Research and Innovation Process Issues

1. Skew and serendipity

Both research and innovation processes are highly skewed. In research, it is axiomatic that a small minority of researchers produces the great bulk of the best research. In innovation, it is well understood that many projects will just about break even; many will fail and lose money; only a very small minority is likely to have huge success, hopefully large enough to compensate for the losses and still to generate a substantial surplus. Venture capitalist activity is tuned to coping with this kind of skew, but it applies equally to innovation programmes. For example, a long-range impact study of the Swedish “competence centres” found economic benefits in the range 5.3 to 11.8 billion Swedish Crowns – of which a single project accounted for 16-24% (Stern, Arnold, Carlberg, Fridholm, Rosemberg and Terrell, 2013).

This skew has to be considered when designing samples in impact analysis. A random sample of beneficiaries, for example, can easily miss the one or two projects that make the big money, leading the evaluator radically to misjudge the performance of the scheme. The skew means that some schemes – especially small ones – may produce poor returns through bad luck rather than bad judgement. The normal portfolio principle applies: namely, that larger portfolios are more likely to capture high return cases and are less vulnerable to individual failures than small ones.

Industry has evolved techniques for coping with skew, notably by weeding out early cases of incipient failure. Thus industrial R&D projects tend to involve progress reviews or “stage gates” where go/no go decisions are made about project continuation. This allows at the least frequent assessment of technical feasibility and lets management bring new information about markets or firm strategy to bear. Consequently, the rate of project cancellation can be quite high. Public funding schemes tend to work with a different logic, focusing on using a high-quality process for identifying potentially successful projects and then, once they are funded, monitoring rather than managing them. Often, there is a reluctance to withdraw state funding because this would imply that a “bad” funding decision had been made at the outset. However, state-supported innovation projects should in principle be “further from market” than companies’ internally-funded ones and therefore more difficult to assess using a close-to-market stage-gate process. State support for innovation often also has goals that go beyond the success of individual projects, aiming to support network formation, the development of absorptive capacity and so on, in which case it should be measured in these terms and not solely in terms of project success.

One way to cope with skew in evaluation and impact analysis is to search deliberately for the successes and then to sample the remainder of the programme – the approach taken in recent years to the evaluation of the EUREKA programme (Georghiou, 1999). This requires a two-step method, in which the first gathers intelligence about successes and allocates them to a separate sample from the other projects in the programme.

2. Effectiveness comparisons and the substitutability of interventions

Impact assessments tend to be done as individual studies using tailored methodologies. These factors alone limit their comparability. However, other issues mean that trying to standardise methods would have a limited effect in making them more comparable.

Qualitative studies, however useful, are hard to compare. But all studies suffer from limits to the impacts that can be tackled – both because the choice of methods means that some impacts are in focus while others are not, and because the effectiveness of individual methods at detecting impacts varies from study to study. For example, many innovation policy impact assessments derive their impact data from information generated by the beneficiaries, often through surveys or telephone interviews. Such methods are poorly adapted to assessing spillover benefits to non-beneficiaries.

Typically, in surveys, only a proportion of beneficiaries claim that there have been economic benefits, and in turn only a sub-set of these are willing to able to provide estimates of benefits. Clearly, the impact assessment only captures a proportion of the benefits – and the percentage covered is inherently unknown. Other methods to varying degrees suffer the same difficulty. The upshot is that comparing impact assessments involves comparing studies whose coverage of impacts is unknown but likely to be different.

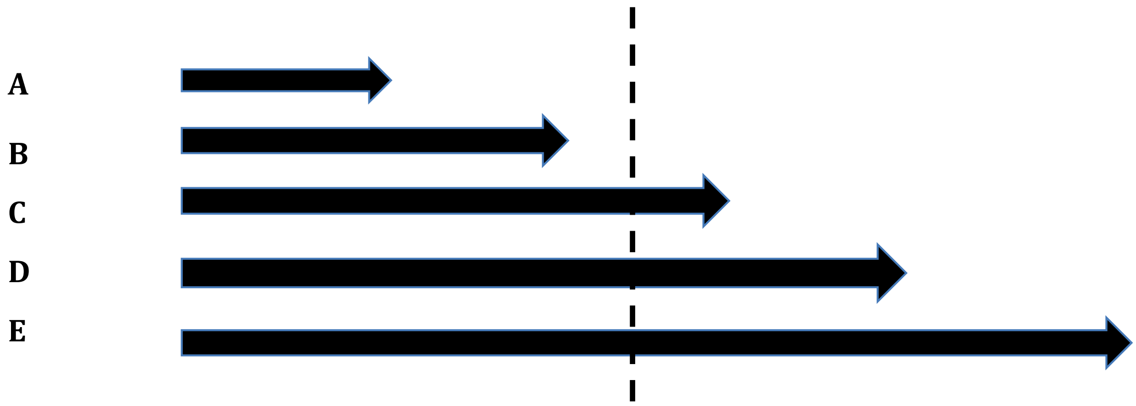
The impacts of individual policies and instruments are also likely to depend upon the context. For example, many countries run 'innovation voucher' schemes, providing small subsidies to SMEs wanting to gain experience with the use of services from research institutes or testing organisations. In countries where few firms have such experience, it would be reasonable to expect the effects on company performance to be greater than in more mature situations, where relations between SMEs and institutes are better established and vouchers are being used by already-experienced or less capable firms.

For good reasons, states aim to limit the cost of intervention – normally with the Finance Ministry central in imposing cost discipline. It is therefore tempting to look for the interventions that have the highest benefit-cost ratios and to focus funding only on those. The difficulties of comparing impact assessments discussed in this and the next sections suggest this might not be wise. A systemic perspective also suggests that this kind of optimisation would be unfortunate: interventions are not necessarily substitutable. Since the innovation systems heuristic (and, indeed, common sense) implies that good economic performance is connected to an extensive combination of intra- and inter-organisational capabilities and flows, the policymaker needs to monitor the health of the system quite expensively and intervene where necessary to correct systemic faults. Some individual interventions may appear to have comparatively low returns while still being necessary to ensure the overall performance of the system. Optimising based only on economic returns would be similar to deciding that the way to make a car go faster is to install a bigger engine and throw away the wheels.

3. Timing impact assessment

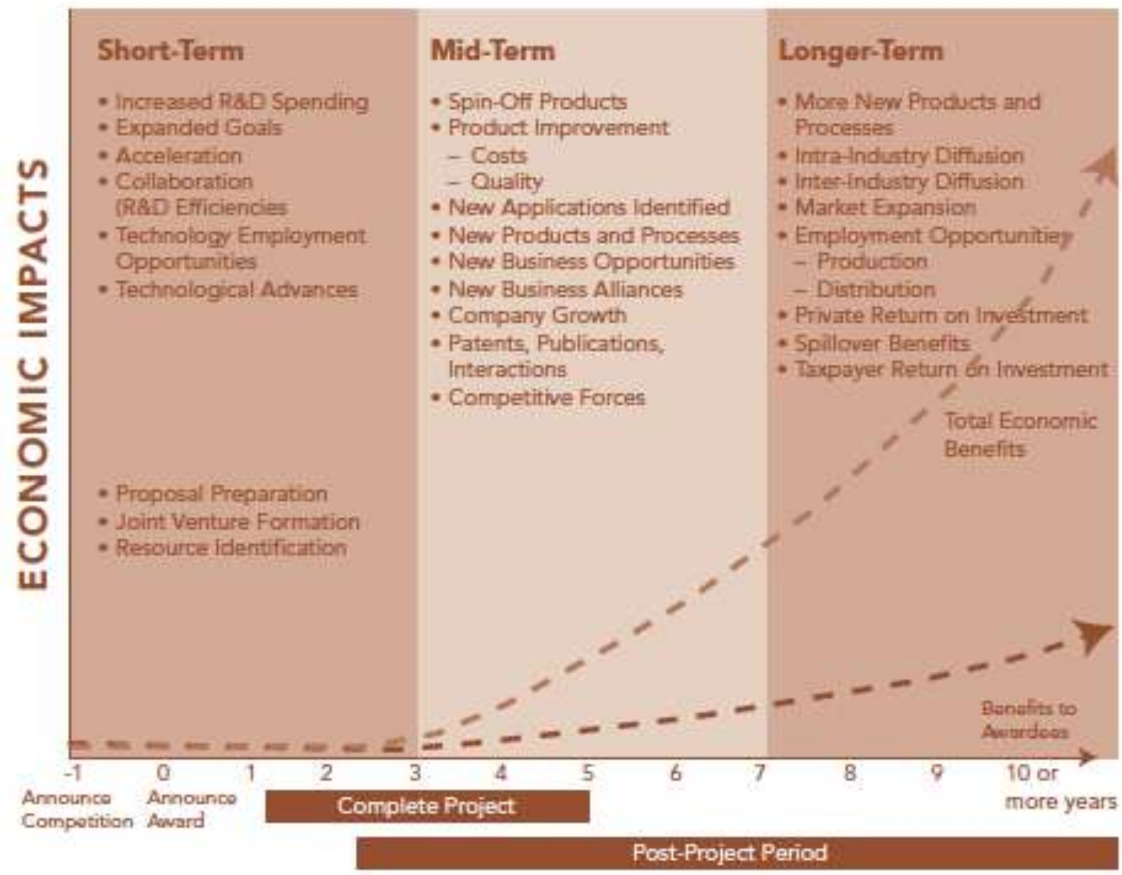
A well-known problem in evaluation is that, in order to make timely policy and especially to secure needed budgets, evaluation results are often needed before the relevant intervention has had any impacts. The question **when** to do an impact assessment is no less problematic and becomes acute if there is a desire to compare the impacts of alternative policies or instruments. In many cases, the time at which an intervention "matures" and produces impact is varies – and may even be unknown. Making a comparative assessment of impacts in the cases shown in Figure 1. at the point shown by the dashed line would clearly present a more positive picture for short-term interventions than for others. On the other hand, overly delaying impact assessment for instruments A and B risks loss of data as they disappear into history.

Figure 1. Timing of impacts and impact assessment



There **are** some systematic difference in how long it takes between impulse and impact. Figure 11 illustrates the expected timescales for the large US Advanced Technology Programme (ATP), which aimed to bring pre-commercial technologies to the stage of industrial use. However, there is also evidence that the time to impact for some research is very long indeed.

Figure 2. Time lags, impacts and proxies for the ATP programme



Source: (Ruegg & Feller, 2003)

DoD's Project Hindsight (Iserson, 1967) looked at twenty weapons systems and identified "RXD (research or exploratory development) events" that contributed to them. Results from undirected basic research typically took about 20 years to implementation while those from directed basic research took about half as long. NSF's TRACES project (Loellbach, 1968) studied "critical events" leading to five innovations: magnetic ferrites; the videotape recorder; oral contraceptives; the electron microscope; and matrix isolation[1]. Applied research events were important in the twenty or so years prior to innovation while basic research events stretched up to 50 years back, with a peak in the interval 20-30 years before innovation. A recent meta-analysis by VINNOVA of the fourteen long term impact studies undertaken so far (Elg and Håkansson, 2011) found that programmes need to be flexible and "patient": long programmes have greater effects on beneficiaries' strategies and learning than short ones. Often, 10-20 years elapse before socio-economic effects of any size are visible.

This document is based on: **OECD Directorate for Science, Technology and Innovation (2014), "Assessing the Impact of State Interventions in Research – Techniques, Issues and Solutions", unpublished manuscript.**

[1] A technique for trapping small amounts of a chemical within an unreactive matrix, easing analysis, for example using spectroscopy.

References

Elg, L. and S. Håkansson, (2011). När Staten Spelat Roll: Lärdomar av VINNOVAs Effekstudier. VINNOVA Analys, VA 2011:10. Stockholm: VINNOVA.

Georghiou, L. (1999, 24 1). Socio-economic Effects of Collaborative R&D—European Experiences. Journal of Technology Transfer , 69-79.

Iserson, R. S. (1967). Project Hindsight (Final Report). Office of the Director of Defense Research and Engineering. Washington DC: Department of Defense.

Loellbach, H. (1968). Technology in retrospect and critical events in sciences (TRACES). Vol. I. Illinois Institute of Technology Research, Contract NSF-C535. Chicago: Illinois Institute of Technology.

Ruegg, R. and I. Feller, (2003). A Toolkit for Evaluating Public R&D Investment Models, Methods, and Findings from ATP's First Decade. Washington DC: US Department of Commerce.

Stern, P., E. Arnold, M. Carlberg, T. Fridholm, C. Rosemberg and M. Terrell, (2013). Long Term Industrial Impacts of the Swedish Competence Centres. Vinnova Analysis, VA2013:10. Stockholm: Vinnova.