technopolis |group|

Some reflections on the developments in the evaluation of research and innovation

Paul Simmonds
Technopolis Group

About me ...

- I have spent much of the last 25 years with Technopolis evaluating research and innovation programmes, in UK / Ireland / EU
  - *We do so as a private contractor, working on commercial terms for government ministries and STI funding agencies*
  - *Technopolis does not apply for research and innovation grants, as a part of our commitment to independence*
    - In terms of my background, I began as a marine engineer before going on to study economic geography and ultimately ended up working as a researcher at SPRU
- In that time, from one perspective, everything has changed and then again, from another perspective we continue to battle the twin methodological challenges of time lags and attribution
- The STI evaluator's toolkit remains a work-in-progress

technopolis |group|

Changes in focus …

- The focus of evaluations (our questions) has shifted
  - *In very simple terms, we have moved from a focus on process and outputs to OUTCOMES and IMPACTS*
  - *Relevance / quality / efficiency are still in the mix, but are very much hygiene factors*
  - *Monitoring systems have improved hugely and remove the need for much of the data collection we would have once invested in heavily*
  - *Effectiveness is the key question (outcomes): has the programme delivered the quantum of wider benefits promised, or would the money have been better spent elsewhere (or not spent at all)*
  - *BUT increasingly clients push beyond measuring outcomes among the supported populations and insist evaluators quantify the nature and extent of wider social and economic IMPACTs*
  - *This is deeply challenging …*

technopolis |group|

Changing evaluation methods

- For STI, theory-based evaluation (TBE) and a mixed methods approach continues to be the mainstay of our working methods
- BUT, the desire for quantification of NET EFFECTS has increased the use of quasi-experimental methods, typically within a TBE framework
  - *Evaluation in the innovation space has been criticised by government economists and the wider policy evaluation community for its poor methodological robustness (as compared with the approaches one finds in health, education, employment, etc.)*
- It's a tough ask
  - *Counterfactual analysis using econometrics and economic modelling is more tractable / credible in large-scale innovation support programmes*
  - *Quasi experimental methods don't cope well with the diffuse and long-term effects of applied research*
  - *Getting to grips with causality and the quantum of all benefits is complex and costly*
  - *STI evaluations struggle to get past level 1 on the Maryland Scientific Methods Scale (e.g. before and after comparisons; simple comparisons with 'approximate' control groups)*
  - *For research impacts, it seems the best approach remains qualitative: backwards tracking, longitudinal impact case studies within a TBE environment*

Changes in the evaluation process …

- Notwithstanding the increased difficulty of the evaluations …
  - *Governments are evaluating a larger proportion of policies and programmes and insist on more consistent and robust methods*
  - *BUT, faster … (6 months rather than 12) and Cheaper … prices have halved in cash terms over 20 years (-75%)*
- Economic consultancies have changed the rules of the game
  - *Econometrics, economic data and economic models provide an edge*
  - *Standardisation of data collection tools*
- Framework contracts are also becoming the norm
  - *Favour larger contractors with tighter management systems / QA*
  - *Thematic experts increasingly brought in to provide orientation and add a little colour, but in a very limited capacity*
- Governments have strengthened their internal evaluation capacity, and methodologists work closely with policy teams

technopolis |group|

Changes in use and useability …

- Evaluations are sometimes giving way to Impact Assessment,
  - *The focus is on identifying and counting benefits rather than worrying overly about effectiveness or efficiency*
  - *Primary purpose: gather sufficient evidence to impress finance ministries; protect or increase budgets*
- In other areas, clients want the best of all worlds
  - *Impact assessment for finance ministries AND*
  - *Ex post evaluations of Effectiveness and efficiency AND*
  - *Formative evaluations to inform organisational learning*
- Effects in practice …
  - *Mixed picture*
    - Evaluations do underpin incremental learning for institutions
    - Less so for programmes which are rather fluid in conception and often change radically from one generation to the next, driven more by political imperatives
    - Some countries and clients are more disciplined than others in transparency around institutional learning and programmatic changes, which provides a powerful feedback loop and incentive mechanism

technopolis |group|

Future developments …

- National guidelines evolving into international norms
- More widespread use of evaluation frameworks created as part of the programme design and used to set up monitoring systems (KPIs), baselines and evaluations
- Greater use of ex ante impact assessments (test policy proposals more robustly)
- Greater use of portfolio evaluation or fitness checks (coherence / complementarity of multiple initiatives addressing a single set of policy objectives; grand challenges)
- Greater cooperation between commissioning organisations and evaluators, moving beyond the idea of evaluation being independent / objective as a function of structural difference
- Quasi-experimental methods will likely increase, but need better data on wider business populations to support matching (e.g. BERD, public supports, innovativeness, trajectories), and use of panels to allow reviews at several points across the STI lifecycle (to cope with the time lag issues). We must do better at linking micro and macro economic perspectives …
- Data science techniques hold out the promise of improving evaluator's ability to cost-effectively work across multiple data sources (inc. unstructured) but not the silver bullet many are anticipating …