic development plan (including competitive poles implemented over the period 2006-2008)	Ex post		√	√	
National Research Flagship Program (AT)	Examination of the large-scale multidisciplinary research partnerships that harness world-class expertise	Mid-term	√	√	√	√
Advanced Network and Research for Industry and Education (CANARIE) - ADDP programme (CA)	Evaluation of the Advanced Applications Development Programme managed by the CANARIE and supporting advanced network applications and related technologies	Mid-term	√	√	√	√

[39] Arnold E., K. Männik, R. Rannala, A. Reid (2008) Mid-Term Evaluation of the Competence Centre Programme.

[40] Belgian Science Policy Office/BELSPO (2010), Belgian Report on Science, Technology and Innovation, June 2010

Energy-cluster.ch (SW)	Assessment of the main activities of the cluster since its establishment in 2004	Mid-term	√	√	√	√
GEN-AU - Genome Research (AT)	An evaluation to determine whether the genome research programme, planned for a period of nine years, in its present form, should be continued	Mid-term		√		√
Innovation Consortiums (DK)	Analysis of the economic impact of this flexible framework for collaboration between companies, research institutions and non-profit actors	Mid-term	√	√	√	√
Pôles de compétitivité (FR)	The assessment covers both the national and a detailed evaluation of each of the 71 clusters	Ex post	√	√	√	√
Scottish cluster Strategy in Biotechnology (UK)	Assess the effectiveness of a number of policies implemented in Scotland to create a biotechnology cluster	Mid-term	√	√	√	√
Basque Cluster Policy (ES)	The evaluation addresses the effectiveness of cluster policy. It makes use of the representativeness of the beneficiaries of the policy cluster in the Basque Country on the one hand, of a comparative analysis of results of competitiveness of beneficiaries and non- beneficiaries of the policy,	Mid-term	√	√	√	√
VINNVÄXT (SE)	Investigate the performance of the three winner regions in relation to each of their action plans	Mid-term			√	√

The majority of the reports present mid-term evaluations of cluster programmes, which range from targeted measures such as the Scottish Cluster Strategy in Biotechnology to much broader and inclusive measures such as the Review of the National Research Flagships in Australia or the Evaluation of the priority actions plan in the Walloon region.[41]

We can see that effectiveness, efficiency and sustainability questions are assessed in almost all programmes. Over 90% of the evaluations (11 out of 12) have in fact tackled these questions, albeit in somewhat dissimilar manners. The related issue of impact assessment was less widespread, reflecting the interim timing of the evaluations (too early to credibly test for impacts), with just seven of the studies having endeavoured to explore and dimension the actual or potential impact of the scheme in question.

Most of the evaluations looked at efficiency from an operational perspective (e.g. the implementation and development of the scheme, the application and evaluation process and operational performance), rather than value for money and the relative efficiency of a given measure's creation of impacts as compared with the equivalent productivity / value of analogous schemes in other regions or countries or alternative policy options. Relevance issues were addressed by 9 out of 12 evaluations.

5.4.3.2 The evaluation methods implemented

Figure 23 Methods used in the evaluation of Networks and clusters

	Qualitative methods	Quantitative methods	Approaches

[41] IWEPS; Institut Wallon de l'evaluation, de la prospective et de la statistique. (2009). Evaluation du Plan d'actions prioritaires pour l'Avenir wallon: rapport de synthese.

Name of the measure evaluated	Individual interviews	Case studies	Peer review	Questionnaire survey	Use of administrative/	Micro-/ macro-economic model	Bibliometrics	Baseline	Counterfactual	Benchmarking	Indicators
Collaboration programme (IE)	√		√	√	√				√	√	
The Centre of Expertise Programme (OSKE) (FI)	√		√	√	√			√	√	√	√
Walloon Competitiveness poles (BE)	√			√	√				√		
National Research Flagship Program (AT)	√		√	√							
Advanced Network and Research for Industry and Education (CANARIE) - ADDP programme (CA)	√	√		√	√						
Energy-cluster.ch (SW)	√			√	√						
GEN-AU - Genome Research (AT)	√	√		√	√			√		√	
Innovation Consortiums (DK)					√	√		√	√		√
Pôles de compétitivité (FR)	√		√	√	√					√	√
Scottish cluster Strategy in Biotechnology (UK)	√			√				√			√
Basque Cluster Policy (ES)	√			√	√			√			√
VINNVÄXT (SE)	√		√	√	√						

Although both qualitative and quantitative methods are evident in most of the evaluations, there is a slight preference for **qualitative methods** at least in terms of the number and balance of different methods and analytical techniques used (see Figure 23). In general, most evaluations focus on process and learning, reflecting their interim status. The methodologies are quite reasonably affected by the short timeframe between the start of the scheme and the commencement of the (mid-term) evaluation.

While there are clear differences in scope and methodological approach, all 12 evaluations include a formative element, identifying lessons learned and searching out opportunities for practicable improvements in focus or operation. Evaluations have been designed in such a way that the common designation of '**participatory evaluation**' seems to be very suitable.

The evaluation design is not rigidly imposed from outside, but gradually takes shape through the collaboration of all the stakeholders and their active participation in the analytical evaluation process. This participatory approach can be especially seen via:

- **Stakeholders' consultation**s have been used by 11 of the evaluations, during the early phase of a study and while the detailed study design was being finalised.

In general interviews have been designed to gather information on the operational and organisational matters of the cluster initiatives, its general functioning and success in relation to the attended results among stakeholder (member of the cluster organisation, regional and national government bodies, universities and others). Based on the agreed evaluations issues and questions, all interviews have used as basis clear interview guides.

- **Questionnaire surveys** have been widely used (11 evaluation reports), typically designed by the evaluation contractor but nearly always finalised in discussion with clients and steering groups. These electronic surveys were typically distributed to members within the cluster organisation.[42]

The starting point for most evaluations (10 evaluations) is a programme of desk research involving a review of programme documents on the one hand and the use of administrative data from programme monitoring on the other hand (involving analysis of planned and actual inputs and outputs; of funding and funding sources [inputs], funding destinations and activities, and standard outputs). Background literature on the programme design, its rationale, planning material, together with other background documents such project files and reports have been in most cases examined and incorporated in the evaluation work. Both mid-term and ex-post reviews have assessed statistical data on beneficiaries provided by the corresponding administrations. In addition to the available ex-ante documentation and administrative data, additional suitable documents have been utilised (e.g. previous workshops results, previous schemes evaluations, etc).

While questionnaire surveys and descriptive statistics using survey and/ or programme documents and monitoring data have been extensively used, other quantitative methods have been rarely employed.

The complexity to quantify certain impacts and difficulty to establish the causal relationship between specific policies and macro-level business performance are most likely to be the justification explaining the low utilisation of these evaluation techniques. The analysis of the effectiveness of the policies by means of indicators and econometrics (1 evaluation) and bibliometrics (0 evaluations) have been rarely used. Several of the mid-term evaluations (i.e. the assessment of evolution in terms of macro-economic impact of the Walloon Competitive Poles) have stated that quantitative approaches – to impact assessment – are not feasible or at the very least too premature. 'An Analysis of Firm Growth Effects of the Danish Innovation Consortium Scheme' evaluation presenting an analysis of the economic impact of 'Innovationskonsortieordningen' (Innovation Consortium scheme, IC scheme) on participating firms in terms of growth and value creation is the exception.

Cases studies are equally uncommon: the CANARIE programme in Canada and the GEN-AU in Austria are the only two evaluations where cases studies were conducted (three and four case studies respectively). The large number and variety of beneficiaries found across the different programmes, and the difficulty to single out a representative of the group could be one of the reasons explaining the modest use of this methodology.

As part of the evaluations, other evaluations methods such as a separate assessment carried out by a set of experts – peer review- and baseline and performance target have been relatively used (both in 5 evaluations).

[42] These are not the only participatory approaches found. For instance, the Review of the Canadian National Research Flagships has organised laboratory visits and the Evaluation of the Cluster Policies in the Basque Country a focus group.

5.4.4 The advantages and limits of the methods used

None of the evaluations disclose the budget for the exercise, and only one indicates the duration of the study (five months)

The advantages of the typical methodological design are twofold:

- The participatory approach is seen as an effective way to assess cluster policies, which tend to involve many different stakeholders (and sponsors) in a wide variety of formal and informal interactions that unfold over many years. The fluidity and variability of this somewhat organic entity does appear to require a less technocratic approach to evaluation. On the one hand it allows actors in the programme to make the actual evaluation process and the results their own, transforming the evaluation into a learning process, the results of which, in a certain sense, belong to them. On the other hand, it considerably increases the probability that the results achieved by the evaluation will be used in an effective way to improve the policy or measure
- Despite the modest use of analytical tools, the methods used have been mostly adapted to the depth of the desired information along with the use of several sources in parallel (background information, data on beneficiaries, telephone surveys). Of particular note is the approach used in the summative Evaluation of Canarie Phase 3, where the evaluation issues were determined in consultation with the client prior to the commencement of the study. These predetermined questions have guided and structured the entire evaluation methodology

The limitations of the 12 evaluations of clusters are fourfold:

- **Periodicity and follow up**: the studied evaluations tend to be carried out within three to five years of the launch of the scheme, when few major effects on the cluster are likely to be detectible, especially if the cluster is an emerging one. Indeed, the ultimate success of a regional cluster policy may not fully reveal itself for 10 or even 20 years, which is clearly too long to wait to carry out any useful evaluation! Rather cluster policies should be assessed on a more permanent basis, tracking initiatives over time in order to reveal the *intermediate* effects and facilitate strategic management rebalancing and increasing support as necessary. In essence, evaluators help administrators to learn from the initiative so that actions can be taken to improve the instrument and the way it is carried through (Sölvell, 2008)
- **Minimal use of analytical tools**: while an emphasis on qualitative research methods and participatory approaches is appropriate here, the limited efforts to carry out substantive impact assessments does seem to be a weakness that policy makers and financiers may object to. The challenge is one of scale, wherein cluster programmes typically have a size and complexity beyond the reach of a normal, one-off evaluation, and yet these schemes are rarely so large that one can expect to detect changes in macro regional statistics
- **Lack of creativeness**: we consider that evaluations should, when possible, be as creative as possible. Evaluators should use their imagination and technical expertise in trying to identify new designs for evaluation that prove most suitable for the characteristics of each policy, the socio-political context and the needs of different stakeholders. A redundancy in the use of interviews and surveys and the very little use of some relatively new techniques, as for instance, network analysis techniques, show the 'straightjacket' and sometimes imperfect methodology used during the evaluations. The elaboration of conclusions merely based on the impression of beneficiaries and stakeholders can certainly lead to inaccurate and/or bias conclusions, not showing the real impact of the programme or initiatives in a given territory. We consider that evaluators should use the models and techniques that adapt best to each situation, regardless of whether they are of a quantitative or qualitative nature. For instance, social network analysis (SNA) ought to be a powerful technique through which to establish a baseline view of

regional relationships and how those linkages and interactions develop during the life of the cluster scheme. Surprisingly, the GEN-AU evaluation is the only evaluation that used this methodology[43]

- **Nothing on international comparison:** only three of the 12 evaluations made any substantive use of international comparators and benchmarking analysis more generally. Learning from peers internationally by introducing methods for international benchmarking could bring great value added to the evaluation methodologies[44]

5.4.5 Identification of good practice

The evaluation of the **Centre of Expertise Programme** in Finland is perhaps the most thoughtful and wide-ranging study out of the 12. The evaluation starts with the initial baseline assessment, from which the programme performance can be tracked back over time. In addition to the baseline scenario, both the performance and business environment data are used to assess the impact of the scheme in four concrete areas (expertise and innovation, business growth and competitiveness, internationalisation, and coordination, management and networking). Performance data includes indicators such as employment, productivity, number of establishment, patents, etc; where business environment data includes indicators of level of expertise, levels of networking and internalisation. Data is collected and maintained in the so-called ' **OskeNyt database**[45]' developed by the Innovation environment team at the Ministry of Employment and Economy[46]. This valuable database is updated annually and is used as basis for evaluation studies commissioned from external consultants.

A second example of good practice is the independent mid-term evaluation of the '**French Competitiveness Clusters'**[47] programme, which encompassed both global policy and each cluster individually[48]:

- Dedicated means, consistency with other public policies (R&D and innovation), cluster selection process, financing support processes, policy management at national and local levels, synergy between actors, first effects on local actors were evaluated on the policy level.
- In the evaluation of individual clusters, the following points were taken into consideration: economic and international strategy; cluster government and animation, evolution of the cluster population; R&D projects and firm-public research-training synergy; territorial settlement and network strengthening,

---

43 For additional information on the Social Network Analysis carried out in the context of the GEN-AU evaluation see next section on identification of best practices.

44 Most of the above-mentioned limits can be tracked back to the very substantial effort - and specialist skills – required to set up monitoring systems that are sufficiently well targeted and comprehensive in their reach (and frequency of report) to provide an adequate platform for robust evaluation. Moreover, good monitoring in such systems-based interventions demands universal support among stakeholders, which in turn demands a seriousness and sensibility around programme commitments that many political authorities struggle to sustain across the necessary time-horizons. Making up for poor monitoring and evaluation systems – after the fact – through one-off evaluations is difficult, time-consuming and costly. Mostly, therefore, agencies will prefer simply to press external evaluators to do the best they can within a conventional specification and funding envelope.

45 For additional information on the OskeNYT database, see http://www.tekel.fi/in_english/tekel_network/tekelnyt/

46 Pelkonen, A., Konttinen, J., Oksanen, J., Valovirta, V., Boekholt, P., & Levasluoto, J. (2010). Osaamisklusterit alueiden voimien yhdistäjänä Osaamiskeskusohjelman (2007–2013) väliarviointi. Ministry of Employment and the Economy.

47 The 'competitiveness poles' concept has been most strongly developed in France from the mid-2000s and was quickly 'replicated' in countries ranging from Belgium (Wallonia region), Greece ('regional innovation poles'), Hungary, etc. Equally, competence centres have become a popular tool across a number of EU countries during the last decade (Compera 2010).

48 BCG, CM International. (2008). Evaluation des pôles de compétitivité: Synthèse du rapport d'évaluation

including structural projects; SME integration and new enterprise creation; Human resources training; and Green development approach.

The evaluation comprised of performance data such as employment concentrations and patents, business environment including levels of networking and cluster governance, and finally descriptive data on member list and company size. All data was gathered trough surveys and interviews on the competiveness clusters[49].

Finally, the evaluation of the **GEN-AU programme** includes a meticulous SNA analysis, which revealed:

- The quality and intensity of co-operative relationships, information and knowledge exchanges, and the exchange of resources between the various GEN-AU projects and their casts of actors
- The identification of strategic and peripheral actors and/or projects, and the pinpointing of factors critical to project success or failure, thus enabling the identification of the innovation potential of co-operative research
- Insight on the functionality of the GEN-AU network enabled the identification of strategic fields of action and offered ways to optimise the programme management

[49] On the basis of positive results of the evaluation, the French government decided to launch a second phase of the competitiveness cluster policy (Cluster 2.0) for a further 3 year-period (2009-2011) with a total budget of €1.5 billion.

6. Conclusions and recommendations

6.1 Overall

Overall, the literature review suggests that, currently, evaluation methods are only partially determined by the particularities of the innovation measure under review.

The great majority of the evaluations appraised by this study, address themselves to the same central questions applicable to any policy evaluation – relevance, efficiency and effectiveness – and most of them deployed a broadly similar core methodology comprising:

- Desk research to test the alignment of the scheme (volume / shape of investments, activities and outputs) as compared with the strategic plan, and to gather definitive statistics on scheme inputs and outputs for incorporation with the subsequent value for money calculations
- Stakeholder interviews to explore opinion on the continuing need for such a measure within the country or region in question, in light of wider developments (events) and more recent policy initiatives (complementary schemes)
- A questionnaire survey to obtain semi-quantitative feedback on the administration and efficiency of the scheme in question and to detail and possible dimension the attributable benefits, social and economic

This is an over-simplification of course. There are plenty of exceptions within the portfolio of evaluation reports that we have gathered together, where one or other of these aspects is missing from the study design. However, these appear to reflect a conscious decision to move quickly and efficiently to answer one question: a strategic review for example, which might focus on the stakeholder interviews; or a pilot evaluation, which might emphasise the desk research and questionnaire survey. Equally, this triptych is an attempt to characterise a core methodology, where many studies deploy two or three other data collection methods or analytical techniques.

We can unpack this over-simplification at two levels, between those measures falling at either ends of the science and innovation spectrum and at the level of the individual innovation measures.

There is something of a split evident in the overarching study design between two clusters of the innovation measures under review, which one might loosely describe as the science and the innovation ends of the innovation support spectrum. In simple terms, the measures that support (pre-competitive) research within the university sector, whether that is strategic research programmes or competence centres, are narrower and more homogeneous in methodological terms: qualitative research methods predominate and the evaluation questions revolve around effectiveness (research quality and community engagement) rather than relevance or efficiency.

By contrast, the innovation end of the spectrum – with its support for the proprietary activities of large numbers of actors – is much more focused on quantitative research methods and economic impact in particular. This second (and much larger) cluster of evaluation reports is also more likely to devote a special effort to researching the net benefits directly attributable to the policy support and wider economic impacts (through the economic multipliers of wages and purchases). Control groups and simple econometric techniques are also very much in evidence. Questions about efficiency are nearly always prominent too, both operational efficiency (service quality) and overall efficiency (value for money of this scheme as compared with any practicable alternative policy option).

Figure 24 Differences in evaluation across science and innovation

Support for non-competitive, strategic research	**Relevance** is not a major focus for many of these studies, however where it is discussed it is usually a question of reviewing stakeholder's current views on the continuing strategic importance of a given topic or theme **Efficiency** is more narrowly concerned with management efficiency, rather than value for money. In a minority of cases, there is also a question of sustainability, as regards the likelihood that a new centre for example might ultimately become self-financing **Effectiveness** is most often concerned with what might be called intermediate outcomes, which is to say the quality and international standing of supported researchers and research (and industrial engagement to a lesser degree). Qualitative research methods predominate, and peer review (panels of international academic and industrial experts) still sits centre stage for many evaluations, albeit expert judgement is almost always informed by a comprehensive mixture of contextual and operational statistics In addition to peer review, there are several **techniques** more in evidence here than in the proprietary innovation support schemes. Social network analysis appears to be emerging as a fashionable new tool, albeit with a deal of uncertainty as regards the calibration or interpretation of results from what is a relatively novel technique. By contrast, bibliometrics is used more widely and with greater confidence (in particular benchmarking a centre or region's performance against citation levels in the field). Several evaluations count different forms of IPRs, from invention disclosures to patents granted, but none of the studies had managed to normalise these data using EPO statistics for example
Support for proprietary research and innovation	**Relevance** is more of an open question as regards the needs and inclinations of any local business community, and studies may devote quite substantial efforts to exploring the need for and appropriateness of the proposed measure. Data collection may comprise desk research to map innovation activity or the provision of support as well as consultations and surveys to test opinion on both supply and demand sides **Effectiveness** is concerned primarily with determining programme impacts and in particular net economic benefits and internal rates of return (i.e. the net present value of the short and medium term benefits attributable to the public investment). There is a clear focus on quantitative research methods and economic impact in particular. This second (and much larger) cluster of evaluation reports is also more likely to devote especial effort to researching the net effect of the scheme evaluated (netting off any improvements that might have occurred in the absence of support to arrive at a sharper view of the benefits directly attributable to the policy measure) and wider economic impacts (through the economic multipliers associated with increased employment / wages and increased purchases within the region or country). There is some interest in determining intermediate effects, for example the behavioural additionality of a particular measure although this tends to be addressed through a single opinion survey and with no attempt to establish the persistence of such changes Questions about **efficiency** are prominent too, both operational efficiency (service quality) and overall efficiency (value for money of this scheme as compared with any practicable alternative policy option). The importance of testing for the counterfactual means that control groups and simple econometric **techniques** (difference-in-difference methods) are also in evidence, as is the use of input-output statistics to estimate wider economic benefits (beyond the beneficiaries). The current methodological battleground is really concerned with the issue of knowledge spillovers, which by definition happen outside the immediate beneficiaries. Micro-economists are exploring the power of combining in-depth impact case studies with broadband surveys to profile awareness of new developments while macro-economists have been attempting to improve their ability to model the effects of changing compositions of R&D investments and labour markets

The following sub-sections unpack these heavily digested points and offer a view on the current state of the art as regards evaluation methods and the innovation measures studied. In the first instance, we have tabulated the main methods and tools identified through the literature review and elaborated on their applicability to the innovation

measures of interest and their particular strengths and weaknesses from an innovation evaluation perspective. This horizontal analysis is then followed by a more in-depth assessment of the tools and methods from the perspective of individual groups of innovation measures.

6.2 Horizontal analysis of evaluation tools and methods

Figure 25 Analysis of evaluation tools and methods and their advantages and limits for the evaluation of innovation support measures

Tools	Description	Application	Strengths	Weaknesses
Desk research	Logical framework analysis, to recreate the intervention logic in order to test its coherence / relevance and to define the key dimensions and metrics to be studied through the evaluation Compilation of financial and other monitoring data (activities, outputs) to analyse actual performance against planned performance Compilation and analysis of secondary data as defined by programme or logframe, to gauge trends in key reference statistics (e.g. regional BERD)	Desk research is used in the very great majority of all evaluations of all kinds of innovation support measure examined here	Robust methods and tools with good provenance LFA is a disciplined approach to definition of appropriate performance measures Permits one to update and align key tests with changed circumstances and the reality on ground Programme data permit objective analysis of Finances, Activities and Outputs Secondary statistics provide an important reference point, with respect to the wider effects of a programme	No significant weakness. Better at recording / analysing inputs and outputs, and less good with outcomes and impacts Getting to grips with outcomes is dependent upon budget holder having had the foresight to establish baselines and gather the full spectrum of data needed through their standard formal monitoring systems Not easy to secure buy-in to newly-defined (unforeseen) tests amongst beneficiaries and other stakeholders
Questionnaire surveys	A bespoke set of questions directed to target populations, typically the recipients of public assistance (the beneficiaries) Usually delivered by telephone, rather than post or online, to maximise response rates The questions will typically profile respondents as well as gathering facts and figures on outputs and outcomes and taking opinions on the attractiveness of the measure and the quality / efficiency of its	Surveys are a centrepiece of the data collection strategy of more than 90% of the evaluations examined here, and are clearly appropriate for use in the assessment of any of the types of innovation support measures under review. Large–scale surveys are widely used in the evaluation of proprietary innovation activities (direct financial support, innovation management support, etc) as a way to collect the view of a large number of firms having benefited from the support measure. Smaller-scale surveys are preferred to evaluate science-based innovation support	Robust methods and tools with good provenance Bespoke. One can design questions befitting the intervention logic (programme theory), where official surveys / statistics might miss the core objectives. One can also invite beneficiaries to document achievements, where official data will almost certainly be too aggregate to reveal any programme-derived change Efficient. One can collect large amounts of highly relevant facts, figures and opinions,	The principal weaknesses are the predominantly subjective nature of the data and the risk of response biases Quality of questionnaire design, which is tough to get right even with piloting. Single snapshot survey deals poorly with before and after. Can become burdensome to administer if one has to commit to maintaining periodical surveys

Tools	Description	Application	Strengths	Weaknesses
	delivery Most questionnaires used a majority of closed questions to permit a degree of quantification and descriptive analysis, and also cross-tabulation Most Qs will also include a small number of open questions too, usually forward looking questions (e.g. what one practicable improvement would you recommend the service consider implementing?)	measures (Strategic research, Science-industry cooperation, etc), where the focus is more qualitative and fewer beneficiaries are targeted.	cheaply and quickly. One can target recipients rather than address questions to very large numbers of people and organisations with no knowledge of a scheme	Response rates among non-beneficiaries will always be low, and rates among beneficiaries collapse with repeat surveys beyond life of support Deals very poorly with the counterfactual
Interviews	Interviews usually take the form of a structured conversation with a pre-defined actor (within a group of actors or stakeholders), covering a set of pre-defined broad themes that tie back closely to the core evaluation questions. While a questionnaire or structured interview has a very formal question set, a semi-structured interview is more flexible, inviting contributors to talk at some length around each of say 5-10 broad issues, in order to allow discussion partners to explain / elaborate on a given phenomenon, and also allowing new questions to be brought up during the interview as a result of what the interviewee says. The specific topics that the interviewer wants to explore during the interview should usually be	Interviews are a core data collection method, used in every one of the evaluation reports analysed in this study. They tend to be used sparingly, to help to launch studies and develop the more structured data collection tools and to consult stakeholders on the continuing relevance of a measure and its coherence with other schemes and strategic developments. They are also used in many cases to probe issues revealed by analysis of questionnaire data. In most of the evaluation reviewed here interviews target stakeholders in implementing authorities, beneficiaries; and depending on the level of analysis funding bodies and non participants. Increasingly interviews are carried out by telephone and supplementary email conversations, with only the most critical interviews dealt with face-to-face (for	Interviews are especially good at allowing respondents to reveal issues and explain matters that one might not have anticipated Flexible structure enables interviewees to respond in own terms, and interviewer to respond as part of a two-way conversation Framework ensures comparability of interviews Seen as most appropriate when unravelling diverse layers and subtle nuances	The single biggest shortcoming with interview in the context of innovation evaluation is the time-limited nature of most people's memory. Interviewing people 5-years after the fact can be problematic and a source of bias Carrying out semi-structured interviews is non-trivial, in that there is a need to allow respondents to develop a narrative while probing and checking their accounts and also covering all of the evaluation questions. Must no pre-empt answers and must cover all issues equally thoroughly They are quite labour and time-intensive, when done properly (collection and analysis) As with the questionnaire survey, interviewees may themselves be unaware of indirect influences of

Tools	Description	Application	Strengths	Weaknesses
	thought about well in advance. It is generally beneficial for interviewers to have an interview guide, and for the questions to be shared with interviewees ahead of the discussion The topics provide the structure necessary to synthesise feedback from multiple interviews, with different groups of actors or stakeholders, and thereby discern any strong patterns in opinions / experiences	reasons to do with timeliness and cost).		research
Case studies	An empirical approach to documenting / describing a current phenomenon within its real-life context, where the boundaries between phenomenon and context are not clearly evident and in which multiple sources of evidence are used (Yin, 1984)	Case studies were a component of the evaluation methodology in a minority of cases in each of the innovation support measures reviewed here. Case study methods were more common – but still not in widespread use – in the evaluations of direct financial support: and the innovation end of the research and innovation spectrum. They are used mostly for illustrative purposes, however in some instances case studies have been used alongside surveys or to support a peer review, to obtain a better balance of breadth and depth as regards the benefits data secured.	Excellent descriptive tool and powerful communication platform Can be used to build up multiple cases and thereby generalise and judge programme performance	The principal shortcoming is the extent to which a small number of cases can be used to generalise and judge a programme's performance This is compounded by the challenge of designing the right sample frame and thereby identifying and researching the right cases They are very time and labour-intensive to prepare, consistently and rigorously
Expert panel (also referred to as 'Peer	Peer review refers to a process	A modified form of peer review – what the	Powerful and credible means by which to	Its principal weakness is the

Tools	Description	Application	Strengths	Weaknesses
review')	wherein one or more domain specialists – experts – will critically appraise a project proposal, scientific article or institution in terms of its originality and excellence	European Science Foundation (ESF) calls informed peer review[50] – was used commonly in the evaluations of research and innovation programmes with a strong academic component (industry science cooperation / strategic research). In this context, the peer review process typically involved an international panel (5-10 people) from several countries and several disciplines and sometimes from the public and private sectors (business). Peer reviews are first and foremost used in the evaluations of science-based support measure, targeting the quality of research. In almost all cases, these panels were provided with substantial prior information on strategy, investment, outputs, context and so on. In some cases, the unit under review was in charge of collecting and preparing the data (participatory approach). Experts were also typically given 5-10 evaluation questions or criteria to work with, and their visits to institutions were organised in such a manner as to permit all of the issues to be covered.	engage with deep specialists working on complex issues at the very frontier of knowledge Incorporates an intrinsic normalisation process, via the international and sectoral experience Reviewers are usually a good source of insight and wisdom as regards issues to address going forward, and possible solutions	subjective nature of the approach, and the anxiety that judgements reflect panel membership at least as much as they reflect the people and institutions being evaluated It struggles to cope with whole research systems, at least not without incurring very large costs and time penalties Its contingent upon the quality and comprehensive nature of the evidence prepared in advance of the exercise
Econometrics	The statistical analysis of micro-economic data obtained through monitoring data and surveys and possible official sources (e.g. annual business enquiry) Most often addressed through a	Evaluators will in some cases attempt to arrive at an aggregate figure for programme impact using various economic modelling techniques. In its simplest form, evaluators will take the directly observed estimates of programme benefits and adjust these using coefficients (derived from very large,	Cost-effective means by which to arrive at a reasonably good estimate of direct benefits attributable to a programme For individual programme evaluations, macro-economic modelling is almost certain to be too aggregate but might be	Does less well with indirect benefits, such as knowledge spillovers, so might understate Does less well with issues like displacement

[50] The ESF maintains comprehensive guidelines for public bodies wishing to conduct peer review, which can be downloaded at www.esf.org/activities/peer-review/

Tools	Description	Application	Strengths	Weaknesses
	combination of methods, typically beneficiary surveys run in parallel with surveys of a control group, asking the same questions in order to be able to estimate additionality	exhaustive evaluations of analogous interventions) for a number of more general effects, including deadweight (the proportion of the benefits that may have arisen without the intervention), displacement (the proportion of the additional benefits accounted for by losses among competitors in a given territorial space) and economic multipliers (the additional benefits to the wider economy that result from increasing output [industry purchases] and increasing wages. In some cases, evaluators will use input-output data [reflecting the structure of the regional or national economy] to better estimate macro-economic effects. In a very small number of cases, evaluators have used quite sophisticated input-output models to estimate flows of new knowledge and technology through an economy in an effort to simulate the effects of knowledge spillovers, using small numbers of exhaustive micro-economic case studies to derive the basic input data to feed into the modelling process.	applicable for collected investments across a programming period	Does less well with more intangible gains
Social network analyses	A social network is a social structure made up of individuals (or organisations) called "nodes", which are tied (connected) by one or more specific types of interdependency, such as a project partnership or IP agreement. Social network analysis (SNA) maps these different nodes and ties within a given population, and can be used to trace the evolution of those connections – density – and the importance of individual actors	Social Network Analysis was used in a very minority of the 58 evaluation reports reviewed here. It was only used in those measures that had both university and business participants, however SNA as a technique is developing rapidly and there is no obvious reason why administrative data (on e.g. partnerships and subcontracts) and survey returns cannot be used to map the evolution through time of networks of businesses and other non-public research organisastions.	Powerful means by which to map changing relationships between individuals or organisations over time	The central weakness of SNA is arguably a function of its novelty, inasmuch as people struggle to infer meaning from the complex maps one generates The other challenge is to obtain comparable contextual data – before, during and after, with which to test for the importance of external factors (counterfactual analysis)

Tools	Description	Application	Strengths	Weaknesses
	within a network - centrality			
Bibliometrics	Bibliometrics is a set of methods used to count numbers of research outputs (e.g. refereed journal articles) and citations to those publications.	Bibliometrics was not widely used in our sample of evaluation reports, although it was used in several of the strategic research programmes as a way to assess research quality. Bibliometrics used refereed journal articles is an increasingly sophisticated technique, which is widely used by evaluators to compare a research scheme's performance on productivity and quality with international norms for a field (e.g. regional performance compared to world average or double world average citation rates for the field in question), for the main supported disciplines, nationally and internationally It can also be used – via the analysis of the organisational affiliations of co-authors – to identify and trace trends in academic industry cooperation	The main strength of journal-based bibliometrics is the very large databases of historical material held by international databases like Thomson Reuters or Scopus, which permits regional agencies to rigorously establish their local researcher's international standing (in almost all fields)	The main weakness from a resaerch and innovation standpoint is that the process is really only relevant to academic work Industry-academic links can be traced, but that is usually a bespoke requirement and quite resource intensive The same would hold for other forms of bibliometrics, for example tracing policy impacts in discussion papers, studies and communiqués It has no relevance or credibility with respect to innovation support more generally
Patent analysis	Patent analyses tend to involve measures of quantity rather than quality, although working with USPTO and triadic patents (EU, US, Japan) can be a useful proxy.	We found no instances of evaluations that had used patent analysis as part of the study design. Given the widespread nature of the ambition to increase rates of technological innovation within innovation programmes, this seems somewhat surprising. However, in several cases, surveys had attempted to count patent numbers attributable to assistance.	Patent applications and patents granted amount to one objective measure of a certain kind of technological outcome With the availability of large-scale online patent databases (e.g. esp@cenet, USPTO) it is possible to examine connections between partners and their prior art. It is also possible to explore a programme's contribution to a technological field more generallyAcademic groups and patent specialists are developing software tools to begin to analyse patent meta data in much larger volumes, permitting easier analysis of technological trends and unfolding relationships within communities	The principal shortcoming with patent analysis is the uncertainty over its significance (of an application), which is a long way from being an innovation Uncertainty and time-lags notwithstanding, the analysis of patent data and their bibliometric information in particular is labour intensive / costly and time consuming

6.3 Innovation management support and dissemination

While one might imagine that innovation management support is the most widespread support measure in the policy makers' portfolio, its almost universal provision and focus on rather small pieces of support (information, advice, training) means the methods used are about as mainstream as one can get.

There seem to be three central lines of enquiry for this class of innovation support measure:

- Comprehensive monitoring data are commonplace and permit quite robust analyses of the marketing and communication of schemes, audiences attracted and customer satisfaction. They also provide good before and after data – semi-quantitative and suitable for statistical analysis – which help to detail the behavioural changes amongst assisted businesses with respect to their changed outlook / awareness and confidence to pursue innovation projects beyond the support. It is unusual for agencies or their service providers to follow-up with regular surveys of their clients in order to determine the extent to which behaviour does change in practice and also persist. This is an approach that might be considered, however there are costs and risks attendant
- The large populations of assisted businesses, does lend itself to large-scale questionnaire surveys, and control groups, in order to run quite sophisticated statistical analyses in order to dimension the aggregate increases in economic output (GVA) and employment. The control groups are especially important with these kinds of smaller interventions
- While several of the evaluations work hard to estimate the wider (indirect) benefits in the form of different economic multipliers, the studies devote relatively less attention to the effects on innovation behaviour or innovativeness more generally

Likewise, the monitoring data and surveys provide a good platform from which to rate the quality, speed and suitability of the services on offer and the associated delivery arrangements, as viewed from the perspective of the target audience (i.e. the fraction of the SME population that has ambitions to grow and yet is not especially innovative and has limited absorptive capacity or wider social networks).

6.4 Intermediary bodies and agencies

A majority of the evaluations of intermediary bodies assessed the **relevance** of the measure, however the methodologies were quite generic and comprised desk research cross-checked with a series of individual interviews, linked in several cases with a site visit. The specific relevance questions comprise tests of coherence (alignment of activities with objectives and planned activities) and added value (need for a measure of the kind in question, which his consistent with the wider innovation policy landscape). For example, in the case of business incubators the evaluators need to find out to what extent the incubator tenant characteristics match the definition of target market and admission criteria.

Questions related to **effectiveness** of the schemes focus primarily on how effective the programmes were in meeting initial objectives. Evaluations of business incubators analysed foremost the survival and growth of the incubated businesses (through monitoring their revenue, employment, profits, exports, etc.). Science and technology park evaluations similarly look at their effectiveness in terms of indicators relating to firms using their facilities (turnover, accessing finance, employment growth and in launching new products etc.). In some evaluations of support for technology transfer programmes was effectiveness simply assessed by measuring perceived effectiveness (asking the beneficiaries whether the projects provided good value for money). In just

one case, the effectiveness of the scheme was compared with international benchmarks.

In evaluations of schemes for support of business incubators, **efficiency** tends to be evaluated by comparing expenditure and budget handled by the management, and by assessing the execution and management of the programme (in planning and design vs. implementation, resource allocation, balance of staffing, organisation and work processes against programme objectives). Similarly to effectiveness, efficiency-related performance of business incubators can be compared to international benchmarks, provided such data is available. For schemes of science parks tends this question to be asked directly to the key stakeholders and tenants of the science park. Programmes for support of technology transfer within our sample did not generally evaluate the efficiency of the intervention.

Impact was a question, which was absolutely central to some of the evaluations. This is partly because the studies were meant to inform the decision-makers on future developments and provide evidence for justification of continued support of the schemes. The methods that the evaluators used varied from a treatment that did not go far beyond a question in a questionnaire to a more complex multi-criteria analysis concluding with direct and indirect economic impacts for 3 different scenarios, using a combination of primary and secondary data sources. This kind of analysis is only possible where the sample is large enough and enough time has passed for the firms to be showing signs of commercial success. Both of these determinants have to be met in case of performing an insightful impact assessment and should be considered by the evaluator (and the evaluation unit) at the initial planning stage of a study. The report in our sample that did use economic impact analysis has, in order to ascertain the level of programme added value, endeavoured to use two control groups: companies that were accepted for incubation but chose not to enter an incubator and incubated companies that left before graduation without the mutual agreement of the incubator and company. Overall we can conclude that in this innovation measure category is such analysis challenging and relatively rarely provides insightful conclusions due to problems of reaching statistically significant results. Combination of empirically collected data and existence of comparable secondary data is therefore inevitable.

Notably, the evaluations of business incubator schemes assessed one additional aspect, the **future sustainability** of the schemes. Financial sustainability of an incubator is seen as a measure of success and was a subject of international comparison between programmes. Many of the sub-questions and indicators related to future sustainability are related to scheme's efficiency and cost-effectiveness and were therefore addressed together.

All of the evaluations included a set of recommendations, based on their findings. The recommendations ranged from opportunities for improvement of design and delivery to development of new indicators for better monitoring of progress.

6.5 Start-ups and spin-offs

On **relevance**, the evaluation reports suggest that a well-run 'startups' scheme would provide evaluators with the results of the original baseline exercise, which had researched the funding landscape and determined that there was indeed a funding gap and that this was impeding the emergence and growth of new, high-growth businesses.

The subsequent evaluation of 'relevance' would look at that evidence base and essentially re-run the exercise to confirm the current situation: is there still a funding gap; to what extent has the provision of public support corrected for that shortfall; and what evidence is there that the market might improve the availability of risk capital if public support were to be reduced or halted?

Desk research to map and profile available funds, within the region and beyond, in order to understand the nature and extent of provision and the existence of a funding gap in respect to medium-scale risk capital (EUR 1 million) available for higher risk, early stage investments. Ideally, the review of the funding landscape should also be

complemented by a demand side assessment as regards the nature and extent of funding requirements and in particular where and on what terms start-ups have managed to secure funding.

The objectives of the programmes supporting start-ups and spin-offs tended to be couched in terms of providing finance to those early stage companies that have been identified as having definite potential to grow and make a profit and enabling them to do so. In most cases, we found that the specific objectives of the programme being evaluated were not SMART (Specific, Measurable, Achievable, Realistic, Time-bound). One of the reasons for this could be that it is a feature of early-stage capital investments because they are known to give notoriously skewed returns, with most projects failing and some giving very high returns. This makes it difficult to predict how many firms will succeed in order to set this as a realistic objective. In general, the way that **effectiveness** was evaluated was to look at how many firms had become successful start-ups and interpret this without the benefit of a benchmark to compare this to. This meant that there was some overlap in the way that effectiveness was assessed and the way that efficiency and impact were assessed. Another difficulty with estimating how many of the firms invested in were successful was the famous 'J-curve' effect, which argues that firms tend to make a smaller profit after the initial investment while they are spending resources rearranging themselves, and then their profits grow much faster than if they hadn't received any funding. This effect means that most evaluations were technically done too early to tell whether the funds had been successful or not.

The way that **efficiency** tended to be evaluated built on the information compiled to answer the question of effectiveness. Because the programme has a direct financial output, an internal rate of return could be calculated. This involved taking the profits the early stage companies had produced, and calculating a ratio between them and the operating costs of the funds. An established finding is that the smaller the value of the funds being given, the higher the operating costs are to manage those funds, and so the lower the efficiency tends to be. This does not necessarily mean that larger funds are better, but is an issue the evaluator needs to factor in.

Impact was a question which some of the evaluations looked at in great depth. This is partly because the programmes have a financial output, and so the question of impact is already partly answered by looking at the question of effectiveness. The methods that the evaluations used varied from a treatment that did not go far beyond how the question of effectiveness was answered to a full-scale econometric analysis. The econometric analysis was made possible where the sample was large enough and enough time had passed for the firms to be showing signs of whether they are going to be successful or not. Both of these issues are up to the evaluator's discretion. The report in our sample that did use econometrics to calculate impact (From funding gaps to thin markets) choose a range of early stage venture capital funds that were all substantially similar in the types of firms they supported and period they had been supported in. They then matched these against ten control firms, which were matched by size and age of the firm, number of employees, profitability and sector of the economy. This gave a total dataset of 7,741 companies.

The final question most of the evaluators sought to answer was the matter of the **future sustainability** of the funds that had been created with the support of public monies. In all cases, this question was addressed through a combination of qualitative research – did fund managers believe that similar volumes of deals would be made available for early stage investments beyond the life of the current publicly-supported initiative – and financial analysis (the internal rate of return of the current portfolio). Whether funds were invested in the form of a preferential loan or equity, there should be a sufficiently positive return made on those investments to retain the interest of the private sector and occasionally produce a sufficient rate of return that a surplus is available that can be reinvested into the funds wherein the funds can tend towards being self financing, for some time at least, without further injections of public cash or guarantees.

6.6 Science-industry cooperation

The notion of science-industry cooperation encompasses both means and ends, with certain SIC schemes encouraging interaction in order to accelerate technological breakthroughs in areas of strategic importance, while other schemes are concerned almost exclusively to bring about a step change in the nature and extent of interaction among the two communities in general. The former often involves major capital investment and the creation of semi-permanent structures that bring together established partners to collaborate on a specific programme of user-oriented research, which have a great deal in common with strategic research programmes discussed elsewhere in this report, while the second group provide very much smaller packages of support – financial incentives – in an attempt to change the habits of occasional or non-collaborators, permanently.

The evaluations of university-industry technology centres tend to have study designs that echo methodologies one might find in evaluations of applied research programmes more generally. There is a focus on research quality – the international standing of people and research outputs – and reliance upon peer review. The peer review process stands apart from more academic peer review, inasmuch as the panel combines both eminent academics and leading industrialists, and there are substantial amounts of contextual and performance data fed into the process as well as a more closely prescribed set of evaluation questions. Given the experts tend to be outsiders that cannot (by design) benefit directly from the programme under review, many experts will be paid a fee or honorarium.

This form of modified peer review is a good, practicable means by which to judge the efficiency and achievements of collaborative research structures, that are rather complex and long-lived entities and which ought to produce manifold benefits across may years and in many locations through hard-to-observe knowledge spillovers. Relevance issues are similarly addressed through the prism of the peer review process, while the issue of effectiveness tends to be concerned primarily with the achievement of a critical mass of people and work of international significance. Certain regional schemes may work with a slightly lower quality threshold; however there is a general presumption that research of average quality is much less likely to generate significant social or economic benefits (beyond the direct employment).

The international benchmarks seem to be a similarly practicable means by which to gauge the relevance and future development of these structures, particularly for the programmes that are rather recent at the time of the evaluation (e.g. Evaluation of the Estonian centres of competence).

When it comes to collaborative projects, the evaluation reports show a marked difference in their basic study design, and focus primarily on economic impacts and value for money. Research quality is almost never considered and the investigation of behavioural effects – and the persistence of those effects – tends to be treated rather lightly, with a series of questions in a questionnaire survey. There are however two exceptions to this rule: the evaluation of the UK Knowledge Transfer Partnerships and the evaluation of the Christian Doppler Research Association, which both explore behavioural effects in more depth and through each of their primary data collection methods addressed to enterprises and academic partners.

It was a surprise to discover that none of the evaluations included within the set of science-industry schemes had used bibliometric analyses (e.g. co-publication of papers by authors at academic and business addresses) or social network analyses (e.g. an evolutionary analysis of the number, density and centrality of connections between academic and business partners within the schemes). The study team believe these would be desirable in evaluations of these sorts of measures: the former provides a very good source of data to help gauge trends in visibility and research quality, internationally normalised, and also trends in co-publications with industry. Network analysis, by definition, permits one to map certain types of relationship among communities and across geographies, while also tracing the evolution of those relationships through time.

The Danish evaluation of Innovation consortia, using a combination of baseline, counterfactual and econometrics, is a good example of how one can look at the economic benefits – and wider effects – of an SIC innovation support measure (funding smaller, commercially focused projects).

6.7 Strategic research

Most of the evaluations within our sample looked at the **relevance** of the Strategic Research programmes they were judging. However, they tended to do this within the context of checking the funding landscape and checking with stakeholders that there was a clear and important 'gap' in the provision of funding for the topic in question. In this sense, the evaluations did not do a 'radical' assessment of the relevance of the funding – by going as far as examining whether the research area being funded really was 'strategic'. The answer to whether there was a clear gap in the provision of research funding was usually effectively assessed through a combination of desk research and surveys. The desk research could be used to do a descriptive statistical analysis to ascertain what the programme was actually funding. It could also be used to look at what other kinds of similar funds exist that might overlap with it. However, this tended to be done less frequently. The alternative was to ask the stakeholders affected by the funding programme whether they felt that the programme filled an important gap and whether they felt that the programme was still needed.

All of the evaluations in our sample looked at the question of **effectiveness**, and all but one looked at the question of **impact.** However, because none of the objectives were SMART (Specific, Measurable, Achievable, Realistic, Time-bound) the distinction between the two tended to blur. In general, the methods and resulting evidence collected were the same, but the kinds of interpretation they yielded varied depending on the questions asked of it.

The most common method used to collect evidence on effectiveness and impact was the questionnaire survey. The researchers were asked whether their work had yielded the expected results in terms of papers and outputs, and this was used to calculate the effectiveness of the programme. Respondents were also asked whether they believed that their work had become better since this new type of funding became available, and whether they found it easier to collaborate with other researchers and make use of any available technology platforms. Finally respondents, including all stakeholders, were asked whether they felt that the whole strategic research area had become a more cohesive and well-managed discipline, and whether it had become stronger internationally. This was used to feed into an assessment of the overall impact of the programme. Interviews were also used to add colour and depth to this kind of information.

Several of the studies under review included analytical techniques that are more commonly applied to evaluations of academic or science-industry collaborations, and would be much less likely to be used in the evaluation of more general innovation measures.

Bibliometrics was used to compare the research that had been funded by the programme in question to publication outputs and citation rates in equivalent disciplines within the same country and internationally. This enabled the evaluators to assess how effective the programme was at funding good research, and what impact the programme had on the international standing of that research.

Another technique that was used within our sample was Social Network Analysis. The data for this analysis was collected via a web-based survey, where researchers who had been funded by the programme were asked who they had collaborated with. This data was then transposed onto a map and used to identify how effective the programme had been in encouraging collaboration with other academic groups and businesses nationally and internationally. It was also used to give an indication of the quality and intensity of the relationships that had been formed and the extent to which ties with

industry had been formed. This fed into an evaluation of the impact that the programme has had on science-industry cooperation.

Only one of the evaluation reports in our sample looked at the question of **efficiency**. This was achieved by calculating a ratio between the programme's outputs in terms of papers produced against the costs of running the programme. The evidence was collected using both programme documents and a survey which asked respondents to quantify the benefits they felt the programme had produced. It is possible that the reason not many of the evaluations have included a question on efficiency is because the outputs of strategic research are hard to measure, as this example shows.

The **Future direction** of the programmes under evaluation was an important area of focus in all of the evaluations in our sample except one. This was reflected by the amount of expertise that was brought into these evaluations. Over half of them were written by, or had a major component written by peer review. As experts in their field, including top researchers and end-users of this type of strategic research, once they had analysed the evidence before them, they were in a strong position to recommend how the programmes should be improved. Another useful technique was to use benchmarking to compare the funding programme in question to similar programmes internationally. This was done in two ways. If an independent consultant wrote the evaluation, then the evidence used to inform the benchmarking exercise was usually compiled through desk research. If the evaluation was written by the peer review committee, then they usually brought their expert knowledge of the funding schemes in their own respective countries to the table and discussed the merits and shortfalls of the programme in their conclusions and recommendations.

6.8 Direct financial support for innovation activities

Direct funding measures tend to focus on one of two quite distinct challenges within the regional innovation system:

- The most common reflects the concern about the universally inadequate levels of research and innovation activity among smaller firms. This is judged to be an issue that cannot be improved materially either by information campaigns (hard to reach audience, which is reluctant to risk trying to innovate) or by fiscal measures (R&D tax credits work well for businesses already committed to carrying out quite a substantial programme of internal development work, but don't really help the undecided or occasional investor)
- The second group of measures is more thematic and intended to produce a substantial increase in the aggregate volume of BERD in subject areas judged to be of strategic importance to a region or country. These targeted measures are preferred to general measures (tax credits) because they permit the state to influence the focus of investment and to engender better and more committed efforts through the competition for funds

The issue of **relevance** tends not to be scrutinised especially closely in either group. The failure of businesses to invest sufficiently in research and innovation is seen as a truism, a point that is so obvious as to be hardly worth mentioning, except perhaps as a reminder to finance departments.

In many cases, the introductory sections to the evaluations cite statistics from the Community Innovation Survey (CIS), which tend to confirm that the majority of small businesses are non-innovators and that only a minority invest in or otherwise formally pursue research and innovation. These are helpful reference data, but arguably a little too aggregate, and possibly skewed in their sectoral distribution, to properly direct an evaluation of an individual programme or measure. There may be value for regional authorities in the more active promotion of the biennial Community Innovation Survey among local businesses, in order to increase responses to a level sufficient to test for regional-specific issues. Equally, it may be possible to exploit existing quarterly or annual business surveys by adding in innovation-related questions.

On **effectiveness**, the evaluation reports are mostly concerned to determine the extent to which public assistance has:

- Produced behavioural changes, which might be expected to persist in the longer term (e.g. a change in confidence and outlook amongst assisted businesses, along with a commitment to continue to invest in innovation in the future)
- Produced measurable economic benefits, in the short term

The first of these two questions tends to be addressed rather simply and directly, through opinion surveys, asking the principal contact at the assisted business to indicate the extent to which the support might has changed perceptions and or behaviour. We found no examples of evaluations having gone back to a panel of businesses over a number of successive years to test the persistence of the behavioural learning, although that would be feasible in some regions and possibly instructive.

On the second issue, the better evaluations are making use of micro-economic impact assessment techniques to estimate the total economic benefits realised by scheme participants – collectively - as a result of this public support. In almost all cases, the quantification of directly attributable economic benefits (increased income and employment) is determined through a questionnaire survey. This will tend to be either a census of the whole population or a large sample, depending upon the size of the scheme and the number of beneficiaries. As a rule of thumb, one would be looking to obtain replies from a minimum of 100 businesses in order to have confidence in the results. In most cases, those surveys will include questions that attempt to wrestle with the added value of the support: how much additional income did you secure this year, which you would attribute to your participation in the scheme; and how likely is it that the development project might have been able to proceed in the absence of support through this scheme.

The more robust approaches combine questionnaire surveys with control groups – a matched population of several hundred businesses that have not received support from this or similar government schemes – to test for the counterfactual situation, exploring underlying trends in innovation investment levels and commercial outputs among these unassisted businesses. Controlling for net effect is especially important in this arena, where public funds could so easily flow to private actors with a strong R&I track record and simply substitute for previous private investment. It is the area where there has arguably been the greatest methodological advances in the last 10-20 years, and provides powerful feedback for policy design (to minimise deadweight).

The state of the art is rather less certain when it comes to extending robustly from the direct benefits of a specific measure through to the wider economic impacts (e.g. through knowledge spillovers), although there are several meso-level econometric techniques coming into general use (e.g. input-output analyses to estimate and apply economic multipliers to the direct benefits). One could argue that estimating wider impacts is the major focus for the development of evaluation methodologies at present, reflecting the central importance of (hard to detect / measure) knowledge spillovers within the rationale for public investment.

As with the general schemes to encourage smaller businesses to invest in R&D, the evaluation of targeted measures tend to focus on the economic impacts realised as a result of the support that has been given. Again, the questionnaire survey and control group (often surveys of unsuccessful applicants) is at the heart of the study design in many cases and often supplemented by impact case studies to illustrate effects and more readily resolve the financial complexities of the attributable benefits, which are often reliant upon substantial further investment. This latter approach – sitting with the business in question and going over its profit and loss account – can produce more robust estimates of attributable costs and benefits (as compared with a project manager's unilateral estimate), however it is a costly process, two or three orders of magnitude more expensive (per data point) than a questionnaire survey.

Evaluations of targeted industrial R&D schemes ought to go far beyond the programme participants to gather relevant data, and official statistics are of very limited value as the targeted nature of the measures means one is seeking to expand investment in niche fields that may not be captured by more aggregate national surveys. So, ideally one would implement surveys of assisted businesses and a statistically significant sample of other regional businesses to determine the baseline and annual trends in BERD investment levels in a given field, adjusted to take account of any more general inflationary or deflationary trends in annual BERD, as revealed within the sample businesses and through the official national annual business inquiry (R&D). Ideally, and in the fullness of time, one would also look to judge the fecundity of that enlarged pool of strategic applied research: to what extent has the region registered a significant improvement – in national or international terms – in several key statistics, disaggregated to the theme level, such as patent registrations or the technology balance of payments?

The better evaluations tend to address the question of **efficiency** in a pretty generic manner, addressing service quality on the one hand (management efficiency) and value for money on the other:

- Inviting participants to rate the fairness, transparency and timeliness of the delivery processes (and possibly inviting people to offer suggestions for practicable and affordable improvements), wherein one can derive both performance ratios (e.g. % of applicants that rate the appraisal as transparent and fair; average time to process claims for payment, etc)
- Calculating the internal rate of return for the scheme, using the net present value of the cashflows realised by all beneficiaries together (positive and negative) in the period to date and projected forwards by 3-5 years

The best evaluations will take these efficiency metrics and compare the statistics with analogous schemes in operation elsewhere in the region or nationally, and possibly with other policy options altogether.

6.9 Networks and clusters

This suite of reports is primarily concerned with the evaluation of cluster schemes rather than the narrower innovation networks, which are often a subsidiary component of cluster policies. Cluster schemes typically address multiple actors and involve multiple types of support measures, designed to strengthen all aspects of the innovation ecosystem.

In a majority of cases, the published evaluations report mid-term reviews, three to five years into the life of the policy, and with a particular focus on operational efficiency and interim results (outputs). The study designs are quite conventional however. While almost all of the studies have analysed the nature and extent of the investments made and sought feedback on process and early benefits from all of the people and organisations assisted, the real heart of the evaluations tends to be the formative discussions with key stakeholders.

The one specific quality that defines these evaluations and sets the measure apart from the rest of the studies is the explicitly participatory nature of the approach, wherein stakeholders are involved in framing study questions and receiving the study findings as well as providing evidence for the evaluation to deliberate. This model is commonplace in several policy domains – such as international development or healthcare service delivery – however it is unusual within the realms of regional innovation or economic development, and reflects the reality of cluster schemes where government is just one stakeholder among several critical constituencies, all of which must be fully committed and able to move forward in concert. This kind of mutual dependency demands a more interactive, open and formative approach, which could serve as good practice for evaluation in other areas, both in terms of its ability to get at real-world issues – good and bad – and its ability to define next steps in a manner that is fully owned by all parties.

Relevance and efficiency questions are dealt with almost entirely through the combination of desk research, questionnaire surveys and individual interviews.

Effectiveness is dealt with rather more narrowly, polling the opinions of beneficiaries on the one hand and exploring the views of stakeholders on the other. Almost none of the reports under review had managed to quantify the spectrum of wider benefits and impacts – past or future – nor had any been able to rigorously dimension the added value of schemes in question. In part, this is a reflection of the nature of the undertaking: cluster programmes tend to be quite complex in their construction and are expected to improve the dynamism and innovativeness of a locale gradually, over a period of many years and even decades. Running an evaluation three or four years into what is typically a 10- to 20-year journey, is always going to be methodologically challenging. There is a scale issue too, wherein cluster programmes typically have a size and complexity beyond the reach of a normal, one-off evaluation, and yet these schemes are rarely so large that one can expect to detect improvements in changing macro regional statistics.

While the authors of several of the evaluations contend that it is almost impossible to quantify the impact of these policies, others suggest that a focus on intermediate outcomes and a commitment to continuous monitoring is a practicable and affordable development that would facilitate programme management and reassure funders. The solution is a commitment to greatly improved monitoring and the early definition of selected, relevant performance indicators that can be tracked continuously – from cradle to grave – and probed and elaborated periodically by external evaluations.

The only novel evaluation technique in evidence among the 12 cluster evaluations is the use of Social Network Analysis (SNA) within the evaluation of the Austrian genomics cluster programme. This exercise produced valuable objective statistics sufficient to reveal, on the one hand, the number and density of connections among regional actors, and on the other, the degree of centrality or peripherality of particular organisations and indeed projects.

6.10 Possible case studies

The following is a list of evaluation reports that the study team have judged to be of good quality and possible candidates for case study. At present, these examples have been compiled around different innovation measures, however it is conceivable that one might drop down a level and select the report or reports that have the best treatment of a questionnaire survey or impact case studies of the estimation of an internal rate of return. This list is intended to complement the recommendations from the survey of managing authorities and thereby develop a longer list of candidates from which to select the final sample of case studies.

Figure 26 List of potential candidates for case studies issued from the evaluation reports reviewed

Innovation Measure	Description
Direct financial support for innovation activities	Norwegian DEMO 2000 programme (NO) IWT grants for R&D projects of companies in Flanders (BE) Evaluation of Grant for Research and Development & SMART (UK)
Innovation management support	The Economic Impact Study of Business Link Local Service (UK)
Intermediary bodies and agencies	Incubator Support Programme Evaluation (NZ) The Mid-term evaluation of the Swedish national Incubator Programme (SE)
Start-ups and Spin-Offs	Evaluation of ERDF supported Venture capital and Loan funds (UK) Evaluation of Community Development Finance Institutions (CDFIs) (UK) Scottish Enterprise Proof of Concept Programme Evaluation – Rounds I to VI: Final Report (UK) From funding gaps to thin markets: UK Government support for early-stage

	venture capital (UK)
Networks & Clusters	Evaluation of the Centre of Expertise Programme in Finland (FI) Mid-term evaluation of the 'French Competitiveness Clusters' programme (FR) Austrian Genome Research Programme (GEN-AU): Mid Term Programme Management Evaluation (AT)
Strategic research	Austrian Genome Research Programme (GEN-AU): Mid Term Programme Management Evaluation (AT) Value for Money Review of Science Foundation Ireland (IE) Mid-term evaluation of the Quantifying and Understanding the Earth System Programme (QUEST) (UK)
Science-industry cooperation	Mid-term evaluation of the Swedish Institute of Excellence Centres Programme (SE) Impact evaluation of the Finnish Programmes for Centres of Excellence in research 2000-2005 and 2002-2007 (FI)

Appendix A Bibliography

Aragón, C., Araguren, M. J., & Iturrioz, C. (2010). *Evaluación de políticas clúster El caso del País Vasco*. Instituto Vasco de Competitividad, Fundacion Deusto.

A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009. American Journal of Evaluation, June 1, 2011, 32: 199-226.

Arnold, E., Bussillet, S., Sowden, P., Stroyan, J., Whitehouse, J., & Zaman, R. (2004). *An Evaluation of the RTDI for Collaboration Programme: main report*. Technopolis Group.

Arnold, E., Deuten, J., & Van Giessel, J. (2004). *An international review of Competence Centre Programmes*. Technopolis Group.

Arnold, E., Männik, K., Rannala, R., & Reid, A. (2008). *Mid-Term Evaluation of the Competence Centre Programme*.

Arnold, E., Malkin, D., Good, B., Clark, J., & Ruiz Yanniz, M. (2009). *Evaluating the National Innovation Strategy for Competitiveness*.

Australian Goverment . (2006). *Review of the National Research Flagships - an initiative of CSIRO*.

Australian Goverment. (2006). *Review of the National Research Flagship - an initiative of CSIRO*.

Bach, L., Cohendet P., Lambert, G. and Ledoux, M.J. (1992). *Measuring and managing Spinoffs: the case of the spinoffs generated by ESA programs*, in: 'Space Economics', Volume 144, Progress in Astronautics and Aeronautics.

Bailetti, A.J, Callahan, J.R. (1992). *Assessing the impact of university interactions on an R&D organization*, in: R&D Management, Volume 22, Issue 2, pages 145–156, April 1992.

BCG, CM International. (2008). *EVALUATION DES POLES DE COMPETITIVITE: Synthèse du rapport d'évaluation*.

Becattini, G. (1990). *The marshallian industrial district as socio-economic notion in Industrial Districts and Inter-firm Cooperation in Italy*.

Belgium Science Policy Office/BELSPO. (2010). *Belgium Report on Science, Technology and Innovation*.

Boekholt, P., Lankhuizen, M., Arnold, E., Clark, J., Kuusisto, J., Laat, B., et al. (2001). *An international review of methods to measure relative effectiveness of technology policy instruments*.

Bonaccorsi, A., Piccaluga, A. (1994). *A theoretical framework for the evaluation of university-industry relationships*, in: R&D Management, Volume 24, Issue 3, pages 229–247, July 1994.

Cooke, P., Eickelpasch, A., Williams, I. F., & Rangers, J. (2007). *Evaluation Report by the VINNVAXT International Review Team*. VINNOVA Report.

Cozzens, S. Popper, S., Bonomo, J., Koizumi, K. Flanagan, A. (1994). *Methods for evaluating fundamental science*, Washington D.C.

Etzkowitz, H., & Leydesdorff, L. (1997). *Universities and hte global knowledge economy: a triple helix of university-industry goverment relations*. Pinter, London.

European Commission. (2006). *Community Framework for State Aid for Research and Development and Innovation (2006/C 323/01).*

Evalsed guide, sourcebook and glossary, online: http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/index_en.htm.

Freeman, C. *Technological Infrastructure and International Competitiveness paper* . OECD Group on Science, Technology and Competitiveness, 1982.

Garofoli, G. (1994). *The industrial distric of Lecco: innovatoin and transformation processes in Entrepreneurship & Regional Development*.

Geizler, E., Rubenstein, A.H. (1989). *University-Industry Relations: A Review of Major Issues*, Chapter 3 in *Cooperative research and development: the industry, university, government relationship*, edited by Albert N. Link, Gregory Tassey, Springer, 1989.

Georghiou, L. and Rossener, D. (2000). *Evaluating technology programs: tools and methods*, Research Policy, 29, 2000, pp. 657-678.

Jordan, G.B. (2010). *A theory-based logic model for innovation policy and evaluation*, in: Research Evaluation, Volume 19, Number 4, October 2010, pp. 263-273(11).

Griliches. Z. (1990). *The search for R&D Spillovers*, Working paper, Harvard University, Cambridge, MA.

Griliches. Z. (1979). *Issues in Assessing the contribution of R&D to productivity Growth*, Bell Journal of Economics, Vol. 10, No. 1, 1979, pp. 92-116.

Hansen, M.B., Vedung, E. (2010). *Theory-Based Stakeholder Evaluation*, in: American Journal of Evaluation 31(3) 295-313

Huang, Q., McDonald, F., Tsagdis, D., & Tuselmann, H. (2007). *Is there evidence to support porter-type cluster policies?* Regional studies vol 41 n.1 .

IDEA Consult and ADE. (2011). *Ex-post evaluation of the IAP programme (phase VI, 2007-2011): main evaluation report*. Belgium Science Policy Office (Belspo), Brussels.

Industry Canada. (2003). *Summative Evaluation of CANARIE Phase 3 Final Report*. Audit and evaluation branch.

IWEPS; Institut Wallon de l'evaluation, de la prospective et de la statistique. (2009). *Evaluation du Plan d'actions prioritaires pour l'Avenir wallon: rapport de synthese*.

Kagi, W., Frey, M., Steiner, S., koci, M., & Hohmann, R. (2007). *Evaluation des Vereins „energie-cluster.ch" Schlussbericht*. Eidgenössisches Departement fu r Umwelt, Verkehr, Energie und Kommunikation UVEK.

Klette, ,. T., Møen, J., & Griliches, Z. (2000). *Do Subsidies to Commercial R&D Reduce Market Failures?*

Kuhn, J. M. (2010). *An Analysis of Firm Growth Effects of the Danish Innovation Consortium Scheme*. CEBR- Centre for Economic and Business Research, Danish Agency for Science, Technology and Innovation.

Margherita, B., Andrea, L. (2006). University–industry interactions in applied research: The case of microelectronics, Research Policy, 35(10), 1616-1630.

McLaughlin, J.A., Jordan, G.B. (1999). *Logic Models: A Tool for Telling Your Program's Performance Story*, in: Evaluation and Program Planning, Volume 22, Number 1, February 1999.

MEANS. (2000). *Guide to Methods for Evaluating Structural Policies*.

Meier zu Köcker, G., & Rosted, J. (2009). *Promoting Cluster Excellence: measuring and benchmarking the quality of cluster organisations and performance of clusters*. PRO INNO EUROPE.

Morgan, K. (2007). *The learning region: institutions, innovation and regional renewal*. Regional studies. 41.

Nightingale, P., & Reid, A. (2010). *Early-stage Investment by the Estonian Development Fund: An appraisal of activities 2007-2009 and scenarios for futures development*.

Nightingale, P., Murray, G., Cowling, M., Baden-Fuller, C., Mason, C., Hopkins, M., et al. (2009). *From funding gaps to thin markets: UK Goverment support for early-stage venture capital*. BVCA, NESTA.

OECD. (2006). *Goverment R&D Funding and Company Behaviour, measuring behavioural additionality*.

Pelkonen, A., Konttinen, J., Oksanen, J., Valovirta, V., Boekholt, P., & Levasluoto, J. (2010). *Osaamisklusterit alueiden voimien yhdistäjänä Osaamiskeskusohjelman (2007–2013) väliarviointi*. Ministry of Employment and the Economy.

Perrin, B. (2002). *How to - and How Not to- Evaluate Innovation*.

Porter, M. (1998). *Clusters and competition:new agendas for companies, goverments, and institutions*. Division of Research, Harvard Business School Boston.

Rosiello, A. (2004). *Evaluating Scottish Entreprise's Cluster Policy in Life Sciences: a descriptive analysis*. INNOGEN Working Paper No. 16.

Salles-Filho, S., Bonacelli, M.B., Carneiro, A.M., De Castro, P.F.D, Santos, F.O. (2011). *Evaluation of ST&I programs: a methodological approach to the Brazilian Small Business Program and some comparisons with the SBIR program*, in: Research Evaluation, Volume 20, Number 2, June 2011 , pp. 157-169(13).

Sölvell, Ö. (2008). *Clusters: Balancing Evolutionary and Constructive Forces*.

Smith, K. (2000). *Innovation as a systemic phenomenon: rethinking the role of policy*. Entreprise and Innovation Management Studies, 1 (1) 73-102.

Wise, E., Lotte, L., & Degn Bertelsen, M. (2009). *The use of data and analysis as a tool for cluster policy: an overveiw of international good practices and prespectives prepared for the European Commission*. European Cluster Alliance.

Appendix B Definition of evaluation criteria, methods and approaches used in the Literature Review

B.1 Common evaluation criteria

The following table provides an overview and explanation of the main data collection criteria commonly used in evaluations.

Method	Description
Effectiveness	Has the intervention met its initial objectives?
Efficiency	Has the intervention delivered its expected outcomes at the minimum cost or, equivalently, with maximizing outcomes for a given level of resources?
Relevance	What is the appropriateness of the explicit objectives of the intervention, with regard to the socio-economic problems the intervention is meant to solve?
Sustainability (further developments)	Has the intervention produced longer-term and wider impacts that persist over time and after the end of the programme?

Technopolis, based on the Evalsed guide, sourcebook and glossary (http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/index_en.htm)

B.2 Common data collection methods for evaluations

The following table provides an overview and explanation of the main data collection methods commonly used in evaluations.

Method	Description
Use of administrative data	Use of data and other information relating to the programme's administration, activities or performance systematically collected during the lifetime of the of the programme, usually by the programme management or administration, although the availability and quality of the administrative data can be variable depending on the programme requirements and the programme implementation mechanisms
Use of secondary data	Use of existing data and documents directly or indirectly related to a programme, which are not produced during the evaluation process. This includes: Desk research of programme documents and other related documents (administrative manuals, application forms, assessment forms, existing evaluation reports and broader policy reports, etc) Literature review (academic publications, grey literature, etc) Collection of statistical data from existing surveys or databases
Individual stakeholder interview	Technique used to collect qualitative data and the opinions of people who are concerned or potentially concerned by the intervention, its context, its implementation and its effects. Several types of individual interview exist, including informal conversations, semi-structured interviews and structured interviews. The latter is the most rigid approach and resembles a questionnaire survey. A semi-structured interview consists of eliciting a person's reactions to predetermined elements, without hindering his or her freedom to interpret and reformulate these elements. Individual interviews target two types of population: Stakeholders and beneficiaries interviews conducted with those who have participated in the programme or policy evaluated 'Non-participant' interviews conducted with those who have not participated in a measure or who have not benefited from the activities or services provided by a measure.
Questionnaire survey	A survey consists in putting a series of standard questions in a structured format to a sample of individuals who are usually selected as being representative of the population under observation. As individual interviews, surveys target either the beneficiaries and stakeholders, or the non-beneficiaries. Surveys are either exhaustive, covering the whole population involved or based on a representative population of the whole population observed. They can be carried out by phone, on paper or online.
Focus groups (also referred to as 'workshops', 'seminars', or 'group meetings')	The focus group is a well-established method of social inquiry, taking the form of structured discussion that involves the progressive sharing and refinement of participants' views and ideas. The discussion is used to identify important themes or to construct descriptive summaries of views and experiences on the focal topic.

	The typical format involves a relatively homogenous group of around six to eight people who meet once, for a period of around an hour and a half to two hours. The evaluator or researcher is in charge of facilitating the group interaction.
Expert panel (also referred to as peer reviews)	An "expert panel" is a specially constituted work group that meets for evaluation. Expert panels are usually made up of independent –often international- specialists recognised in the fields covered by the evaluated programme. In the evaluation process, they are usually used as a mechanism for synthesising information from a range of sources, drawing on a range of viewpoints, in order to arrive at overall conclusions. Results are usually based on reaching a consensus of opinion in arriving at a value judgement on the programme and its effects.
Case studies	Methods of inquiry that focus on detailed data collection and analysis and which focus on a restricted number of participants/beneficiaries. It involves in-depth study of a phenomenon in a natural setting, drawing on a multitude of perspectives. These multiple perspectives may come from multiple data collection methods (both qualitative and quantitative), or derive from multiple accounts of different actors in the setting. The phenomena may concern individuals, programmes, organisations, projects, groups of people or decision-making processes.
Descriptive statistics analysis	Use of basic descriptive statistics to analyse the data and describe an intervention or a situation (e.g. uptake analysis, meaning the extent to which target beneficiaries have taken up the support provided by the intervention; or counterfactual analysis comparing subjects who were exposed to an intervention with a comparison group who were not exposed)
Micro-economic models	Micro-economic modelling refers to modelling behaviour/performance of individual economic actors, most often businesses but also households, consumers, etc. In the context of evaluation, micro-economic modelling would be used to try to understand the effects (or lack thereof) of public interventions on the behaviour of a business (or other economic actors). The usefulness of the model depends on whether it can be generalised.
Macro-economic models	A macroeconomic model is a tool used to present a holistic view of the operation of an economy, usually in the form of a computer-based system. It is a means of collating research on the economy in a systematic and policy-relevant way, and depends on the availability of such research. The goal of a macroeconomic model is to replicate the main mechanisms of an entire economic system, which may consist of a region (such as the Italian Mezzogiorno), a nation state (such as Poland), or a collection of nation states (such as the 27 members of the EU). The only requirement is that the entity being modelled is large enough to display the distinctive properties that are the subject area of macroeconomics.
Input/output analysis	Method used to characterise economic activity in a given time period, and to predict the reaction of a regional economy to stimulation, for example, from increased consumption or changes in government policy.
Cost benefit analysis	Tool for judging the advantages of the intervention from the point of view of all the groups concerned, and on the basis of a monetary value attributed to all the positive and negative consequences of the intervention (which must be estimated separately). Cost-benefit analysis is used mainly for the ex ante evaluation of large projects.
Multicriteria analysis	Multicriteria analysis is used to make a comparative assessment of alternative projects or heterogeneous measures. With this technique, several criteria can be taken into account simultaneously in a complex situation. The method is designed to help decision-makers to integrate the different options, reflecting the opinions of the actors concerned, into a prospective or retrospective framework. Participation of the decision-makers in the process is a central part of the approach. The results are usually directed at providing operational advice or recommendations for future activities.
Bibliometric or patent database studies	Searches of scientific publications (and sometimes their citations) and patents from bibliometric and patent databases.
Social Network Analysis	Analysis that aims to map the social interaction between the subjects of an evaluation including the beneficiaries. It considers the participants as a social structure made up of individuals (or organizations) called "nodes", which are tied (connected) by one or more specific types of interdependency. The SNA aims at assessing the intensity of the interdependency between the individuals.
Indicators	A characteristic or attribute, which can be measured to assess an intervention in terms of its outputs or results. Output indicators are normally straightforward. Result indicators may be more difficult to derive, and it is often appropriate to rely on indirect indicators as proxies. Indicators can be either quantitative or qualitative. Context indicators relate to the environment for the programme.

Technopolis, based on the Evalsed guide, sourcebook and glossary (http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/index_en.htm)

B.3 Common analytical approaches and methods

There are two main analytical approaches used during evaluation to assess the impacts of the policy or measures under review:

- Counterfactual Impact Evaluation (CIE), which tends to focus on a single type of impact and attempts to quantify the differences in the improvements observed in one group (beneficiaries) with the trends in the wider population;
- Theory-Based Impact Evaluation (TBE), which tends to be more qualitative and acknowledges the complexity and time lags associated with the kind of innovation systems policy makers are seeking to influence. TBE places rather more emphasis on the identification of intended outcomes, and then exploring the extent to which those various anticipated intermediate effects have been realised and give confidence that a measure is working well and is more or less likely to deliver on its ultimate impacts. It is possible to combine TBE with counterfactual analysis.

These approaches, as well as the main analytical methods used under both approaches, are presented in the two tables below.

1. Counterfactual impact evaluation (CIE)

- **Approach that compares the state where no intervention has (or is assumed to have) taken place and the state where there has been an intervention. The question of attribution (i.e. how and to what extent is what occurred attributable to the programme?) is central to this approach. Since by definition we can never observe the counterfactual situation, we can never observe effects with certainty.**
- **Counterfactual Impact evaluation typically produces numbers and gives a causal interpretation based on empirical evidence and some assumptions.**

The main statistical analytical methods in use in the evaluation reviewed are listed below.

Difference-in-differences	Difference-in-differences or double differencing is based on the availability of outcome data (for example, firm sales) for beneficiaries and non-beneficiaries (assisted and non assisted firms), both before and after the intervention (say, the year preceding and the year following the receipt of assistance). It is based on a baseline approach, assessing the difference between the situation before and after the intervention.
Propensity score matching	The matching strategy is based on the possibility of observing all the relevant characteristics X of both beneficiaries and non-beneficiaries and to pick the non-beneficiaries that "look like" beneficiaries according to these characteristics.
Discontinuity design	The strategy is based on the idea of discontinuity in treatment around a threshold, which applies mainly to those situations in which some units are made eligible for the intervention and others are made ineligible by some well defined rule, typically some administrative rule...
Use of instrumental variables	The instrumental variables identification strategy is based on the idea of involuntary variation (in the official jargon instrumental variables): those situations in which the receipt of treatment is partially determined by an extraneous factor.

Technopolis, based on the Evalsed guide, sourcebook and glossary (http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/index_en.htm)

2. Theory-based impact evaluation

- **Contrary to the counterfactual approach, the theory-based impact evaluation focuses on the notion of causality or contribution (i.e. demonstrate whether or not the evaluated intervention is one of the causes of observed change by asking the question 'How do the intervention contribute to the observed changes').**
- **Theory-based evaluation discuss and apply the program theory idea (i.e maps out the causal chains from inputs to outcomes, often using a schematic representation of the intervention logic of the programme)**
- **It produces a narrative and insights into why things work or not, rather than numbers.**
- **It requires rigour in the analysis of causal chains and can involve the systematic identification and investigation of alternative explanations for observed impacts.**

Two analytical methods in use in the evaluation reviewed are listed below (other methods are in use and have been put into theories but are not developed here since we found very few examples of their use in the evaluation reviewed here).

Contribution analysis	A contribution analysis, which relies upon a unique chain of logical arguments that are verified through careful fieldwork. The contribution analysis is most often based on the intervention logic and tries to determine if the chains between the different inputs, outputs, outcomes and impacts are working or not, and why.
Theory-based stakeholder evaluation	In contexts characterised by dissension and conflict, the constructio n of one unitary intervention theory as in the contribution analysis is often not realistic. Theory-based stakeholders evaluation therefore aim at comparing, clarifying, and separating the different viewpoints of primary stakeholder groups based on the contribution analysis approach. [51]

Technopolis, based on the Evalsed guide, sourcebook and glossary (http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/index_en.htm)

[51] Hansen, M.B., Vedung, E. (2010). Theory-Based Stakeholder Evaluation, in: American Journal of Evaluation 31(3) 295-313

Appendix C List of evaluation reports used by sections

C.1 Innovation management support and dissemination

Title of the evaluation report (author)	Year	Commissioner	Type of evaluation	Country
Strategic and operational review of the Scottish Enterprise Facilitator pilot (Ekos Ltd)	2009	Scottish Enterprise Highlands & Islands Enterprise The Scottish Government	Ex post	United Kingdom
Innovation Agents - New approaches to innovation in SMEs: Lessons from the interim evaluation of the pilot project Regional Innovation Agents (DAMVAD)	2009	Danish Agency for Science, Technology and Innovation	Mid-term	Denmark
Economic Impact Study of Business Link Local Service: Final Report (University of Warwick, Aston Business School and Kingston University)	2006	UK Department for Business, Enterprise and Regulatory reform	Mid-term	United Kingdom
The Foundation for Finnish Inventions and the Performance and Effectiveness of its Network of Invention Advisers (Pirjo Kutinlahti – Mika Nieminen – Kirsi Hyytinen – Jari Konttinen – Juha Oksanen – Niina Elo)	2006	Finnish Trade and Industry Department	Mid-term	Finland
Enterprise, Insight, Impact, Evaluation: Review of the Make Your Mark Challenge, Make Your Mark Clubs and Ambassadors Programme (Training and Employment Research Unit (TERU) University of Glasgow)	2011	Enterprise Insight	Ex-post	United Kingdom
Evaluation of the Campaign uni:invent (JOANNEUM RESEARCH Forschungsgesellschaft mbH – Institut für Technologie- und Regionalpolitik (InTeReg)	2008	Ministry BMWA (Economy and Employment) and BMWF (Science and Research)	Mid-term	Austria

C.2 Intermediary bodies and agencies

Title of the evaluation report (author)	Year	Commissioner	Type of evaluation	Country
Incubator Support Programme (Ministry of Economic Development Industry and Regional Development Research, Evaluation and Monitoring)	2008	Ministry of Regional Development	Mid-term	New Zealand
BITS Incubator programme and the Intelligent Island Incubator (Allen Consulting Group)	2003	The Department of Communications, Information Technology and the Arts	Ongoing	Australia
Swedish National Incubator Programme (Inno Germany AG)	2008	VINNOVA	Mid-term	Sweden
West of Scotland Science Park (EKOS)	2009	Scottish Enterprise	Ongoing	UK
Kent Science Parks (Centre for Strategy and Evaluation LLP)	2008	Swale Forward in partnership with Swale Borough Council KCC, SEEDA and the Thames Gateway Kent Partnership	Ongoing	UK
Science and Technology parks (POLICY & ACTION GROUP UNICONSULT Sp. z o.o.)	2008	Ministry of Regional Development	Ongoing	Poland
Canadian Initiative for International Technology Transfer	2006	Natural Resources Canada	Mid-term	Canada
Regional Office Technology Transfer Programme	1993	DTI (now BIS)	Ongoing	UK
Knowledge transfer programmes funded through the science budget (SQW)	2005	DTI (now BIS)	Mid-term	UK
Support for Technology Transfer (Ramboll Management)	2008	The Saxony State Ministry for Economy and Labour	Mid-term	Germany
TechnoKontakte (Seibersdorf Research)	2003	BMWA	Mid-term	Austria

C.3 Start-ups and spin-offs

Title of the evaluation report (author)	Year	Commissioner	Type of evaluation	Country
Scottish Enterprise Proof of Concept Programme Evaluation – Rounds I to VI: Final Report (Price Waterhouse Cooper)	2006	Scottish Enterprise	Mid-term	Scotland
Effects of EXIST from the perspective of grantees (Fraunhofer Institute for Systems and Innovation Research)	2008	Federal Ministry of Economics and Technology (BMWi)	Mid-term	Germany
Evaluation of ERDF Supported Venture Capital and Loan Funds (Centre for strategy and Evaluation Services, Scottish Government Social Research)	2008	Scottish Government Social Research	Mid-term	Scotland
Early assessment of the ~ Impact of BIS Equity Fund Initiatives (Centre for Enterprise and Economic Development Research, Middlesex University Business School)	2010	Department for Business, Innovation and Skills (BIS)	Early Stage	UK
Evaluation of Community Development Finance Institutions (CDFIs) (GHK)	2010	Department for Business, Innovation and Skills (BIS)	Mid-term	England
From funding gaps to thin markets: UK Government support for early-stage venture capital (Paul Nightingale, Gordon Murray, Marc Cowling, Charles Baden-Fuller, Colin Mason, Josh Siepel, Mike Hopkins and Charles Dannreuther)	2009	NESTA	Mid-term	UK
External Evaluation of the Pilot Scheme CREA Concerning Support for Venture Capital Companies Financing SMEs in the Seed and Start-up Phase: Final Report (The Evaluation Partnership Limited)	2006	The European Commission Enterprise and Industry Directorate-Genera	Ex-ante	Europe
Finnish Industry Investment Ltd: An International Evaluation (Markku Maula and Gordon Murray)	2003	Ministry of Trade and Industry	Mid-term	Finland

C.4 Science-industry cooperation

Title of the evaluation report (author)	Year	Commissioner	Type of evaluation	Country
First evaluation of the Berzelii centra programme and its centres (Vinnova)	2009	Vinnova	Mid-term	Sweden
Mid Term evaluation of the Institute Excellence Centres Programme (Vinnova)	2009	Vinnova	Mid-term	Sweden
Mid-Term Evaluation of the Competence Centre Programme (Technopolis)	2008	Ministry of Economic Affairs and Communications	Mid-term	Estonia
Impact Evaluation of the Finnish Programmes for Centres of Excellence in Research 2000–2005 and 2002–2007 (Mari Hjelt, Paavo-Petri Ahonen and Piia Pessala)	2009	Academy of Finland	Ex-post	Finland
National and Regional Economic Impacts of Engineering Research Centres: A Pilot Study (SRI International)	2008	National Science Foundation	Ex-post	United States
Assessment „Zukunft der Kompetenz-zentrenprogramme (K plus und K ind/net) und Zukunft der Kompetenzzentren“ (Fraunhofer-Institut fur Systemtechnik und Innovationsforschung)	2004	Ministries BMVIT (Transport, Innovation and Technologie) and BMWA (Economy and Employment)	Mid-term	Austria
Mid-term assessment of the Centre of expertise Programme (2007-2013) (VTT, Technopolis)	2010	Ministry of Employment	Mid-term	Finland
An Analysis of Firm Growth Effects of the Danish Innovation Consortium Scheme (CEBR - Centre for Economic and Business Reasearch)	2010	Danish Agency for Science, Technology and Innovation	Mid-term	Denmark
Evaluation of the Christian Doppler Research Association (Technopolis, Joanneum Research, Frauenhofer ISI, KMU Forschung Austria)	2005	Ministry BMWA (Economy and Employment)	Mid-term	Austria
Knowledge Transfer Partnerships Strategic Review (Regeneris Consulting)	2010	Technology Strategy Board	Mid-term	England
The effectiveness of the innovation voucher (2004 and 2005): impact on innovative inputs and innovative output of companies (CPB Netherlands Bureau for *Economic Policy Analysis)*	2010	CPB Netherlands Bureau for Economic Policy Analysis	Ex-post	Nether-lands

C.5 Strategic research

Title of the evaluation report (author)	Year	Commissioner	Type of evaluation	Country
Evaluation of Genome Canada: Final report (KPMG)	2009	Genome Canada	Mid-term	Canada
Evaluation of the Functional Genomics Programme in Norway (FUGE) (Independent Panel)	2006	The Research Council of Norway	Mid-term	Norway
Austrian Genome Research Programme (GEN-AU): Mid Term Programme Management Evaluation (Joanneum Research, TIA Consulting, Inc., Austrian Institute for SME Research)	2005	Federal Ministry for Education, Science and Culture	Mid-term	Austria
Mid-term evaluation of the Quantifying and Understanding the Earth System Programme (QUEST) (Independent Panel)	2008	Natural Environment Research Council	Mid-term	UK
Evaluation of BBSRC 'genomics' research (Independent Panel)	2011	BBSRC	Mid-term	UK
Value for Money Review of Science Foundation Ireland (Indecon International Economic Consultants)	2008	Department of Enterprise Trade and Employment	Mid-term	Ireland
The value of strategic basic research - Effect measurement of the STWW, GBOU and SBO projects with an economic focus (Technopolis Group)	2008	IWT	Ex-post	UK

C.6 Direct financial support to innovation activities

Title of the evaluation report (author)	Year	Commissioner	Type of evaluation	Country
A look into the Black Box: What difference do IWT R&D grants make for their clients? (Idea Consult)	2006	IWT	ex-post	Belgium
Evaluation FFF – Impact Analysis Background Report (Joanneum Research**;** WIFO – Austrian Institute of Economic Research**;** KOF – Swiss Institute for Business Cycle Research)	2004	Ministry for Transport, Innovation and Technology (BMVIT)	mid-term	Austria
The effectiveness of the innovation voucher 2004 and 2005 - Impact on innovative inputs and innovative output of companies (CPB Netherlands Bureau for Economic Policy Analysis)	2007	Netherlands Bureau for Economic Policy Analysis (CPB)	Ex-post	Netherlands
Evaluation of the DEMO 2000 programme (NIFU STEP)	2005	Norwegian Ministry of Petroleum and Energy (MPE)	ex-post	Norway
Interim Evaluation of the AWS Technology Programme (Technopolis Group)	2006	Ministry for Economy and Work (BMWA)	mid-term	Austria
Evaluation of Grant for Research and Development & SMART (PACEC)	2009	DIUS / LDA	Ex-post	UK
The impact of the State's enterprise supports on the competitiveness of Estonian economy (National Audit Office of Estonia)	2010	Ministry of Economic Affairs and Communications and Ministry of Finance		Estonia

Appendix D Country abbreviations used in the Literature Review

Figure 27 List of country abbreviations (EU countries)

Short name in English	Country code
Belgium	BE
Bulgaria	BG
Czech Republic	CZ
Denmark	DK
Germany	DE
Estonia	EE
Ireland	IE
Greece	EL
Spain	ES
France	FR
Italy	IT
Cyprus	CY
Latvia	LV
Lithuania	LT
Luxembourg	LU
Hungary	HU
Malta	MT
Netherlands	NL
Austria	AT
Poland	PL
Portugal	PT
Romania	RO
Slovenia	SI
Slovakia	SK
Finaland	FI
Sweden	SE
United Kingdom	UK

EU, Europa Institutional Style Guide: http://publications.europa.eu/code/en/en-370100.htm

Figure 28 List of country abbreviations (Non-EU countries)

Short name in English	Country code
Australia	AU
Canada	CA
New Zealand	NZ
Norway	NO
Switzerland	SW
United States	US

technopolis |group| United Kingdom
3 Pavilion Buildings
Brighton BN1 1EE
United Kingdom
T +44 1273 204320
F +44 1273 747299
E info@technopolis-group.com
www.technopolis-group.com