

# Some thoughts on evaluation

Torbjørn Hægeland

# The ideal evaluation setting – from the perspective of an evaluator

- Objective: To establish whether something "works" or not
- Strategy: Perform a randomised experiment
  - Divide the relevant population in two parts A and B – by lottery
  - Give A the "treatment" (e.g. support, access to a scheme)
  - B receives no treatment.
  - Difference in outcome between A and B is causal effect of treatment
- This strategy is rarely used in policy evaluation
- Policy measures mostly general or targeted
- Must look for quasi-experiments or natural experiments
- Be cautious in interpreting findings as causal effects

# The typical evaluation setting

- Tax incentives for R&D - and other policy measures - are often general in nature, based on a belief that the scheme is going to work in the desired way, and therefore should be available to all.
- However, policy makers usually want to or are obliged to evaluate the scheme
  - Because they are to some extent uncertain whether it works or not
  - Part of a compromise
- The main challenge for evaluators is that the more general a scheme is, that is the more equal the scheme treats different firms, the more complicated is the evaluation in terms of providing answers on causal effects.

# The typical evaluation setting – (cont.)

- The reason is that a higher degree of “generality” or “equal treatment” brings evaluations further away from the ideal setting where firms, under otherwise similar conditions, could be compared according to whether they are eligible for the scheme or not.
- Simple comparison of users and non-users will not work – self-selection based on cost-benefit considerations
- (Same problem applies in targeted schemes – here also selection by the granting agency)
- Ideally, evaluations should be built into the design of the scheme itself. However, an evaluator typically comes in after the scheme has been designed and implemented.

# What is SkatteFUNN?

- A GENERAL MEASURE:
  - A system for tax deduction of R&D expenses in all enterprises
- 18 percent tax deduction of R&D expenses up to NOK 4 millions. If joint project with an approved R&D institution, the cap is NOK 8 million
- Smaller enterprises 20 percent deduction
  - <250 employees, turnover < € 250 mill or balance sheet total < € 27 mill, not owned more than 25 percent by larger enterprise
- Launched in 2002 – SMEs only
- Extended to all enterprises in 2003

# Some key figures

	2002	2003	2004	2005	2006
Number of applications	3287	4739	4225	3176	2545
Approved applications	2798	3532	2762	2177	1772
R&D approved (mill NOK)	4104	7459	8155	7065	6677
Tax deduction (mill NOK)	690	1257	1382	1103	1082
Pct. paid as grants	82	78	76	74	73

# Main issues in the evaluation

- Input additionality: Does SkatteFUNN generate more R&D?
- Behavioural additionality: Does SkatteFUNN generate change in R&D behaviour in enterprises?
- Returns: How do SkatteFUNN projects pay off?
- Real R&D or reclassification of costs?
- Does SkatteFUNN stimulate knowledge transfer from R&D institutions to enterprises?
- How does SkatteFUNN work together with other R&D stimulating measures?
- Administrative costs and efficiency

# Data needs

- To study the effects of a tax incentive scheme for R&D, one needs a fairly long time horizon because the total effects of R&D cannot be expected to be visible in the short run.
- The data should cover a sufficiently long period both prior to and after the launching of the scheme.
- What should be regarded as “sufficiently long” can in principle not be established *ex ante*.
- In practice, one has to make a compromise based on the availability of historical data, the cost of gathering new data, and the time horizon for the evaluation project.
- Granting agency must devote significant resources to data collection

# Data sources

- Extensive use of comprehensive micro data bases for firms and plants
  - Many based on administrative registers
  - In addition: Manufacturing censuses, R&D surveys innovation surveys etc.
  - Use of common identifiers facilitates extensive linking of different data sources
  - Results in a unique panel data base for firms and plants, covering also a long period prior to the launching of the SkatteFUNN scheme

# Data sources (cont.)

- The SkatteFUNN base (RCN)
  - contains all relevant information from firms' applications for tax exemption, such as description of the project etc, from the processing of the application in RCN, and from the firms' reports after the completion of the supported R&D project.
- Data from tax authorities
- R&D surveys questions related to the SkatteFUNN scheme.
  - This is done mainly to provide data for official statistics, but the information will be useful also for evaluation purposes.
- Specific surveys covering behavioural additionality and administrative aspects of the scheme.
- These data are available to other researchers on the same conditions as other micro data

# Example of evaluation strategy: Input additionality

- A key point in the evaluation
  - Important criterion in itself, cf. policy goals
  - Additionality a "necessary condition for success"
- Counterfactual question: What would R&D investments have been without SkatteFUNN?
- Non-experimental setting – no formal control group

# Identification

- To answer what would have happened without SkatteFUNN requires analysis of microdata with a control group that can tell us what forms would have done in the absence of the scheme
  - Relevant comparison is not **”present vs past”**, but **”present vs counterfactual present”**
- Other macro-conditions than skatteFUNN may affect R&D investments.
- The design of the scheme precludes any **”smoking gun”** evidence

# Additionality (cont.)

- Before-after studies
  - Predict R&D investments using firm-specific variables, business cycles, tax rules etc.
  - I.e: Try to Control for the effects of other factors on R&D, attempting to isolate the effect of the SkatteFUNN scheme
- "Quasi-experiments"
  - Utilize discontinuities in the scheme
    - ◆ Differences between small and large firms (particularly in 2002)
    - ◆ Limits on the tax deduction gives difference in incentive effect

# Main strategy

- Firms that would have done more R&D than the cap (4 mill NOK) in the absence of the scheme did not get any extra incentive to increase R&D investments
- We do a "difference-in-difference"-analysis
  - Assume that firms above the cap is a valid control group telling us what the growth in R&D for firms below the cap would have been in the absence of the scheme.
  - Compare the change in R&D investments for firms above and below the cap

# Problem I

- The relevant is who would have been above or below the cap *in the absence of the scheme*
- Classify firms according to average level before the launching of the scheme
  - **Not perfect:** Changes not due to the scheme contaminate the treatment and control groups which underestimates the effect. "Regression-to-the-mean" pulls in the opposite direction
- More sophisticated prediction models give similar results

## Problem II

- Would R&D growth be equal above and below the cap in the absence of the scheme? SkatteFUNN?
  - **Solution:** Remove largest and smallest R&D performers
    - ♦ Criterion : Max R&D  $>0$  and  $< 40$  mill
    - ♦ **But:** The more we do to solve problem II, the more we accentuate problem I
  - **Not perfect:** Cannot distinguish the effect of SkatteFUNN from other macroeconomic shocks that affect the treatment and control group differently

# Econometric framework

- Estimate the demand for R&D
- SkatteFUNN is considered a discrete change in the marginal price of R&D for firms below the cap
- Fixed effect regression analysis
  - “Do firms do more than they use to when ...”
- Varieties of the specification

$$\ln(FoU) = \alpha + \beta_1 SF_{it} + \beta_2 SF_{it} * Du4 + \gamma'x_{it} + \alpha_i + \varepsilon_{it}$$

- In sum, we interpret the empirical evidence to be consistent with the Norwegian R&D tax credit being effective in stimulating private R&D investments.
- Results are broadly consistent across data sources and the model specifications.
- However, one should bear in mind that the identification strategy is not fully “waterproof”, and that a causal interpretation of our findings relies itself on assumptions that are not innocuous

# Some thoughts on evaluation

- Tax incentives for R&D are comprehensive support measures – and the tax foregone can be considerable.
- This calls for evaluations that might map results, show how desirable outcomes are met and whether the tax incentive causes increases of R&D investments.
- When designing the scheme it would therefore be helpful to take into consideration that the scheme is going to be evaluated
- The operating agency could then put in place solutions for databases that are adequate and could be used together with national statistics.
- Adequate resources should also be put aside to ensure that the evaluations are of high quality. This might be costly – but the benefits for policy making could be significant.

# Some thoughts on evaluation (cont.)

- The timing of an evaluation should take into account a range of issues, such as:
  - how long it will take before the tax incentives is taken fully into use by businesses;
  - how long it will take for the expected policy effects to emerge;
  - how long it will take for the necessary data to become available; and
  - when the results are required for policy purposes (including any public commitments that might have been made)
- The risk of attempting an evaluation too early is that misleading results might be produced, for example showing no policy effect when it is simply too soon for effects to show up.
- On the other hand, however, it is important for an evaluation plan to recognise that results are needed to inform improvements to the policy and so, evaluations should also aim to produce some results as early as possible, without compromising on the former point

## Some thoughts on evaluation (cont.)

- Good data are crucial for evaluations. The lack of appropriate data severely limits the potential of an evaluation.
- Before starting the design of the evaluation project – and even before implementing the scheme, governments or operating agencies should go carefully through the data requirements for the evaluation and compare them to the data actually available.
- General principle for data issues: As far as possible, use already existing/available data to avoid extra red tape and to be able to have long time-series of data at a modest cost.

# Some thoughts on evaluation (cont.)

- Creating an evaluation database
- A proper evaluation has to rely on many different data sources,
  - Combining different types of information gathered with different methods.
  - Many analyses require information from different sources. The ability to link data from different sources is therefore crucial.
  - Consistent use of common identifiers across data sources is very helpful.
  - Using the idea of creating an integrated evaluation database will be helpful as a guiding principle in the process of gathering the data.
- The unit of analysis should be the entity that faces the incentive inherent in the scheme.
  - This will in most cases be the fiscal legal unit, implying that any data on plant or line-of- business level should be aggregated to this level.

# However...

- Even with the best of data and methods, the possibility of identifying causal effects depends fundamentally on the amount of exogenous variation in access to the scheme
- Certain knowledge of causal effects implies excluding some firms that would otherwise would have received support from the scheme
- This has short-term costs of several types, but may have long-term benefits in terms of better policies in the long run.
- Design of schemes, organization of evaluation and impact of findings also depend on policy makers and other agents' willingness to actually learn something (and not just to have their priors confirmed)