

Getting the most of evaluation in the age of RCTs and big data: an EU perspective

Bradford Rohmer
Kristiansand, 21 September
2018



Agenda

1. What I'm going to talk about
2. Background
3. Illustrative examples
4. Concluding observations
5. Where to go from here: questions and discussion

What I'm going to talk about

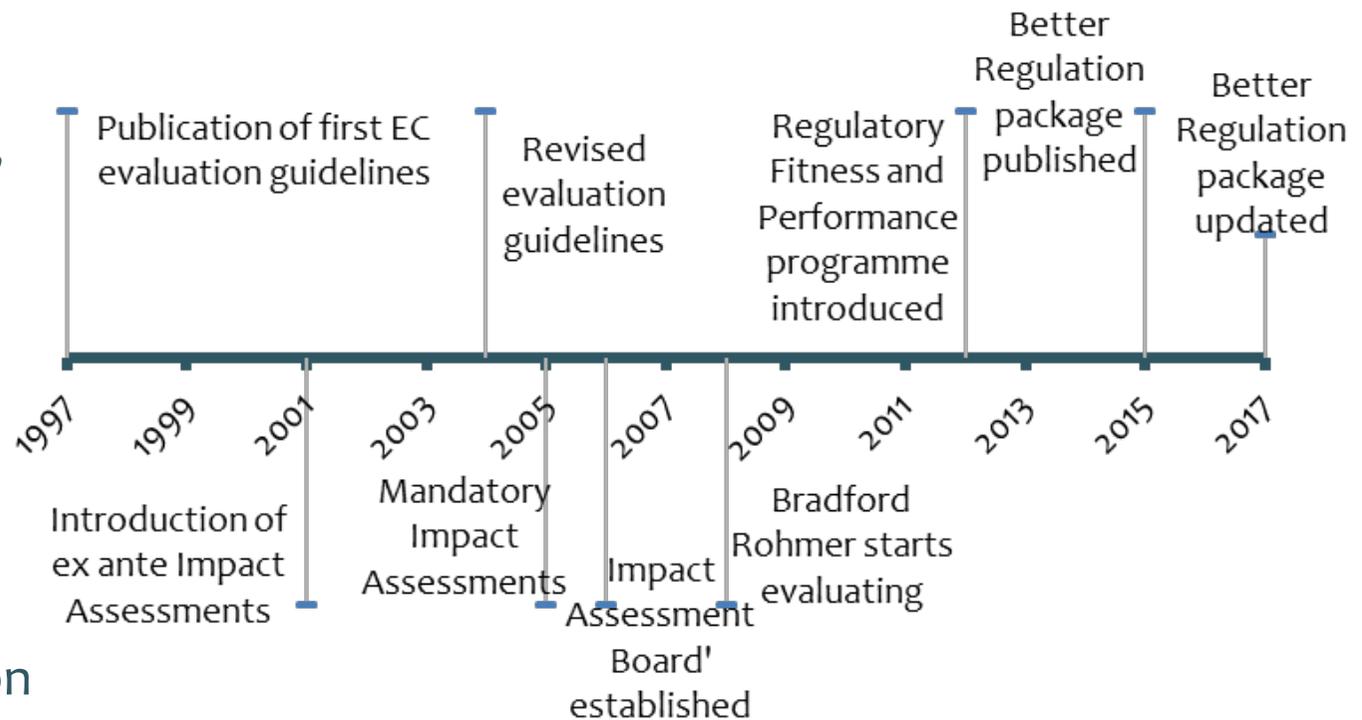
Advances in evaluation thinking and what it means in the EU context:

- Institutional and governance trends
- Putting new methodologies into practice
- Good and bad practices
- Practicalities for us “practitioners” / consultant mercenaries

Background – evaluation in the EC

Three ‘eras’ of evaluation in the European Commission:

- **From 1997:** evaluation mainly as obligation
- **2000s:** central guidance, introduction of ex ante impact assessments, going beyond mere accountability
- **2015 onwards:** Better Regulation package aimed at standardising and improving evaluation practice and governance



Background – evaluation advances

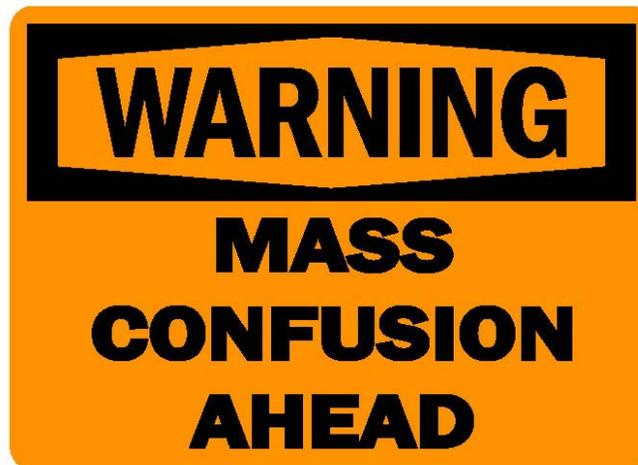
Realist Evaluation

Regression discontinuity analysis

Theory of change

General Elimination Methodology

Contribution Analysis



Big data

Multi-criteria analysis

Randomised control trials

Computable general equilibrium modelling

Difference-in-difference

Outcome Mapping

Propensity score matching

Qualitative Comparative Analysis

Illustrative examples

- The next slides focus on examples of what these trends have meant in practice
- A lot of this depends on who commissions an evaluation, and for what purpose
- To help show this I've defined three 'types' of evaluation commissioners
- No one really fits 100% in a specific 'type'; indeed most people have elements of all three



The nerds



The politicians



The bureaucrats

The nerds

Key features:

- Convinced about evaluation and its learning potential
- Keen to use exciting (especially quantitative) evaluation methods
- Often far from policy- and decision-making

Pros:

- Willing to think outside the box and try new things
- Supportive of daring consultants

Cons:

- Overly optimistic view of possibilities
- Difficulty linking evaluation to policy needs
- Can misdirect energy that would be more usefully applied to 'traditional' methods



The nerds

The politicos

Key features:

- Influential at the senior level
- Direct interest in the results of an evaluation
- Little experience of evaluation/social research

Pros:

- Very engaged throughout evaluation process
- Can help open doors, ensure consultants get needed data, etc.
- Ensures evaluation results are really used

Cons:

- Difficult to engage on methodological aspects
- Care more about the outcome than impartiality or methodological robustness
- Can be prone to manipulating evaluation governance processes or findings



The politicos

The bureaucrats

Key features:

- Focused on processes, rules, hierarchy
- Often have long experience with evaluation and / or the commissioning authority

Pros:

- Commitment to ensuring all procedures and guidelines are followed
- Can help secure evaluation independence

Cons:

- Can be too rigid and impractical, especially given the iterative nature of social research
- Sometimes unconcerned with how the evaluation fits into the big picture



The bureaucrats

In other words...

“Have you considered regression discontinuity analysis to reverse engineer control and treatment groups?”



“The success of our programme isn’t clear enough – send me a revised version by tonight or we terminate the contract!”

“My hierarchy has instructed me to make sure you include 8 case studies in the evaluation.”

Concluding observations

- New and advanced methodologies really add value, but they rely on certain conditions, early planning and the right monitoring / baseline data
- Guidelines and standardisation help, but these don't automatically fix all structural problems
- It takes a long time to build a real evaluation culture, but 'champions' in high places can help
- This puts a lot of pressure on the needs, motivations and capacity of individual clients / commissioners of evaluations
- Competitive bidding procedures are in many ways good and probably necessary, but they also create perverse incentives for practitioners, and make it hard for each side to learn from each other
- The upshot is that consultants point out lots of problems (cf. this presentation) but have a hard time doing anything to fix them

Where to go from here?

- Do the issues I described sound familiar to the situation in Norway?
- What ways of dealing with them have worked well / less well?
- What is the experience of impact evaluation and novel / advanced methodologies in Norway?
- What support could evaluation practitioners give to commissioners of evaluation to improve the situation?



Thank you

Bradford Rohmer

bradford.rohmer@oxfordresearch.se