

# Central Innovation Manual on Excellent Econometric Evaluation of the Impact of Public R&D Investments

“CIM 2.0”

By

Thomas Alslev Christensen, PhD and Head of Department for Innovation Policy  
and Research Analyses,  
Hanne Frosch, Special Advisor  
David Boysen-Jensen, Head of Section,

Danish Ministry of Science, Innovation and Higher Education  
Danish Agency of Science, Technology and Innovation,  
Bredgade 40,  
1260 Copenhagen K,  
Denmark  
[www.ufm.dk](http://www.ufm.dk)

## Indhold

1	Preface.....	3
1.1	Acknowledgement.....	4
2	PART I – What is CIM (2.0)? .....	4
3	Overview of important standards and minimum requirements – The 9 principles .....	8
3.1	Principal 1 – Data quality and harmonisation of data collection .....	9
3.2	Principal 2 – Selection of comparable companies and/or individuals to control groups .....	10
3.2.1	Control groups may be selected using a so-called ‘propensity score matching’ and ‘nearest neighbour’ method. ....	12
3.3	Principal 3 – Use of the difference-in-difference method and balanced panel data .....	13
3.3.1	The difference-in-difference method .....	13
3.3.2	Balanced panel data .....	14
3.3.3	Standards for calculating economic effects .....	15
3.4	Principal 4 – Treatment of outliers.....	15
3.5	Principal 5 – Long-time series.....	16
3.6	Principal 6 – Robustness test.....	16
3.7	Principal 7 – Impact indicators should be made relative .....	16
3.8	Principal 8 – Peer review of results.....	16
3.9	Principal 9 – Failures and stress tests .....	16
4	PART II - Standard for performance objectives: Key performance indicators	17
4.1	Ex ante evaluation .....	18
4.1.1	Input and output indicators .....	18
4.2	Baseline measurement at ex post evaluation .....	18
4.2.1	Outcome and impact indicators .....	20
5	PART III – Overview of the most important key performance indicators, impact assessments and results in Denmark.....	20
5.1	Results of impact assessments in Denmark .....	20

## 1 Preface

Excellent methods for the measurement of the impact of research and innovation policies is key for evidence-based policy-making. The field of quantitative impact assessments is in an on-going development. In 2011 the Ministry of Science, Innovation and Higher Education launched its first manual on excellent econometric evaluation<sup>1</sup>: “*Central Innovation Manual on Excellent Econometric Evaluation of the Impact of Interventions on R&D and Innovation in Business (CIM)*”. Due to new research experience, new impact assessments, improved data collection and longer time series it is time to update this manual in a version 2.0. CIM 2.0 establishes a set of agreed minimum analytical and data quality requirements and key performance indicator standards for econometric impact studies of public research and innovation programmes and policies. Furthermore, it provides an overview of the most important key performance objectives and indicators as well as of econometric impact assessments of the Danish research and innovation system.

### Why CIM?

*Policy makers and programme owners are constantly looking for the best and most solid methods to identify and verify the impact of policies. The Central Innovation Manual (CIM) is a guide on excellent methods for the conduct of economic as well as non-economic econometric impact assessments.*

#### ***Structure of the manual***

The report is in three parts. Part I is a manual on excellent econometric evaluation of the impact of interventions on R&D and innovation in businesses, part II sets up standards for key performance indicators and part III presents an overview of key performance indicators and econometric impact assessments of Danish research and innovation policy.

#### ***Part I: Manual on excellent econometric evaluation of the impact of interventions on R&D and innovation in businesses***

Nine principles which are important standards and minimum requirements for excellent econometric impact assessments are formulated on experiences with impact assessments of Danish R&D and innovation programmes and the latest research in the field.

#### ***Part II: Standards for key performance indicators***

Part II gives an overview of a comprehensive set of input, output and impact performance indicators to measure the effect of research and innovation programmes e.g. in the cases where the main purpose of the programme may be non-economic activities or limited access to quantitative micro data and long-time series. Although CIM 2.0 lists standards for impact assessments, the intention has been to do this in such a way that there is room for flexibility.

#### ***Part III: Overview of the most important key performance indicators, impact assessments and results in Denmark***

Firstly, this part provides an overview of the most important econometric impact assessments on the Danish research and innovation programmes and on the R&D systems and their impacts. Secondly, part III also contains an overview of the most common key performance indicators used in the studies of the Danish research and innovation programmes.

---

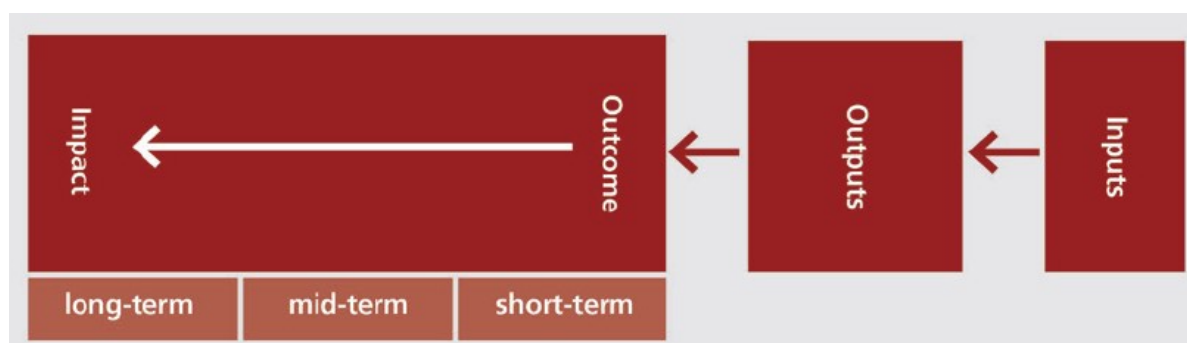
<sup>1</sup> <http://fivu.dk/en/publications/2012/central-innovation-manual-on-excellent-econometric-impact-analyses-of-innovation-policy-cim?searchterm=CIM>

## 1.1 Acknowledgement

This manual has been developed in close cooperation with university researchers and private researchers. Furthermore, the manual has been discussed at workshops and seminars between 2010 and 2013 at Copenhagen Business School (CBS) and in the Ministry of Science, Innovation and Higher Education and we have received valuable comments from several ministries, including the Ministry of Finance, the Danish Ministry of Business and Growth, the Danish Ministry of Climate and Energy, the Danish Ministry of Food, Agriculture and Fisheries, the Danish Ministry of Foreign Affairs and the Danish Ministry of the Environment. A special thank shall be given to Professor Anders Sørensen and Ph.D. Johan Kuhn at Copenhagen Business School (CBS) and to Michael Mark from DAMVAD Consulting A/S.

## 2 PART I – What is CIM (2.0)?

Internationally, policy makers and programme owners have been working on monitoring and evaluating research and innovation programmes for the past couple of decades. The rigorous measurement of the impact of innovation policies is a key for evidence-based policy-making.<sup>2</sup> Due to many economic challenges and international competition between countries and who wants to be among the leading countries in terms of R&D and innovation, policy makers are demanding impact of and accountability from research, development and innovation programmes and policies. Impact assessments can provide robust and credible evidence on performance and on whether a particular program achieved its desired outcomes and impacts.



Impact assessments are part of a broader agenda of *evidence-based policy making*. This growing global trend is marked by a shift in focus from inputs and outputs to outcomes and impacts. The increased focus has been encouraged, among others, by OECD,<sup>3</sup> who has focused great attention on the area through a coordinated effort among most of the 27 EU countries, Korea, Norway, Switzerland, Russia, Turkey, South Africa and most of the countries in South America.

### What is CIM 2.0?

*Central Innovation Manual on Excellent Econometric Evaluation of the Impact of Interventions on R&D and innovation in Business (CIM 2.0) focuses on econometric outcomes and impact assessment methods when analysing investments in public and private research, development, education and innovation.*

<sup>2</sup> Communication from the Commission to the European Parliament, the Council, The European Economic and Social Committee and the Committee of the Regions: “*Measuring innovation output in Europe: towards a new indicator*”, final 624, 13 September 2013

<sup>3</sup> OECD (2008), Science, Technology and Industry Outlook.

The vision of the Danish Agency of Science, Technology and Innovation<sup>4</sup> is that the econometric impact assessments of Danish public research and innovation programmes and policies will be among best practice in international contexts over the coming decade.<sup>5</sup> The objective of CIM 2.0 is to establish minimum requirements and standards on methods and data that are necessary for the implementation of excellent econometric impact assessments of the innovation and research policy.

CIM 2.0 is the result of the evaluation strategy of the Danish Agency of Science, Technology and Innovation, which started as a 5-year research and innovation project about performance measurements in the innovation field. The main elements of the 5-year project are the more than 20 evaluations (see chapter 4.1) which have been conducted from 2007 to 2011. The research project has been extended to 2014 and includes now more than 30 econometric impact studies and evaluations.

CIM 2.0 has been implemented in accordance with the most recent and best econometric research methods, which facilitates publication of methods and results in the most respected international journals in the relevant fields.

### **Text box 1. Definition of impact assessment**

Impact assessments are a particular type of evaluation that seeks to answer cause-and-effect questions. Unlike general evaluations, which can answer many types of questions, impact assessments are structured around one particular type of questions: *What is the impact (or causal effect) of a program or a policy on an outcome of interest?* This basic question incorporates an important causal dimension: We are interested only in the isolated impact of the program or policy, that is, the effect on outcomes that the program or policy directly causes. The basic form of impact assessment will test the effectiveness of a given policy or program. Qualitative data, monitoring data and evaluations are needed to track program and policy implementation and to examine questions of process that are critical to informing and interpreting the results from impact assessments. In this sense, impact assessments and other forms of evaluation are complements for one another rather than substitutes.

The primary target groups of CIM 2.0 are programme owners and policy makers in the Danish ministries, government agencies and regions working with programmes and policies on research and innovation. Other target groups are external expert stakeholders, evaluation experts (from organisations, regions, research, technology, education and knowledge institutions etc.) , researchers, trade unions and business associations who are interested in following and having a dialogue with the Ministry of Science, Innovation and Higher Education on how to conduct impact evaluation studies and how to document the effects of the innovation policy. In other words anyone interested in econometric impact evaluation in ministries and agencies.

CIM 2.0 is not identical to the work done in other countries<sup>6</sup> since the key objective is to establish a clear set of minimum requirements for how to conduct so-called excellent econometric impact assessments of innovation policy. CIM 2.0 establishes a framework for a “standard” impact assessment procedure which makes it possible to conduct excellent impact assessments of interventions on research and innovation programmes in business that makes it possible to compare the impact of different programmes. Hence, CIM 2.0 is not an attempt to establish practical guidance on a broader number of methods on how to evaluate the wider impact of research and

<sup>4</sup> In the Ministry of Science, Innovation and Higher Education: [http://fivu.dk/en?set\\_language=en&cl=en](http://fivu.dk/en?set_language=en&cl=en)

<sup>5</sup> In the reports ‘Clusters Are Individuals – Benchmarking Insights from Cluster Management Organizations and Cluster Programs’ by Kompetenznetze Deutschland (VDI/VDE Innovation + Technik), 2011, and ‘Service innovation: Impact analysis and assessment indicators’ by the European Commission's Pro-Inno Net EPISIS, 2011, the Danish Ministry of Science's econometric performance measurements are singled out as being among international best practice.

<sup>6</sup> See *Guidance on evaluating the impact of interventions on business*, Department for Business, Innovation and Skills (BIS), august 2011

innovation programmes on business. In this way CIM 2.0 complements existing documents and reports.<sup>7</sup>

Although CIM 2.0 lists requirements and standards for impact assessments, the intention has been to do this in such a way that there is room for flexibility and diversity. The purpose of the manual is also to contribute to the international discussion and the knowledge about excellent methods for performance measurements in research, innovation and business policy.

Since excellent impact assessments requires good researchers and evaluators, financing of research and evaluation projects, and unlimited access for researchers to relevant quality data on public programmes and policies, close interaction between researchers and evaluators on the one side and the contractors, programme owners and policy makers on the other side must be established. Excellent econometric impact assessments of public programmes and policies must be built on mutual respect and trust.

The success factors of an excellent impact assessment might also depend on to what extent the evaluation is focused on a relevant and limited set of impact variables and tailored in terms of methods. Furthermore, the relevance and outcome of an evaluation will also depend on the consistency between the evaluation approaches and the requirements of the programme sponsor. In this context a close relationship and high level of interaction between the programme sponsor and the evaluators could lead to the formulation and implementation of an effective and useful evaluation (useful in the sense that the study contributed to policy learning).

This view is also reflected in a large European study on the evaluation practices and methods in the European countries. For instance this point was made by Technopolis in an evaluation of the econometric impact study of the Danish Innovation Consortium Scheme (see the text box below). The study of the Danish Innovation Consortium Scheme can be seen as an example of the type of excellent econometric impact assessments of R&D and innovation policies which is the subject matter of this manual.

**Evaluation brief by Technopolis on: “An Analysis of Firm Growth Effects of the Danish Innovation Consortium Scheme”**

*Interviewer and case study author: Paul Cunningham, Manchester Institute of Innovation Research*

*1. Introductory information*

*The Danish Innovation Consortium Scheme is a national subsidy scheme operated by the Danish Council for Technology and Innovation (RTI) in cooperation with the Danish Agency for Science, Technology and Innovation (DASTI). The econometric impact evaluation of this scheme was titled: “An Analysis of Firm Growth Effects of the Danish Innovation Consortium Scheme” and it was an interim evaluation, commissioned by the Danish Agency for Science, Technology and Innovation. The evaluation was carried out between September 2009 and April 2010 by J.M. Kuhn from the Centre for Economic and Business Research, Copenhagen Business School. The evaluation budget was approximately €40,200. The evaluation examined the economic impact. In a counter-factual analysis, the evaluation examined 220 firms which had participated in at least one Innovation Consortium using a firm-register dataset. The focus is on firm level developments in two success parameters: gross profit and employment. Innovation Consortia subsidise and facilitate cooperation between private firms and research and knowledge institutions, thus contributing to the dissemination of knowledge and research results from the research sector into the business sector.*

*2. Summary*

*The objectives of the evaluated measure were as follows: through the creation of innovation consortia*

---

<sup>7</sup> E.g. *The role of evaluation in evidence-based decision-making*, Department for Business, Innovation and Skills (BIS), august 2010, and *The Green Book – Appraisal and Evaluation in Central Government*, Treasury Guidance, London, United Kingdom, and *The Magenta Book: guidance notes for policy evaluation and analysis*, Government Social Research Unit, HM Treasury, London, United Kingdom (October 2007)

*(consisting of a minimum of two companies, a research institution and a knowledge dissemination party), to promote joint projects that develop and bring research-based knowledge to maturity to the benefit of Danish companies and the Danish business community – especially SMEs. The objectives and main questions of the evaluation were: to determine the economic impact of the measure (in terms of employment growth and gross profit) and on survival rates on participating firms compared to a control group of non-participants.*

*The evaluation methodology was the culmination of a specific research project approach sponsored by the commissioning agency and the evaluator. The approach and main research methods involved a control group comparison on specific growth factors (namely, gross profit and employment). The key information sources included information derived from programme monitoring data (on participants), earlier evaluations (in 2007 and 2008), company financial reports (from an external public domain dataset) and mergers and acquisitions data (again in the public domain).*

*The key findings of the study were that firms that participated in the Innovation Consortium scheme experienced significant increases in the growth of gross profit and employment when controlling for pre-participation growth and developments in the growth of firms in the control group. However, the effects were less marked for larger participating firms than in the case of smaller companies (gross profit <\$20m or <150 employees). No differences in survival rates were detected. In terms of the lessons for evaluating innovation support measures, the (relatively simple and limited) analysis approach was found to be favourable for this type of programme, although regular updates to the data might add additional sophistication – e.g. by identifying those firms that benefit most or which Innovation Consortia work better than others. It should be noted that other attributes of the measure (management and process issues, participant interactions) were also evaluated through separate targeted evaluations.*

### *3. Conclusions and lessons learned*

*The evaluation was successful in that it developed a methodology highly consistent with the objectives set by the commissioning body (the Danish Agency for Science, Technology and Innovation) which addressed a rather narrow set of questions, i.e. to assess the economic impact of the IC Scheme on participating companies. This level of consistency was achieved by the close relationship and level of interaction demonstrated between the evaluators and the contractors and through the opportunity to apply the methodology in two separate evaluations, allowing further refinement of the approaches used. The personal engagement and high level of interest of the key sponsor (the Head of the Centre for Strategic Research and Growth within the Danish Agency for Science, Technology and Innovation) could also be highlighted as a major contributory factor to the success of the methodology. Overall, the entire process was described as “very smooth” by the evaluator. With regard to the approach itself, the programme manager noted that the evaluation had delivered a high degree of utility: it had provided essential answers to specific questions that were required for political justification and budgetary requirements and it had provided a clear quantitative evidence of the programme’s success. Moreover, it had led to the development of a methodology that was both replicable and transferable. As noted it was highly dependent on the availability of good quality comprehensive data on both the participating firms and on the control group of firms required for the counterfactual analysis. This also makes the requirement for such high quality information a potential limiting factor to the adoption of the approach in other contexts. In the absence of existing data sources, it would be necessary to expend more evaluation resources in order to develop the necessary data. Moreover, it would not be routinely updated as is the case in Denmark. However, it should be noted that the evaluator let it be known that the use of the method was also being explored as a potential methodology for evaluating Eureka programme participation. It was also noted that the same approach was being applied in other Danish innovation support schemes such as the Industrial PhDs Scheme, International Collaboration Projects and the Innovation Assistance Scheme.*

*A careful selection process was utilised in defining the group of control firms, which ensured that these were very similar to the participating firms in terms of their characteristics prior to their involvement in the Scheme. This reduced (but did not entirely remove) the possibility that extraneous factors might be responsible for any observed deviations in the economic growth variables between the two groups as a consequence of participation in the Scheme. Since the evaluation methodology was relatively straightforward, its use by other evaluators would not be particularly limited by a need for advanced or specialist competencies. By admission, the evaluation did not address a number of issues that are typically significant for this type of innovation support measure, such as participant satisfaction, uptake statistics and the rationales for participation, management and administrative process issues, knowledge transfer aspects, impacts on other IC participant partners, outputs (e.g. publications, patents) and their associated quality issues, etc. The contractor estimated that it only delivered about 20% of the total amount of information required for the full evaluation of the IC Scheme; however, this was fully recognised. Thus, separate accompanying evaluations have also been*

conducted, using different evaluators and over different timeframes into these specific aspects.

*This approach carries lessons for the evaluation of programmes more generally in that it offers the opportunity to conduct separate, more targeted evaluation studies, performed by different evaluators with distinct capabilities and expertise rather than undertaking a single one-off more comprehensive evaluation. This is particularly useful in cases where timing issues may be important as it avoids having to make a trade-off between carrying out an early-stage evaluation to obtain the management/process information which might indicate the need to change the way in which the scheme is administered, or a mid-term evaluation in which emerging networking and knowledge transfer/output effects might be assessed or finally a longer term evaluation in which various aspects of impacts might be assessed.*

*To summarise, the main conclusions are:*

- *The evaluation was highly focused on a limited set of impact variables and thus was fully tailored (in terms of methods and timing) to assess the programme effects in this context.*
- *The evaluation's success was in part attributable to the consistency between the evaluation approaches and the requirements of the programme sponsor.*
- *The close relationship and high level of interaction between the programme sponsor and the evaluators led to the formulation and implementation of an effective and useful evaluation (useful in the sense that the study contributed to policy learning).*
- *The success of the principal study methodology was predicated on the existence and availability of comprehensive and robust firm-level data.*
- *The study was able to utilise this data to produce reliable comparison data sets of participants and non-participants.*
- *Given the preconditions outlined above, the methodology appeared to offer good scope for its application to the evaluation of other similar schemes.*
- *The broader evaluation approach using variable timings and appropriate methodologies offers lessons for future evaluations.*

*Further information: <http://fivu.dk/en/publications/2010/an-analysis-of-firm-growth-effects-of-the-danish-innovation-consortium-scheme>*

### 3 Overview of important standards and minimum requirements – The 9 principles

The central challenge in carrying out effective impact assessments is to identify the causal relationship between the project, programme and the outcomes of interest. Hence, *CIM 2.0* requirements on excellent econometric impact assessments are (1) a high quality of data, (2) the most recent research-based statistical methods and (3) a high quality of control groups. On this basis, guidelines are set out in *CIM 2.0* in the form of nine principles phrased as minimum requirements on excellent impact assessments.

#### 9 principles: Minimum requirements on excellent econometric impact assessments

<b>Data</b>	<b>1. Data quality and harmonisation of data collection</b>
	<p>Establish standards for data collection, including standards for input data and registration in databases for each research and innovation programme.</p> <ul style="list-style-type: none"> <li>• Standards for data collection are harmonised across all research and innovation programmes in the Danish Agency for Science, Technology and Innovation's database (The Innovation Denmark Database).</li> <li>• Input data should be of the highest quality so it can be linked to national and international statistics and databases.</li> </ul>
<b>Design</b>	<b>2. Selection of comparable companies and/or individuals as control groups</b>
	<p>Selection of comparable companies and/or individuals for control groups must be based on matching as many relevant observed characteristics as possible. The very highest requirements on quality and interpretation of data for control groups must be stipulated.</p> <p>Selection of control groups:</p> <ul style="list-style-type: none"> <li>• Use of the propensity score and nearest neighbour matching method for</li> </ul>



	<p>selection of the most comparable control group</p> <ul style="list-style-type: none"> <li>Alternative methods like lottery and ranking could also be considered, e.g. in robustness tests (see principle 6)</li> </ul>
	<b>3. Difference-in-difference method</b>
	Use of the difference-in-difference method and balanced panel data.
	<b>4. Treatment of outliers</b>
	Outliers must be handled in accordance with the most established international methods in the fields of economic research and econometric methods.
	<b>5. Long-time series</b>
	Ensure high data quality with long time series of at least 6-15 years with a minimum of data gaps in the time series. National statistics on company data and personal data as well as DASTIS' Innovation <i>Denmark Database</i> for research and innovation programmes are to be established with time series of up to 20-25 years depending on the instrument applied.
	<b>6. Robustness test</b>
	Robustness tests are recommended in studies with long time series and many observations. In case of data limitations, e.g. limited time series and observations, it is a requirement that impact assessments be carried out using methods that thoroughly test the robustness of the results.
<b>Inter-pretation of results</b>	<b>7. Relative impact indicators</b>
	The key performance indicators are to be made relative in order to avoid comparison of uneven entities, e.g. through differences in growth rates or other relevant types of ratios.
	<b>8. Peer review of results</b>
	The quality and utility value of impact assessments must be discussed with independent research organisations that are not behind the analyses, e.g. through peer reviews, research seminars, policy maker workshops etc. Preferably, the results of the impact evaluations should be suitable for acceptance by the most reputable international journals.
<b>Critical issues</b>	<p><b>9. Failures and stress tests</b></p> <p>Impact studies also contribute to policy learning. Policy makers and programme owners should accept critical issues, failures of results, collapse of projects and programmes etc. as a part of an evaluation. The contractor of the evaluation should try to learn of failures and not only look for success stories. A close relationship and high level of interaction and mutual trust between the programme sponsor and the evaluator increase the possibility of the formulation and implementation of an effective and useful evaluation, useful in the sense that the study contributed to policy learning and new knowledge about the weaknesses and strengths of the programme.</p>

### 3.1 Principle 1 – Data quality and harmonisation of data collection

Good quality data are required to assess the impact of a research or innovation. At a minimum input data are needed to know when a programme or a policy starts and who participates in the activities. Input data are data collected by programme owners, often at the point of service delivery, as part of regular operations. Both projects that approved and rejected projects should be registered systematically. Data from rejected projects can be useful in selecting a control group. As a standard the following input data are collected for all research and innovation programmes:

- Variables for each project: Name of programme, project title, grant status (rejection or approval), application year, start date for the project, end date for the project, total budget and total grant
- Variables for the participating partners in each project: Company registration number (CVR-number), name, postal code, region, number of employees, industry (NACE code).

- Variables for the participating individuals in each project: Civil registration number (CPR), name, sex, age and education.

The input data for every research and innovation programme in the Danish Agency for Science, Technology and Innovation is harmonized and are part of the InnovationDenmark database.

Furthermore, collection of input data is included in the design of every new research and innovation programme and not when a programme needs to be evaluated.

It is important to determine the data needed and the sample required to precisely estimate differences in impacts between the participants in a programme and the control group. Input data are not sufficient for impact assessments, and therefore national or/and international statistics are used. The following national statistics are used in connection with the impact assessments:

- R&D statistics (Statistics Denmark)
- Accounts statistics (Statistics Denmark)
- Community Innovation Survey (CIS) (Statistics Denmark)
- Education statistics (Statistics Denmark)
- Patent statistics (Statistics Denmark)
- Labour market statistics (Statistics Denmark)
- Salary statistics (Statistics Denmark)
- The Danish Commerce and Companies Agency's Central Business Register / *Købmandsstandens Oplysningsbureau/Experian A/S* (Danish Business World's Information Agency)
- The research indicator (The Danish Agency of Science, Technology and Innovation)

The company registration number or civil registration number is some of the most important input data to collect for each participant because that is the linkage between the input data collected for each programme and national statistics.

It is important as a minimum that sorting of observation and cleaning of data is done in the same way for research and innovation programmes that are to be compared<sup>8</sup>.

### **3.2 Principle 2 – Selection of comparable companies and/or individuals to control groups**

The impact of participating in a research and innovation programme is often indirect and therefore difficult to measure and identify. It is difficult to isolate the actual impact that may be the result of many and varying external factors. It is also difficult to identify the causality of participating in a research and innovation programme.

The selection of control groups is important in relation to the issue of causality. When selecting company or individuals control groups, companies or individuals must be chosen that are more or less equally likely to participate in the programme, yet they do not.

When selecting control groups, it is also important to consider that the companies or individuals that participate in a programme are to be compared with other companies or individuals that are not participating, but are similar in as many relevant parameters as possible that may be of significance to the impact of the analysed programme.

---

<sup>8</sup> Examples of a comparison of programmes is the comparison of ordinary PhDs and industrial PhDs found in DASTI/FI(01/2011), and the comparison of companies participation in EUREKA projects and innovation consortia found in DASTI/FI(15/2011).

Matching methods typically rely on observed characteristics to construct a control group. Finding a good match for each programme, a participant requires approximating as closely as possible the variables or determinants that explain that individual's decision to enrol in the programme.

Even though companies are comparable in observed characteristics, they can be very different in unobserved characteristics such as strategy etc. Hence, the group of companies or individuals participating in a given programme is not necessary comparable with the group of companies or individuals that did not participate. A difference in outcomes between the two groups may be attributable to differences in firms' characteristics as to the programme. The correlation between participation and outcome in the case of selective use is therefore an unreliable estimate of the programme's causal effect.

### **Why is the control group important?**

*The results of an impact assessment cannot be better than the choice of the control group of the evaluation.*

Therefore, a standard is recommended for the establishment of control groups based on a minimum observed characteristics. In this way, a basis is established for making it probable whether there is a causal connection between the factor to be analysed and the performance objective, along with the basis for measuring the isolated impact, but this also depends on the programme that is to be analysed.

<b>Minimum requirements for selecting a control group of companies</b>	<b>Minimum requirements for selecting a control group of individuals</b>
<ul style="list-style-type: none"> <li>• Educational level of the company's employees</li> <li>• R&amp;D intensity</li> <li>• R&amp;D department</li> <li>• Export intensity</li> <li>• R&amp;D investments</li> <li>• Profit, surplus or contribution margin</li> <li>• Company size</li> <li>• Industry affiliation</li> </ul>	<ul style="list-style-type: none"> <li>• Education</li> <li>• Institution</li> <li>• Company size</li> <li>• Industry affiliation</li> <li>• Gender and age</li> <li>• Any other socioeconomic variables, such as salary, background etc.</li> </ul>

It should be assessed whether a control group should be established solely based on R&D-active companies, or whether innovative companies and non-innovative companies should also be included. If the control group consists solely of R&D-active companies, this must be justified, e.g. by the fact that the analysed activity or the analysed programme is not an activity that all companies can launch overnight, but that it is restricted to R&D-active companies only.

This is a strict assumption, which will undoubtedly exclude companies that were predisposed for the analysed activity. Conversely, it may also be a conservative assumption that helps ensure robustness in the results, as it avoids comparison with companies where the probability of participation in the activity in question is very small. By collecting input data for every regional and national research and innovation programmes it is also possible to make a control group of companies or individuals that has not participated in similar programmes as the programme under evaluation.

The minimum requirements are that as many different characteristics as possible are taken into account. However, it is important to avoid including too many explanatory variables, which may give overlapping results, either individually or in combination. By including too many identical variables, there is a risk that multicollinearity will occur along with too great a correlation between the explanatory variables. This means that the parameters become insignificant and the result becomes biased. An example is if R&D intensity is included along with R&D investments, R&D department and company size, as there is interdependency between these variables.

There are different ways to select a control group of companies or individuals which more or less have the same possibility in participating in a research and innovation programme as the once that do.

Above we have described a randomly chosen control group of companies or individuals on observed characteristics. But the observed characteristics can be combined with other factors.

The control group of companies or individuals can be chosen from the companies or individuals that applied, but did not receive funding. There can still be differences between the top companies or individuals in the funded sample and the worst companies or individuals among the unfunded ones so the control group still suffers from selection bias and unobserved heterogeneity.

In the case where the applications get a score the control group could be companies or individuals with projects situated just above and below the funding cut-off, where it is similar ex ante companies or individuals except in their probability of funding.

Another way to find control groups that do not participate, but could have done so, is to have a lottery between the applicants. That is the simplest way to isolate the effect of the programme because it is totally randomly who is funded and who is not. That is possible for research and innovation programmes where there is an evaluation with a minimum score but no ranking, or no evaluation or score of the applications other than they have to fulfil the administration criteria.

### **3.2.1 Control groups may be selected using a so-called ‘propensity score matching’ and ‘nearest neighbour’ method.**

Based on a large number of impact studies and research articles which have been conducted by Danish researchers, R&D consultants and the Danish Council of Technology and Innovation the recommended standard method is the “Propensity Score Nearest Neighbour Matching Method”, which is used to establish and delimit, on a one-to-one scale, the group of R&D-active companies (or innovative enterprises) that participate in an instrument, and a statistically comparable control group of R&D-active companies (or innovative enterprises) that do not participate, but could have done so. It is impossible to find a control group that is completely identical.<sup>9</sup>

The probability models for companies' participation in an instrument, which are used for identification of the factors that have an impact on whether the R&D-active companies are included in the instrument in question, are set out as logistic regressions and used in connection with the Propensity Score Matching method.

In most cases, it will be an advantage to put together a control group that has as many control companies as possible – based on the law of large numbers. Therefore, one-to-one is a minimum requirement, but the standard should be one-to-many. Furthermore, this should be supplemented by balance tests in order to analyse the difference between the treatment group and the control group.

The so-called propensity score matching method is used to match companies or individuals who have participated in the analysed activity with comparable R&D-active or innovative companies or individuals who have not participated in an equivalent activity. The idea of the method is that for a company T, which has the desired activity, a company C is found among the other companies in the relevant statistics, which in a number of statistical parameters resembles company T by having the same probability (‘propensity score’) of taking part in the relevant activity, except that in actual fact,

---

<sup>9</sup> Examples of application of this method are found in DASTI (Innovation: Analyse og evaluering 01/2010): “*Productivity impact of business investments in research, development and innovation*”, DASTI (Innovation: Analyse og Evaluering 02/2011): “*The economic impact of business-research collaboration*”, and DASTI (Innovation: Analyse og Evaluering 03/2013), “*Economic impacts of Business Investments in R&D in the Nordic Countries – A microeconomic analysis*”.

company C has not participated in the activity. In this way, company T (which is designated ‘treatment’ or ‘participating’ company) can be compared to a similar company C (which is designated ‘comparison company’ or ‘control company’), which has been found in the statistics. Statistically, company C must resemble company T in regards to industry, company size, export pattern, staff education, profit, contribution margin and composition as well as R&D activities or innovation activities.

Other methods like establishing control groups based on information from the ranking of projects in the evaluation process can also be used. The advantage of this method is that the control group consists of observations with enterprises and researchers which applied for participation in the programme or initiative but where the application was rejected based on the evaluation criteria. However, obviously not all applicants should be part of the control group. Controls should be based on those applications which had almost the same ranking score in the evaluation process but were rejected due to financial limitations of the support programme. Observations based on the treatment group and the controls should only be from those projects which are just above and just below the threshold leaving the best and the worst projects out of the sample.

However, such an approach requires an excellent and trustworthy ranking system based on a large number of projects which have been evaluated and a large number of observations close the threshold. This type of quality data based on ranking information may exist in the EuroStars-programme of the European Union, the framework programmes of the European Union and in the Innovation Consortia Programme of the Danish Ministry of Science, Technology and Innovation. However, there is no scientific evidence that a control group based on high quality ranking observations produces a better control group than the “Propensity Score Nearest Neighbour Matching Method”.

#### **What is the recommended method to establish a control group?**

*The recommended standard method is the ‘Propensity Score Nearest Neighbour Matching Method’. Other methods like using ranking information can also be used but there is no scientific evidence that a control group based on ranking information produces a better control group than the “Propensity Score Nearest Neighbour Matching Method”.*

It naturally follows that it is not possible by using this or other methods to find a control group that is completely identical in all partially unobservable factors. Another selected control group may give different results. It is therefore important to be able to interpret the characteristics found in the control group.

### **3.3 Principle 3 – Use of the difference-in-difference method and balanced panel data**

The possibilities depend on the design of the innovation policy instruments. For instance, some innovation policy instruments may be able to open up for considerably more precise estimates of the effects than the matching described above and the difference-in-difference method described below. This depends on whether, for instance, a regression-discontinuity design is a possibility.

#### **3.3.1 The difference-in-difference method**

One of the recommended central statistical methods that have been used is the difference-in-difference method. This method is used to calculate differences in the development at the treatment group and the control group of statistically identical companies or individuals without the analysed activity.<sup>10</sup>

---

<sup>10</sup> Examples of application of this method are found in DASTI (Innovation: Analyse og evaluering 01/2010): “*Productivity impact of business investments in research, development and innovation*”, and DASTI (Innovation: Analyse og Evaluering 02/2011): “*The economic impact of business-research collaboration*”.

The difference-in-difference method is based on comparison of changes in output (the performance objective). The model therefore looks as follows:

$$\delta = Y_1^T - Y_0^T - (Y_1^C - Y_0^C)$$

In which  $\delta$  is the effect of the activity, which is calculated on the basis of the difference between the development in the performance indicator, called Y, at the treatment group (T), defined as the performance indicator at time 1 minus the performance indicator at time 0, and the development in the performance indicator at the control group (C), defined as the performance indicator in time 1 minus the performance indicator in time 0. Whether there is a significant difference between the two can be tested subsequently by means of e.g. standard t-tests or linear regression.

### Box 1 Central analysis method: Difference-in-difference

*Difference-in-difference:*

*(a) before-after comparison of companies that participate in the scheme (participant)*

*(b) before-after comparison for companies that do not participate in the scheme (control)*

*See whether (a) is more positive than (b).*

*T1 – success parameter of participant before.*

*T2 – success parameter of participant after.*

*C1 – success parameter of non-participant before.*

*C2 – success parameter of non-participant after.*

*The difference (T2-T1)-(C2-C1) measures the difference in the increases.*

### What is the key method for impact analysis?

*The key method for analysis of the difference between the participants in the policy initiative and the control group which does not participate is the difference-in-difference method.*

### 3.3.2 Balanced panel data

The effect of companies' research and development investments on added value and productivity per employee is a dynamic process, which may vary over time. Cross-sectional analyses based on a single year are not adequate for analysis of the variation over time. Furthermore, there may be unobservable effects on the individual company, which the models are not able to take into consideration. The before and after comparison that results from applying the difference-in-difference method means that panel data (cross-sectional data over time) and methods are needed to check these unobservable effects.

Large companies are included in the research and development statistics every year, while samples from among small and medium-sized enterprises are selected randomly. The result is a very 'unbalanced' panel. For some companies, observations are available for all years, while others only have data for one or a few years. Because of this, it is recommended that the panel data set should be put together as follows:

- Panel data analyses are only to be made for companies with at least two observations. In order to ensure that the analyses are as representative as possible, all companies with two or more

observations are to be included. If the data basis allows for this, the requirements may be made more stringent, so that only companies with three or more observations are included. Naturally, this will reduce the number of companies in the analysis.

- The following approach is recommended for missing observations in time series: If a single observation is missing in a time series, the single missing observation should be estimated. If two or more years are missing in the time series, the most recent continuous part of the time series should be kept.
- Extensive changes in the variables may indicate a merger or division of the company. Such changes may have a disproportionately large effect on the results. It is recommended that the standard in part of the international literature be followed, and that companies with annual growth rates in added value, fixed assets, number of employees or R&D capital of less than - 50 % or more than 300 % be removed. This is in accordance with the standard set out in international literature.
- It is recommended that sensitivity analyses be carried out when basic data are changed.

### 3.3.3 Standards for calculating economic effects

The Cobb Douglas productivity function is used as a standard for indicating the effects of a given instrument in pounds and pence in the form of increased productivity per employee, profit etc. This is typically modelled as an OLS regression.<sup>11</sup>

Depending on the chosen key performance indicator (the analysed success variable), changes of levels over time may also be relevant. An example of changes in levels would be changes in the number of employees and in the level of employment.

An example of relative changes would be the survival rate of companies or the employment quotas. Examples of changes in growth rates are the growth in productivity per employee, the growth in turnover or the growth in added value in the companies. In general, the standard for calculating the economic effects depends on the key performance objectives that are assessed and estimated.

When selecting background factors, it is important to consider how the individual background factors influence both outcome and treatment. For instance, there may be a time-related challenge with background variables, which might be affected by treatment in a model that includes lagged variables.

### 3.4 Principle 4 – Treatment of outliers

Outliers are observations which do not follow the pattern of the other observations. In order for results to be as representative as possible, econometric models should be able to measure effects in a wide range of companies. However, extreme values may distort the effects and reduce precision. In some cases, there may be good reasons for removing extreme values. An example is young companies where large and risky investments are made, which affect the companies' added value for a short period of time. Such companies will potentially experience extreme increases from one year to another.

However, whether or not extreme values should be removed depends on the purpose of the analysis and the innovation policy instrument. Therefore, a careful assessment of outliers should be made for each analysis and each instrument before any decision to exclude these from the analysis is made.

Furthermore, data have been found to include extreme values measured against e.g. companies' average productivity per employee, added value, employment etc. These are assumed to be incorrect

---

<sup>11</sup> Examples of the application of this method may be found in DASTI (Innovation: Analyse og Evaluering 02/2011): *“The economic impact of business-research collaboration”*, and DASTI (Innovation: Analyse og Evaluering 03/2013), *“Economic impacts of Business Investments in R&D in the Nordic Countries – A microeconomic analysis”*.

registrations, which are connected either to the company's added value or to the number of full-year equivalents. Regardless of where the incorrect registration is found, it is recommended that such values be removed from the data. However, there may be other methods, e.g. to include or exclude extreme data to see whether this has any effect on the results, or to consider medians etc.

### **3.5 Principle 5 – Long-time series**

Impact assessment is complex as a linear connection between the analysed activity and a subsequent effect is hardly ever found. Therefore, there are a number of conditions that may make it difficult to measure the impact, such as potential time layers before the impact sets in, different starting points for the companies, differences between the companies' characteristics and the companies' experience and competences in relation to the instrument.

As a standard, the econometric models must therefore be able to make allowance for time lag between the analysed activity and the impact thereof. The impact may set in with varying delays.

### **3.6 Principle 6 – Robustness test**

The researchers checked the robustness of results to various matching procedures. They stress that their estimates might be biased because of some unobserved characteristics. Indeed, when using matching methods we can never rule out bias caused by unobserved variables, and that is their most serious limitation.

### **3.7 Principle 7 – Impact indicators should be made relative**

It is not always easy to identify and delimit effects. Furthermore, differences occur in valuation depending on players and stakeholders. In this connection, an example is related to a company's market value. One suggestion is to use the market's valuation of the individual company as a measure for the price or value of the total 'tangible' and 'intangible' assets. However, this would require the companies in the analysis to be quoted on the stock exchange. Therefore, this method is not used, as most companies are not quoted on the stock exchange.

When effects in companies are to be analysed, it is recommended that a key performance indicator be used, which is made relative in relation to labour input. By making the indicator relative in relation to labour input, it is ensured that the effects cannot be attributed to an endless supply of labour.

### **3.8 Principle 8 – Peer review of results**

The quality of impact assessments must be discussed with independent research organisations that are not behind the analyses, e.g. through establishing steering groups or conduction peer reviews, research seminars, policy maker workshops etc. The intention is to carry out impact assessments based on the best and most accepted international research-related and statistical methods. Preferably, the results of the impact assessments should be suitable for acceptance by the most reputable international journals or at high-level international conferences.

### **3.9 Principle 9 – Failures and stress tests**

Impact studies also contribute to policy learning. Policy makers and programme owners should accept critical issues, "stress tests" of the programme, failures of results, collapse of projects and programmes etc. as a part of an evaluation. The contractor of the evaluation should try to learn for failures and not only look for success stories. A close relationship and high level of interaction and mutual trust between the programme sponsor and the evaluator increase the possibility of the formulation and implementation of an effective and useful evaluation, useful in the sense that the study contributed to policy learning and new knowledge about the weaknesses and strengths of the programme. The impact analyses should not only be restricted to areas and impact indicators where there is a high likelihood for success of the programme but the contractor of the evaluation and programme owners should also look for failures and areas where the programme does not create desired impacts and activities.



## **PART II - Standard for performance objectives: Key performance indicators**

Although CIM 2.0 lists principles and minimum standards for impact assessments, the intention has been to do this in such a way that there is room for flexibility. This is partly because the recommended “Propensity Score Nearest Neighbour Matching Method” will not be the most relevant method for all types of research and innovation programmes and policy initiatives. This will be the case, for instance, when the direct business-related performance objectives are not the object of analysis, but where the study has a wider focus. In the case of other programmes and initiatives, it will be less relevant to make impact evaluations of the listed economic performance targets, because the main purposes of a programme and initiative may be other non-economic activities.

This is the case e.g. when impact evaluations are carried out in relation to clusters and innovation networks, research platforms and networks etc. where the main objectives are not necessarily economic performance targets alone, but may also include non-economic behaviour-regulating performance objectives. It is therefore important that concrete impact evaluations take into consideration what the objective of a given programme or initiative is.

At the same time most of the countries do not have the same possibilities as Denmark and for instance Norway, Sweden and the Netherlands, due to limited access to national wide quantitative micro data and very long time series. In the majority of the countries, it is difficult to establish the micro data for the whole economy basis needed for carrying out solid and validated quantitative econometric analyses that can document and calculate the effects of businesses' research and innovation policy historically.

Therefore, the manual also includes an overview of the non-economic performance indicators for the most important research and innovation programmes and policy initiatives of the Ministry of Science, Innovation and Higher Education.

The Ministry of Science, Innovation and Higher Education has headed an international working group on performance indicators in the EU Pro-INNO project called EPISIS. This working group had participants representing government agencies, ministries and researchers from 8 countries including Denmark, Sweden, Norway, Slovenia, Austria, Germany, United Kingdom and Finland as well as the European Commission. Good practice on evaluations and performance indicators was exchanged, and a manual was elaborated with recommendations for indicators that can be used for setting out performance objectives and key performance indicators.<sup>12</sup>

The EPISIS report from 2011 on Impact Analysis and Assessment Indicators was also an integrated part of the work of a group of policy makers, programme owners and researchers from Denmark, Estonia, Finland, Germany, Iceland, Lithuania, Norway, Poland, Sweden, United Kingdom and the United States of America which met in Copenhagen and Berlin in 2011 and 2012 in order to develop a common understanding on cluster and network policies and world-class benchmarking, monitoring, evaluation and impact assessment tools and indicators. As a result of this work a number of recommendations on possible key performance indicators, impact assessment methods and benchmarking and monitoring tools were recommended in the publication “*Let's make a perfect cluster policy and cluster programme – smart recommendation for policy makers.*” Based on international best practice a perfect evaluation league consisting of benchmarking, monitoring and impact analyses in the following three levels was recommended in the publication.<sup>13</sup>

---

<sup>12</sup> EPISIS Report (2011), *Service innovation: Impact Analysis and Assessment Indicators*.

<sup>13</sup> Gerd Meier zu Köcker, Thomas Lämmer-Gamp and Thomas Alslev Christensen: “*Let's make a perfect cluster policy and cluster programme – smart recommendation for policy makers,*” Berlin/Copenhagen 2012, VDI/VDE-IT and DASTI, October 2012, page 39-47.

### 3.10 Ex ante evaluation

An account of objectives and expected effects of the programme should be given for each research and innovation programme in separate performance descriptions. Thus, the performance description should include, among other things, an ex-ante evaluation of the programme.

On this basis, the Danish Agency for Science, Technology and Innovation sets out key performance indicators for each research and innovation programmes. These indicators can be key performance indicators in the form of so-called input, output, outcome and impact indicators. It is important to set out the key performance indicators in the design of every new research and innovation programme and not when a programme needs to be evaluated.

The assessments and recommendations for the selection of indicators follow the work from the EPISIS project<sup>14</sup> and the work in connection to the recommendations for cluster programmes and policies<sup>15</sup>, which focus on innovation in services and creation of effective R&D and innovation cluster programmes, respectively, but can be inspiration to other programmes, as well as national legislation.

#### 3.10.1 Input and output indicators

The aim of the input indicators to R&D and innovation is to capture the purposeful act that a company undertakes, in order to innovate. Thus, the input indicators measure the intentional part of the innovation process in companies, and is represented by the resources and effort, which companies spend and put into the process. There are many sources of the intentional part of an innovation process, and the input and output indicators are accordingly divided into groups representing management, employees, networks, etc. It is an important point that the input indicators should measure whether companies are working *intentionally* to promote innovation, and that it does *not necessarily* follow that this process leads to innovation. Therefore, an indicator should not be judged upon whether or not the result is certain to be innovation. Rather, it should be considered whether or not the input indicator actually measures the *intentionality* in the process of innovation in a company.

### 3.11 Baseline measurement at ex post evaluation

Emphasis is placed on ensuring baseline measurements of the efforts in order to be able to document the situation before the launch of research and innovation programmes and the situation if the programme had not been implemented, so that the effect of the research and innovation programmes can be estimated in relation to the situation where the programmes did not exist.

In this connection, the most recent research-based methods are applied by choosing advanced control groups that represent the situation if the programme had not been implemented. If the analysis includes a sufficiently large number of indicators, the propensity score matching method can be used for making baseline measurements, cf. principal 2. On this basis, ex post evaluations can be carried out with estimations of the effects of the programmes.

On this basis, the Danish Agency for Science, Technology and Innovation sets out key performance indicators for each research innovation programme. They can be key performance indicators in the form of both so-called outcome and impact indicators.

The assessments for the selection of indicators also follow the work from the EPISIS project mentioned above.

---

<sup>14</sup> EPISIS Report (2011), *Service innovation: Impact Analysis and Assessment Indicators*.

<sup>15</sup> Gerd Meier zu Köcker, Thomas Lämmer-Gamp and Thomas Alslev Christensen: “*Let’s make a perfect cluster policy and cluster programme – smart recommendation for policy makers*,” Berlin/Copenhagen 2012, VDI/VDE-IT and DASTI, October 2012, page 39-47.

**Table 4.1 Recommended input, output, outcome and impact indicators to measure the effect of research and innovation programmes**

Input	Output	Outcome	Impact
Investments	Activities	Results	Effects
Development	Behaviour		
<ul style="list-style-type: none"> <li>• Public R&amp;D expenditure in relation to GDP</li> <li>• Public R&amp;D expenditure by scientific fields</li> <li>• Public R&amp;D expenditure for selected strategic areas</li> <li>• Private R&amp;D expenditure in relation to GDP</li> <li>• Private R&amp;D by sector</li> <li>• Offshoring of research</li> <li>• Private expenditure for innovation (excl. R&amp;D expenditure)</li> <li>• EU's 7th Framework Programme</li> <li>• Funding and grants from the European Research Council</li> <li>• Funding and grants from the European Institute of Innovation and Technology</li> <li>• External funding of university research</li> <li>• Flow of resources from research-funding to research-performing</li> <li>• R&amp;D employees in the public sector</li> <li>• R&amp;D employees in the public sector by scientific fields</li> <li>• R&amp;D employees in the private sector</li> <li>• R&amp;D employees in the private sector by industry</li> <li>• Private expenditure for innovation</li> <li>• Share of companies with high skilled employees</li> <li>• Share of academics in the workforce</li> </ul>	<ul style="list-style-type: none"> <li>• Granted PhD certificates</li> <li>• PhD certificates by scientific fields</li> <li>• Average annual growth in number of publications (15 countries most active in publishing)</li> <li>• Number of research publications</li> <li>• Number of research publications in leading scientific journals</li> <li>• Share of research publications with cooperation by scientific fields</li> <li>• Share of research publications with cooperation by type of partners</li> <li>• Research publications with co-authors</li> </ul>	<ul style="list-style-type: none"> <li>• Citations of scientific journals</li> <li>• Impact of publications with respect to scientific fields (Nordic countries)</li> <li>• EPO-patents and licences</li> <li>• Public research institutions license agreements, patent applications and spin-outs</li> <li>• Share of companies with innovation activities</li> <li>• Share of PP innovative companies</li> <li>• Share of PP innovative companies with innovation cooperation</li> <li>• The 100 most R&amp;D active companies ("EU Industrial R&amp;D Investment Scoreboard")</li> <li>• Innovations and new technologies</li> <li>• Start-ups</li> </ul>	<ul style="list-style-type: none"> <li>• Growth in labor productivity</li> <li>• Growth in Total Factor Productivity</li> <li>• Growth in real wages</li> <li>• GDP growth</li> <li>• Employment growth</li> <li>• Growth in trade turnover or gross profits</li> </ul>
Economic investments	Non-economic investments	Non-economic additionalities	Economic additionalities

How to distinguish between input, output, outcome and impact indicators to measure the effect of research and innovation programmes can best be explained by looking at table 4.1 which provides an overview of the different types of indicators.

### 3.11.1 Outcome and impact indicators<sup>16</sup>

Outcome indicators deal with the newness of the innovation, the outcome indicators can be said to measure intermediate outcome of the innovation process. Therefore, outcome is about the value creation activity rather than the actual value capture of the innovation process, which is measured with impact indicators.

The impact indicators seek to measure the value that a service innovation generates and benefit the company, the customers as well as society as such. For the company, the value of the service innovation can lie in the ability for the company to sell a new service to their customers, but the value can also appear in the form of reduced costs or expenditures internally. For the customer, the value comes from the offering of an improved service compared to the existing offerings or from a service that is new to the market, hence offering the customer a larger variety of services to choose from. The latter is known from academic literature on international trade as the concept of “love of variety” (Krugman, 1979b). The increased value for society follows from the value experienced by companies and customers. Thus if a company e.g. is able to increase productivity, this is in turn valuable for society as it will have a positive impact on GDP. Furthermore, innovation by companies increases competition in the market, ensuring that the strongest companies survive.

## 4 PART III – Overview key performance indicators, impact assessments and results

### 4.1 Results of impact assessments in Denmark

More than 17 impact evaluations of various research, innovation and education programmes have been conducted since 2008. The impact assessments have been carried out by independent researchers or organisations and were commissioned by the ministry, the agency, the research councils or by independent institutions. 12 major impact assessments of innovation policy instruments were conducted alone in 2010 and 2013<sup>17</sup>. Furthermore, DASTI is expected to publish 5 new impact studies in 2014.

Examples of impact assessments are the following: The productivity impact of Danish business R&D and innovation investments, the Innovation Consortia Scheme, the knowledge pilot (Innovation Assistant scheme), the Incubator Programme, the Industrial PhD Scheme, the Innovation Networks Denmark Programme, EUREKA projects and business-research collaboration projects between universities and enterprises and the Danish GTS-system.

#### Box 5.2 Overview over impact assessments of programmes and policies

Focus area	Cluster and network policies
Study no. 1	<i>The independent impact evaluation of Cluster Policy in Denmark (The Innovation Networks Denmark Programme) (DASTI 18/2011)<sup>18</sup></i> : The programme support the establishment and running of cluster and network organisations. Among 1200 non-innovative enterprises participating in the programme, the likelihood of becoming innovative increased 300 percent compared to 1200 statistically identical enterprises not participating in the Innovation Networks Denmark infrastructure. Among R&D-active or innovative enterprises participating in the programme, the likelihood of initiating their first R&D collaboration project with a research institution increased 300 percent compared to statistically identical enterprises not participating in the programme. A new study (see study no. 18) finds that companies participating in the Innovation Network programme, on average, grow approximately 4 percentage points faster,

<sup>16</sup> EPISIS Report (2011), *Service innovation: Impact Analysis and Assessment Indicators*.

<sup>17</sup> <http://fivu.dk/en/publications/2012/central-innovation-manual-on-excellent-econometric-impact-analyses-of-innovation-policy-cim>

<sup>18</sup> <http://fivu.dk/en/publications/2011/the-impact-of-cluster-policy-in-denmark> and <http://fivu.dk/en/publications/2011/innovation-network-denmark-performance-accounts-2011>

	with respect to total factor productivity, the two years following participation compared to similar non-participating companies.
<b>Focus area</b>	<b>R&amp;D collaboration projects between business and research</b>
Study no. 2-4	<i>Three independent impact evaluations (DASTI 06/2008, DASTI 03/2010, DASTI 01/2011<sup>19</sup> and Kaiser &amp; Kuhn (2012)) of the Danish Innovation Consortium Programme</i> (public grants to large research collaboration projects between several enterprises and research institutions and technology institutes) show that there are statistically significant impacts for enterprises as well as for individual researchers depending on the key performance indicators to be analysed. Key performance indicators are for gross profits, individual employment, and employment in enterprises, patenting activity, salary and total factor productivity. Some of the analysis show positive and statistically significant impact for small and medium sized enterprises with respect to labour productivity, patenting activity and employment. None shows impact on total factor productivity or on large companies. One study shows positive, statistically significant impact on the level of salary of researchers at the research institutions. Gross profits increased by EUR 2.7 millions in the average enterprise participating in an innovation consortium over a period of nine years after the innovation consortium started. Enterprises did not receive public grants.
Study no. 5	<i>An independent impact evaluation (DASTI 17/2011) of international research and development collaboration projects (EUREKA-projects)</i> was conducted in 2011 <sup>20</sup> . The impact of EUREKA-participation with respect to labour productivity, employment, turn-over and exports were analysed. The analysis shows a positive, statistically significant impact on growth rates in labour productivity, employment, turn-over and exports compared to statistically similar enterprises not participating in EUREKA-projects. EUREKA-participation also results in significantly higher growth rate in exports and employment compared to enterprises only participating in the Innovation Consortium Programme (and not in international projects).
Study no. 6	<i>An independent impact evaluation (DASTI 02/2011) on national research and innovation collaboration projects between enterprises and universities or GTS-institutes</i> (in Danish) was conducted in 2010 and 2011. Projects are projects without or with grants from public research funding bodies. More than 1,500 R&D-active enterprises engaging in one or more R&D-collaboration projects with research and technology institutions in the period 1999-2006 were compared to more than 1,500 statistically identical enterprises without collaboration found among 20,000 Danish R&D-active enterprises. The labour productivity is 9 percent higher for the average enterprise with R&D collaboration compared to statistically identical R&D-active enterprises without any collaboration in the analysed period. The analysis also looks at differences across branches, types of enterprises and types of research institutions. Impacts are higher in large enterprises than in small enterprises. Impacts are also higher in export enterprises compared to non-exporting enterprises. Finally, impacts increase with the level of skills in the enterprises.
<b>Focus area</b>	<b>Education and academics (candidates and PhD's) in the business sector</b>
Study no. 7-9	<i>Three independent impact studies of the Danish Industrial PhD Programme (DASTI 2007, DASTI 01/2011<sup>21</sup> and The Effect of the Industrial PhD Programme on Income and Employment 2013)</i> show positive, statistically significant impacts. 200-300 participating enterprises as well as for 400 participating researchers depending on the key impact indicators are analysed. The programme provides a subsidy to enterprises hiring PhD-students to work with a PhD-project. Key performance indicators are labour productivity, individual employment, and total employment in enterprises, patenting activity, individual salary and total factor productivity. The 01/2011-analysis shows positive and statistically significant impact for small and medium sized enterprises with respect to labour productivity, patenting activity and employment compared to statistically similar enterprises without Industrial PhD-projects. Patenting activity

<sup>19</sup> <http://fivu.dk/en/publications/2010/an-analysis-of-firm-growth-effects-of-the-danish-innovation-consortium-scheme>

<sup>20</sup> <http://fivu.dk/en/publications/2011/economic-impact-of-international-research-and-innovation-cooperation>

<sup>21</sup> <http://fivu.dk/en/publications/2011/analysis-of-the-industrial-phd-programme>

	nearly doubled and employment is nearly 2 persons higher per PhD-project per year. Both analyses show positive impact for individual employment and salaries in enterprises. None shows impact on total factor productivity or on large companies. Other studies of the Industrial PhD scheme on income and employment impacts have also been conducted in 2012/2013. <sup>22</sup>
Study no. 10-11	<p><b><i>An independent impact evaluation of the Danish Knowledge Pilot (Innovation Assistant) Scheme (DASTI 04/2010)</i></b> shows that there are positive but no statistically significant impacts for enterprises. Gross profits increased EUR 156,000 in average over three years after the knowledge pilot project started. The scheme provides a subsidy to SME's hiring academics of up to EUR 20,000. Key performance indicators analysed are gross profits, total employment and survival rates of enterprises. Because of the lack of sufficient data and observations a new independent impact evaluation has been conducted.</p> <p><b><i>An evaluation of the Danish Innovation Assistant Programme (Videnpilotordningen) (DASTI 12/2013<sup>23</sup>)</i></b> shows that there are positive short-term employment effects for the innovation assistants, but no statistically significant impacts for enterprises. The scheme provides a subsidy of up to EUR 20,000 for SMEs hiring university graduates. Key performance indicators analysed are gross profits, added value, return on assets, labour productivity, total employment and survival rates of enterprises. These results are confirmed by a new study (see study no. 18) that finds that companies participating in the Innovation Assistant programme, on average, grow approximately 3 percentage points faster, with respect to total factor productivity, the two years following participation compared to similar non-participating companies.</p>
Study no. 12	<b><i>An independent study of the impact of PhD-candidates on productivity in enterprises (DASTI 2012, prepared by CEBR – Centre for Economic and Business Research at CBS, Copenhagen, 23. September 2011)</i></b> shows that the average labour productivity in enterprises with minimum one PhD-candidate is approximately 34 percent higher compared to enterprises with the same mix of educations and skills but without a PhD-candidate. The impact of PhD-candidates seems to be smaller in small enterprises than in larger enterprises. The average labour productivity difference for small enterprises with and without PhD-candidates is 11 percent. The salary of PhD-candidates is approximately 10 percent higher than the salary of non-PhD-individuals with same educational background, age and sex and working in the same type of enterprise and business sector.
Study no. 13	<b><i>The Report on 'Productivity and higher education' has been conducted by the Centre for Economic and Business Research (CEBR) for the Danish Business Research Academy (DEA) in 2010.</i></b> The effect of different types of highly-educated working capacities on the productivity (added value) in 138,372 Danish enterprises over a nine year period (from 1999 to 2007) is analysed. The analysis shows that the productivity for each individual becomes increasingly higher, the longer the person's educational background is, regardless of the field of education. Education within social sciences results in the highest individual productivity. Technical, health sciences and life-science result in a slightly lower productivity than the social sciences. One percentage point increase in the share of employees with a long-cycle higher education will cause an increase in the Gross National Product by approximately 1 percent.
Study no. 14	<b><i>An analysis of the macroeconomic benefits of attracting international students (2013).</i></b> The Danish Ministry of Science, Innovation and Higher Education has commissioned an analysis of the socioeconomic benefits of attracting international students to complete full Master's degree programmes in Denmark. The analysis has been carried out by the Danish Rational Economic Agents Model (DREAM) group. The analysis shows that attracting international students with a Bachelor's degree level of education to Denmark to complete a full Master's programme is socioeconomically beneficial. This is true both of students who are subject to fees when attending a Master's programme in Denmark and of students from other EU/EEA countries which attend Master programmes in Denmark based on financing from the Danish state. Attracting 1,000

<sup>22</sup> <http://fivu.dk/en/publications/2013/the-effect-of-the-industrial-phd-programme-on-employment-and-income>

<sup>23</sup> <http://fivu.dk/en/publications/2013/an-evaluation-of-the-danish-innovation-assistant-programme-en-effektmaaling-af-videnpilotordningen>

	additional international students annually results in a lasting improvement of public finances of between DKK 0.4 billion and 0.8 billion. The revenue primarily results from tax revenue generated by graduates who choose to stay in Denmark. The DREAM model is built and the calculations are made on the basis of experience with the behaviour of consumers, businesses and so on. An important assumption in these particular calculations is that the share of students which stay in Denmark after their graduation remains the same as in recent years. <sup>24</sup>
<b>Focus area</b>	<b>Commercial exploitation of public and private research and inventions</b>
Study no. 15	<i>An independent impact evaluation of the Incubator Programme (DASTI 01/2010)</i> shows that there are no statistically significant impacts for more than 300 enterprises and more than 300 entrepreneurs. The scheme provides public risk capital to the establishment of new knowledge intensive enterprises. Key performance indicators analysed are individual salaries, revenue, added value, total employment and survival rates of enterprises. Because of the lack of sufficient data and observations a new independent impact evaluation is conducted in 2014. The focus of the upcoming study is impacts at the level of enterprises.
Study no. 16	<b>Public research – effects on innovation and economic growth (2012<sup>25</sup>)</b> The Danish Ministry of Science, Innovation and Higher Education has commissioned an analysis of the effects on innovation and economic growth of public research. The report which is written in Danish builds on results from a number of international analyses which gives an overview of the types of contributions from publicly funded R&D to economic growth. The results are as follows: The estimated yearly benefit to the Danish economy from public investments in R&D is estimated at 20-40 percent. A DKK 1 billion R&D-investment result in growth in the Danish economy measured as gross domestic product (GDP) of DKK 1.2-1.4 billion. It must be mentioned that the causalities behind these estimates are difficult to assess. The relation between publicly funded R&D and education is a key issue in this context. Engineering, natural sciences and health sciences seem to make the most obvious types of contributions. The time elapsed from when research results are discovered and translated into industrial commercialization is estimated to vary between 6 and 20 years, depending on the research field. The period seems to have shortened over time. The report is based on three underlying reports which can be found via the links below. The first underlying report is authored by PhD Rikke Nørding Christensen, Professor Svend Erik Hougaard Jensen, CBS, Professor Keld Laursen, CBS and Professor Michael S Dahl, AAU and contains a review of Danish and international literature about the economic growth effects from publicly financed R&D. The second underlying report is authored by PhD Rikke Nørding Christensen and Professor Svend Erik Hougaard Jensen, CBS, and contains an analysis of the types of contributions from publicly financed R&D to the Danish economy from different research fields. The third underlying report is made by Senior Analyst Johan Moritz Kuhn and Senior Analyst Martin Junge, CEBR, and contains analyses of the PhD-employment, wage level and productivity in Danish firms.
<b>Focus area</b>	<b>Meta-evaluations</b>
Study no. 17	<i>DASTI (Innovation: Analyse og Evaluering 03/2013), “Economic impacts of Business Investments in R&amp;D in the Nordic Countries – A microeconomic analysis”</i> estimates labour productivity elasticities of R&D investments. The estimated R&D labour productivity elasticities are positive and significant for the enterprises in Denmark, Finland, Norway and Sweden. Furthermore, the sizes of the elasticity are within the normal bounds as described by the literature. When comparing the R&D elasticities, it is apparent that the elasticities vary across the Nordic countries. Denmark and Sweden has the lowest R&D elasticities of 0.091 and 0.107 respectively. There is no significant difference in the elasticities of the two countries. Conversely, Norway and Finland have the highest R&D labour productivity elasticities of 0.129 and 0.167 respectively. When comparing the rate of return of the last invested euro in business R&D across the Nordic countries, it is apparent that the rates of return vary across the Nordic

<sup>24</sup> <http://fivu.dk/en/publications/2013/analysis-of-the-macroeconomic-benefits-of-attracting-international-students>

<sup>25</sup> <http://fivu.dk/en/publications/2012/public-research-2013-effects-on-innovation-and-economic-growth>

	<p>countries. Swedish companies have the lowest return of investments. The median enterprise in Sweden obtains a rate of return of the last invested euro in R&amp;D of 16.4 percent. The return of R&amp;D investments is 34.2 percent for Denmark and 22.7 percent for both Finland and Norway. A test for different return of investment shows that the estimated rate of returns of business R&amp;D is significantly different across the Nordic countries. This result indicates that Danish enterprises realize a significantly higher marginal return of investment in business R&amp;D compared to enterprises in Sweden, Norway and Finland. The report also shows that the results vary significantly across business sectors and size of enterprises.</p>
Study no. 18	<p><b><i>The Impact on Productivity Growth of the Danish Innovation and Research Support System (forthcoming)</i></b> The Danish Agency for Science, Technology and Innovation has commissioned an analysis of the joint effect of the Danish innovation and research. The analysis is the first of its kind to perform a joint estimation of the economic impact of the Danish innovation and research programmes on companies' total factor productivity growth. The analysis compares the productivity growth of companies that received support from a programme with similar companies that did not received support. The analysis finds that companies participated in a programme, on average, grow 2.5 percentage points faster the two years following participation compared to non-participating companies. Furthermore, the analysis shows that programmes (i.a. Innovation Network, Innovation Voucher) that are designed to spur an increase in knowledge via Collaboration and Counselling and support increase participating companies' productivity growth by 2 to 3.75 percentage points. The estimation sample contains around 1,100 observations covering the period 2000 to 2011. The estimation sample is created by combining data from the InnovationDenmark database (see page 9) and worker-firm matched registry data from Statistics Denmark.</p>