Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs,
Education and Research EAER

Swiss Science Council SSC

Note about SSC procedures and methodological choice in policy and programme evaluation

(SSC, in February 2018)

The SSC is strongly involved in policy and programme evaluation as part of its mandate. It has to respond to various needs of the Government in terms of programme evaluation and assessment in different kinds of contexts and at various time periods.

There is an increasing literature in economics on policy evaluation which aims at reaching the highest academic standards – that is to say, using formal methodologies which can allow evaluators to discern causality: if A, then B. Isolating the effect of a *treatment* is the gold standard in science and policy, and should be the obvious goal of any policy evaluation exercise. Other more descriptive or qualitative approaches, although useful for producing relevant information, cannot signify any causal association in terms of the change or evolution in the characteristics of a treated population.

Technically, the most sophisticated evaluations of public programmes – able to measure causality – are based on randomised trials, regression discontinuity or instrumental variables methods. The first of these approaches compares firms (or teams or individuals) that were randomly selected to receive support with those that did not. The second compares outcomes just above and below the threshold of qualification for public support. The third approach identifies the programme effect through an exogenous component of variation in qualification.

However, in many instances such techniques are not easily applicable or just not applicable at all. Let us take the example of a large-scale programme aiming at supporting the move of a whole scientific community towards a new research field. A rigorous evaluation of such a programme is undoubtedly very difficult for obvious measurement problems inherent in evaluating large-scale and long-standing programmes:

- There is no pure treatment effect (for such a programme, there are interrelated sequences and processes that make it difficult to link the programme to outcomes [such as increase in publications or patents] with any degree of precision).
- Large-scale programmes are affecting complex systems (an entire research community), in which the programme is an important element, but only one of many important elements.
- There are selection effects that are difficult to control (the research community that applies for large-scale programme funding appears to be better organised than average people. Moreover, successful applicants tend to receive multiple awards). Of course, awarding good communities makes sense; however, it does complicate evaluation.

However, the inability to undertake evaluations that are rigorous enough to satisfy the highest academic standards is not a good reason to stop governments as well as agencies from implementing certain evaluation routines – which are useful early-warning signals and can flag blatant programme failures. *A little bit of light is better than total darkness* and the kind of evaluation which is undertaken in the case of large-scale and complex programmes at least allows outcomes to be assessed against a particular benchmark – the baseline established by ex-ante expectations. This is what the Council is doing when it proceeds to evaluate large-scale and long-standing programmes. **While not claiming to discern causality, it contributes to the identification of strong achievements as well as significant underperformance relative to that benchmark, which would then call for discussions and debates with the programme beneficiaries and the policy makers.**