

ENVIEVAL

Grant Agreement Number 312071



Strengths and weaknesses of Counterfactual Evaluation

Jyrki Aakkula, Janne Artell & Heini Toikkanen
MTT Agrifood Research Finland

04.July .2013, INEA, ROME










Contents



- 1) Basic concept of counterfactual evaluation
- 2) Common Monitoring and Evaluation Framework (CMEF) and counterfactuals
- 3) Observations from the review of RDP evaluation reports
- 4) Approaches to counterfactual evaluation and examples of advanced methods
- 5) Lessons learnt and essential issues for future counterfactual evaluation of environmental impacts of RDPs

© MTT Agrifood Research Finland

16.7.2013

2

1) Basic concept of counterfactual evaluation



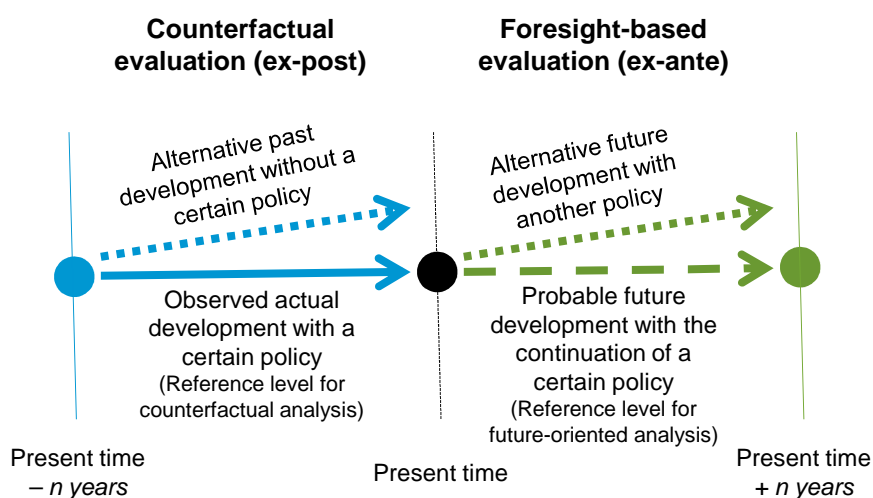
Counterfactual approach

- The 'counterfactual' measures what would have happened to beneficiaries in the absence of a certain intervention (policy instrument), and the impact is estimated by comparing possible alternative (i.e. counterfactual) outcomes to those observed under the intervention (policy instrument)
- Observed actual development is a reference level, where alternative possible developments are compared to
- In other words, counterfactual evaluation provides scenarios depicting possible alternative pasts

1) Basic concept of counterfactual evaluation




Evaluation and time horizon



2) Common Monitoring and Evaluation Framework (CMEF) and counterfactuals

Common Monitoring and Evaluation Framework (CMEF) and counterfactuals



- “In a first step, impact should be estimated at the level of direct and indirect beneficiaries by programme evaluator on the basis of output and result indicators, survey data, experience and evaluations from previous programming periods (for calculation of double counting, deadweight, displacement and multiplier effects). This should be cross-checked against the **counterfactual** situation and contextual trends in programme area.” (CMEF Guidance document, p. 14)
- “As evaluation looks at change over time, the establishment of the **counterfactual** is a central issue for all evaluations. In this context the “base-line indicators”, established by the SWOT analysis and ex-ante evaluation at the time of programming, need to be mentioned. The base-line indicators are an important reference point for the evaluation of impacts of single measures and programmes as a whole.” (CMEF Guidance note B – Evaluation guidelines, p. 5)
- “Process and synthesise available data and information, and - where necessary – handle data gaps by modelling or other extrapolations. Apply a measurement against the **counterfactual** as well as target levels.” (CMEF Guidance note B – Evaluation guidelines, p. 11)

© MTT Agrifood Research Finland 16.7.2013 5

3) Observations from the review of RDP evaluation reports

Counterfactuals in RDP studies




Type of public good addressed	Type of counterfactual approach applied				TOTAL
	Comparison of participants and non-participants	Comparison of participants before and after	Quantitative modelling	Not specified	
Climate	4			2	6
Water quality	12		4	3	19
Biodiversity wildlife	8			5	13
Biodiversity HNV	1			4	5
Soil	6		4		10
Landscape	5				5
Animal welfare	12	4			16
TOTAL	48	4	8	14	74

- Counterfactual approach mentioned quite often, methodology uncertain
- Used data sources vary and are vaguely reported
 - FADN & IACS data most used
- Impact indicators often insufficiently defined

© MTT Agrifood Research Finland 16.7.2013 6


4) Approaches to counterfactual evaluation and examples of advanced methods



Counterfactual approaches

		Comparison of participants and non-participants during the programme	Comparison of participants (and non-participants) before and after the programme
Quantitative analysis	Statistical / econometric methods	Most preferred way of doing counterfactual analysis (exemplary methods: propensity score matching, regression discontinuity design)	Most preferred in cases where a measure/programme is all-inclusive
	Quantitative modelling	Applicability depends on model characteristics and behavioural assumptions	Most applicable in ex-ante analysis
Qualitative analysis	Expert / stakeholder views	Provide useful additional information on likely impacts but in a form difficult to quantify	Provide useful additional information on likely impacts but in a form difficult to quantify
	Qualitative modelling	Able to indicate likely impacts, not necessarily able to confirm them	Able to indicate likely impacts, not necessarily able to confirm them

4) Approaches to counterfactual evaluation and examples of advanced methods



Propensity Score Matching (PSM)

- PSM uses information from a group of eligible participants that do not participate in a programme to identify what would have happened to those who actually participate in the absence of the programme
- By comparing how outcomes differ for participants relative to observationally similar non-participants, it is possible to estimate the impact of the programme
- PSM is useful especially in cases where there is enough information on all the characteristics that affect eligible participants' decision to join or not to join the programme
- Correspondingly, PSM is less useful in cases where it is likely that the decision to participate or not to participate depends also on such characteristics of which there is not enough information available

© MTT Agrifood Research Finland 16.7.2013 8

4) Approaches to counterfactual evaluation and examples of advanced methods



Regression Discontinuity Design (RD)

- Regression discontinuity (RD) method exploits information on eligibility rules related to programme participation
- Eligibility rules provide a distinct cut-off point below which participation is not possible
- Comparing participants (i.e. those who are eligible) and non-participants (i.e. those who are not eligible) in the same neighbourhood, i.e. very close to the cut-off point above and below, provides a way to take into account observed as well as unobserved characteristics which affect the programme outcome
- The RD approach is vulnerable to bias
 - if the eligibility rules bend in practice or if they change over the evaluation period
 - if the number of observations is limited in the neighbourhood of the cut-off point
- The DR approach cannot be applied if programme participation is open to all without any limiting eligibility rules

5) Lessons learnt and essential issues for future counterfactual evaluation



Lessons learned (1)

- Design of counterfactuals should always be an integral part of the RDP planning process
 - Ex-ante evaluation should already pay attention to how well the requirements of counterfactual assessment are taken into account
 - High quality ex-post counterfactual evaluation is possible only if ex-ante evaluation has addressed counterfactual issues in a sufficient way
 - Especially the use of advanced statistical/econometric counterfactual methods (or any other counterfactual approach) becomes difficult if their data and methodological requirements are not taken into account in the early phase of RDP planning
 - Counterfactual thinking should cover all phases of the evaluation process: structuring, observing, analysing and judging
 - Selection of indicators is crucial, emphasis should be on result and impact indicators
 - Because RDP evaluations are not necessarily well-resourced, the emphasis has been on input and output indicators for which data is more readily available

5) Lessons learnt and essential issues for future counterfactual evaluation



Lessons learned (2)

- Aim should always be quantitative assessment of impacts when possible
 - Choice of a counterfactual evaluation method is in a decisive role, priority should be given to statistical and econometric methods because they can more reliably and transparently provide quantitative estimates for impacts
 - However, qualitative counterfactual methods are useful especially in the early phase of RDP design
 - Elimination and analysis of biases in comparison settings improve considerably the quality of evaluation results
 - Nature of an environmental impact (public good) to be evaluated affects considerably applicable methods and methodology
 - Independent of a counterfactual evaluation approach used, the causality between policy measures and environmental impacts is hard to identify reliably when multiple intervening factors affect simultaneously

5) Lessons learnt and essential issues for future counterfactual evaluation



Major challenges of (counterfactual) evaluation of RDPs (1)

- In many cases monitoring and data requirements (especially data about result and impact indicators and characteristics of participants) of the counterfactual approach are excessive which limits the use of advanced statistical/econometric techniques
 - How to promote data collection which matches to the needs of advanced statistical/econometric techniques that are preferable to use in counterfactual evaluation?
- Identification of all the causal relationships how various RDP measures affect the rural environment improves the quality of counterfactual and other evaluation
 - How to develop the identification and determination of relevant causal relationships?
- RDPs are only one set of applied CAP measures, and these other CAP measures (as well as other EU level policies) may also have an impact on the quality of the rural environment
 - How to differentiate between impacts of RDP measures and impacts of other policy measures?



Major challenges of (counterfactual) evaluation of RDPs (2)

- Sometimes even an advanced counterfactual analysis is not enough. For policy-makers knowing impacts is a necessary but not a sufficient condition: without some kind of commensuration across different impacts it is not possible to rank policy outcomes or alternatives which involve trade-offs
 - How to further develop the output of counterfactual evaluation to better match the needs and expectations of policy-makers?
- It seems that despite of the CMEF guidelines counterfactual approach is occasionally completely neglected or carried out vaguely
 - How to promote more rigorous execution of counterfactual approach in RDP evaluations?