

**Summary report on the 2020 Meeting of the Community of Practice
on Counterfactual Impact Evaluation of ESF interventions
(CoP-CIE-ESF)**

10&12 November 2020

This year the Community of Practice was devoted to discussing all phases of the policy cycle, focusing on the real experiences of Managing Authorities (MAs) in the past 7 years in carrying out CIE of ESF funded interventions. The choice of the topic represented a constructive way of ending the current programming period and preparing for the new one.

In particular, the meeting was structured as an online open discussion between the MAs, first in small groups, and then in a plenary section, based on their work in the field on evaluation. The discussion focused on the points that they identified as best and worse experiences of the past seven years in carrying out CIE. Small groups discussions were very fruitful, and touched upon several topics which can be summarized in five broader areas: (i) selection of interventions and methods; (ii) data; (iii) evaluation process; (iv) contractors; and (v) communication of results and engagement of policy makers. In the sections below, we report the main points of discussion touched during the CoP. The aim of this document is to keep track of the current challenges, on how some Members States have solved some of them and on the steps that still need to be done in the future.

1. Selection of interventions and methods

MAs often have the need to evaluate set of actions, i.e. entire thematic objectives or operational programmes, rather than single operations. CIE methods are more tailored for evaluations of specific interventions with clear target populations, selection criteria, and expected outcomes.

- CIE methods indeed are not able to deliver on evaluations of heterogeneous measures, which might have different intended outcomes, selection criteria etc. For this reason, evaluations using CIE cannot be thought as a one-size-fits-all solutions, rather as a way to “zoom-in” and better quantify impacts of specific lines of intervention.

The choice of the intervention to evaluate using CIE is not always clear-cut. This is especially relevant when the amount of resources to carry out evaluations are limited, since CIE evaluations tend to be data greedy, and time consuming. The choice is generally driven by considerations in terms of data availability, policy relevance, and feasibility of the CIE approach. On the latter, some MAs might not have the internal capacity to judge the degree of feasibility of a CIE evaluation and external support might be needed.

- The support offered by CRIE with the QAS services also covers these points. Discussing with CRIE the different interventions ongoing, can help identifying the more suitable for a CIE evaluation.

CIE methods are not able to “tell us” why an intervention works or does not work. On top of this, the specific set of selection criteria and eligible populations make it difficult to generalize the evaluation results.

- CIE could be complemented with other evaluation methods which are better tailored to answer this set of questions
- However, it is often the case that some intervention are more effective on sub-groups of participants. Identifying on which of these sub-groups the effects of the policy under consideration is higher/lower can definitely be done using CIE methods. The characterization of these sub-groups in terms of observable characteristics (e.g. age, duration of unemployment, sex, region etc.) could be an ideal starting point to assess “where” the intervention works.

ESF-funded operations often overlap, fully or partially, with already existing policies designed by MS, and financed with national resources. If a comparison group has to be identified, say for a training programme, then the question becomes “how to identify control units who do not participate, since there are already existing national training programmes that are very similar in terms of content”

- Even in those cases in which the only difference between the two programmes is the funding source, there is always a sub-population of non-beneficiaries. The evaluation process is not different from the standard case: identifying non-beneficiaries (of both programmes) using (a) good quality data; (b) selection rules of the intervention. In the case in which ESF-funded operations differ from existing programmes, the focus of the evaluation could be to assess the effectiveness of the additional/different elements characterizing the two programmes, having in mind the programme-specific assignment rules. Put differently: treat the two programmes separately, and evaluate their relative effectiveness using different sets of control units.

2. Data

Having access to the right data is fundamental to carry out a good CIE evaluation. This implies having data not only on the beneficiaries of the intervention, but also on a group of non-treated units, which should be used as control. In addition, there might be the need to use longitudinal data, which allow to measure a baseline (before the intervention take place) and the proper outcome (sometime after the intervention has been completed). Moreover, other characteristics which are likely to influence both participation and outcomes, the so-called “control variable”, are often fundamental in the evaluation process. The use of administrative or registry data, if

available, has been suggested as the best practice for finding the right data for CIE evaluation. However, access and use of administrative data is not that simple. Two main problems were identified: how to exchange and get access to administrative data from the data owner, and how then share this data with the external contractors in charge of the evaluation.

Exchanging data between institutional partners

The exchange of data between national partners is frequently an issue, especially when data on beneficiaries have to be matched with data on the target population. The problem is quite often related to data protection rules (interpretation of GDPR rules by DPOs). There is also a lack of understanding of the type of data required to perform some evaluations, i.e. why are individual microdata needed? Aren't aggregate data sufficient?

The different Member States proposed some solutions they adopted to overcome the data issue:

- i) Ad ad-hoc memorandum of understanding to overcome the obstacles generated by the different set of rules set out by different DPOs. The major challenge was the necessity to pre-specify a set of characteristics to identify ex-ante the pool of statistical units in the database of the institutional partner.
- ii) The exchange can be based on a gentlemen agreement, but some institutions refuse to collaborate because of the lack of a legal basis and mandate. A drawback of this option is that the "partner" might change and the process needs to start from the beginning.
- iii) Regulatory solution. The exchange of data for evaluation purposes between different institutions is regulated by a State act which provides the legal basis for data exchange. The process was lengthy, and will need to be renewed with the new programming period.
- iv) Involvement of National Statistics Institute. They have the know-how, the legal mandate, and the institutional credibility to access and merge microdata from different sources and to provide the MA with pseudo-anonymized data for evaluations.
- v) The MA has the legal mandate to use the microdata, set out in national legislation. They collaborate closely with the Tax Authority who share relevant data with them. GDPR is respected, since data are used for evaluation purposes and public benefit. They have internal guidelines within the Ministry to work on data. External evaluators received anonymized data, and they have the obligation to use it only for the evaluation.

In addition, in the **next Common Provision Regulation (CPR)** there will introduce a stronger legal basis for the use of administrative data for monitoring and evaluation purposes, so that there should be no longer need of a specific act to access data as long as it is specified that these are used for evaluation purposes.

A further technical issue with data is the difficulty in matching different data sources from different partners potentially due to non-harmonization of data collection practices by different stakeholders.

A final remark on data is that planning data collection and start gathering info on which existing data could be used is key.

Exchanging data with external contractors

Even when access to administrative data is possible, other complications arise when the MA has no internal capacity to carry out the evaluation and data need to be transferred to external contractors. It is been mentioned that institutional partners, holding the data, are reluctant to exchange data to third parties.

Some Members States have proposed the following solutions:

- i) Give anonymized data to external contractors.
- ii) Support of legal advisors

3. Evaluation process

A major concern is that **the evaluation units are often not involved in the design phase of interventions**. This is likely to have consequences on the quality of evaluations for different reasons:

- 1) The design of the policy or its implementation might make the policy not evaluable using CIE. Some MA are suggesting to plan CIE where possible: “Plan the intervention, so that it is possible to evaluate it!”.
- 2) The evaluation units are not fully aware of all the technical details of the implementation phase (the knowledge rests within the beneficiaries), and this lack of information might impact the choice of the appropriate methods, or make the use of CIE not feasible

However, even if the evaluation units can’t participate to the design phase, something could be done to ensure that evaluation is still feasible:

- i) A dialogue between policy makers and evaluators in all the phases of the evaluation process.
- ii) Engaging stakeholders with internal capacity on evaluation: finding project organizations working on implementation with expertise in evaluation is important.
- iii) Make key stakeholders (e.g. the PES) a strategic partner of the evaluation. One Member State reported the experience of a fruitful collaboration between the MA and the PES made possible by the common objective of a rigorous evaluation based on CIE analysis

4.a Contractors

MAAs have highlighted several problems in finding good contractors, able to perform CIE.

Poor knowledge of CIE among contractors. There is not enough capacity to carry out CIE on the market. Evaluators tend to apply simplified methods (not CIE) and it is difficult afterwards to follow and understand how some conclusions were reached.

Engaging expert in CIE is hard. Some limitations are due to public procurement rules, as MAs are not free to choose the contractors they would like to.

Some of the solutions adopted are the following:

- i) Specifying which interventions should be analyzed using CIE in the ToR, explaining why these interventions are chosen.
- ii) Common market for contractors: organizing common capacity building events to inform potential contractors about MA's expectations in terms of CIE among neighboring countries that share similar experiences. The event is intended to raise awareness among evaluators about EC evaluation criteria (effectiveness, efficiency,..).
- iii) Collaboration with CRIE was started to increase knowledge of CIE among contractors. CRIE has been helping in two ways: a) as a "supervisor" of the work carried out by the contractors related to CIE (i.e. providing feedback throughout the evaluation process); b) as a "co-evaluator" working together with contractors. In this latter case CRIE was responsible for the CIE evaluation part only, and the contractor for the remaining parts of the evaluation. This helped capacity building among contractors.
- iv) If contractors are not able to bring their own expertise, contact with academia can help.

Interactions with contractors after selection can sometime be problematic too. Sometimes MAs are called to both judge the work carried out by contractors and work with them to carry out the evaluation. It can be hard to balance the two roles of evaluators and judge in the collaboration with contractors

In addition, it can be hard to follow the actual evaluation work after the contractor has been selected. Involving academia and other experts throughout the evaluation process can help. CRIE can also provide this type of support within the QAS service.

4.b The role of Academia

One possible solution to the lack of CIE knowledge among external evaluators is to consider academic institutions or researchers as potential contractors. Researchers should know the proper methodologies and should be able to carry out good CIE evaluation.

However, some MAs pointed that also engaging researchers from the academia has some negative sides:

- Some countries have highlighted that the service provided by experts in academia is not of high quality, and they don't deliver with high academic standards.

- Even in the case of specific CIE requirements in the ToR, in some Member States, contractors from academia always carry out the same type of analysis, and do not propose alternative or more valid solutions.
- When it comes to outsourcing, notwithstanding the existence of potential good quality academic partners, they barely submit an application for tender.
- In public procurement they normally hire contractor that outsource the work to experts in academia, who outsource to their PhD students (double outsourcing), which can be a costly and long process.
 - As an alternative, in one MS it was decided to have ad-hoc teams to hire PhD students directly since they are the ones doing the job.

How do we increase buy-in from academia?

- i) Involve researchers from academia as experts both in the steering group, at the beginning of the evaluation, and in the final phase, reviewing the final reports
- ii) To stimulate researchers' interest, it could be possible to include in the ToR a requirement specifying that the evaluation study has to be publishable (although the publication may be completed with some delay with respect to the publication of the evaluation report)
- iii) Data exchange: researchers are typically interested in good data for good publications. Data exchange could be a good way to increase buy-in and interest in institutional collaborations.

5. Communication of results and engagement of policymakers

Knowledge of CIE methods among policymakers is scarce. This makes it difficult to make the policymaker aware of (i) the importance of evaluations aimed at isolating causal impacts, (ii) of how these findings are different from the ones that can be obtained using other types of evaluations, (iii) of how different findings can be used to shape future policies. This also has effects on the degree of buy-in of policymakers on CIE evaluations.

Policymakers often have expectations about the results of some interventions. Their involvement and trust in the evaluation process (possibly not only in the communication phase) could increase their understanding of the evaluation design, its potential drawbacks, and the findings of the evaluation. MAs find it difficult to engage policymakers with every type of evaluation, not only CIE. There is the **need of building up the culture of evidence-based policy.**

However, **demand for evaluations from policymakers** is on the rise in some MS, here are some reported experiences:

- i) Policymakers are required to take more and more into account solid evidence in their activity, and the European Commission demand for evidence-based policymaking is constantly on the rise. This combined pressure that helps the management realize they need evidence.
- ii) Some policy makers specifically pushed for the use of CIE, as it can be a powerful tool to argue for increase funding for effective programmes.

iii) Stakeholders can be more engaged if they have ownership of the evaluation. Recommendations stemming from the evaluation should be thoroughly discussed with stakeholders, and relevant institutional partners (e.g. other ministries) This would ensure that the results of the evaluation inform future policy design.

Results of evaluations are useful to the extent that the generated knowledge is shared with policymakers, but also with the general public. To this end a proper communication strategy should be envisaged.

The meeting took place online, on 10 and 12 November 2020. It was organized by the European Commission Joint Research Centre (JRC) Centre for Research on Impact Evaluation (CRIE) together with DG Employment, Social Affairs and Inclusion (DG EMPL). 29 representatives from ESF Managing Authorities and Research Institutes attended from AT, BE, BG, CZ, EL, FR, HR, HU, IE, IT, LV, MT, PL, PT, SE and SI.