fteval JOURNAL
for Research and
Technology Policy Evaluation

July 2025, Vol. 57, pp. e16, 1-15
DOI: 10.22163/fteval.2025.708

THE SITUATION OF R&I POLICY EVALUATION IN COST INCLUSIVENESS TARGET COUNTRIES

KLAUS SCHUCH, NINO GOGADZE, VITALII GRYGA, ATTILA HAVAS, RENATA ANNA JAKSA, OLHA KRASOVSKA, ARZU KURŞUN, SOFIA LIARTI, ERIONA SHTËMBARI, MIROSLAV ŠTEFÁNIK, LENA TSIPOURI AND ALBENA VUTSOVA
DOI: 10.22163/FTEVAL.2025.708

ABSTRACT

The paper summarises the main findings of a discussion between professional evaluators and scholars on the situation of R&I evaluations in the so-called COST Inclusiveness Target Countries. It shows that the situation depends to a large extent on the state of the respective evaluation markets in the countries and sectors concerned. Since the implementation of EU evaluation rules and practices can give the R&I evaluation sector a modernisation boost, it makes a difference how often one is confronted with European evaluation requirements and thus also a difference whether a COST Inclusiveness Target Country is a member of the EU or not. The paper addresses the consequences of a missing or fragmented evaluation market, offers some alternative approaches (and their limitations) to overcome them and, based on the identified challenges, formulates recommendations for both governments and R&I policy administrations as well as R&I policy evaluators. The aim of this opinion paper is to shed light on the often-critical situation of R&I policy evaluation in the COST Inclusive Target Countries and to identify possible solutions for improvement.

Keywords: R&I evaluation, evaluation practices, COST Inclusiveness Target Countries, evaluation market, R&I policy intelligence

1. INTRODUCTION

Evaluations in the field of research and innovation (R&I) constitute a standard instrument in many European countries to review different issues at different points of time (ex-ante, interim, terminal and ex-post) such as the design, relevance, efficiency, processes and dynamics, effectiveness, coherence, impact and sustainability of various R&I policy interventions. These interventions can be new or existing programmes, policy instruments or other monetary and non-monetary measures. Organisations such as universities, research organisations and research funding agencies, as well as systems such as the competitive research system of a country, can also form the object of R&I policy evaluations.

Evaluations are an important element of the intelligence portfolio available for reflective and evidence-informed R&I policy making. However, it should not be assumed that the established, mostly sectoral (e.g. related to social policy, public health policy, development assistance or other policy areas) or national evaluation "systems", to use an approximation term, are similar to one another. On the contrary, they differ not only in terms of scope and scale, but also in terms of customary practices and in their respective stage of development, the (legal) and practical degree of commitment (including follow-up activities of the assessment) and the underlying evaluation culture.

At the *"REvaluation '24"* conference, organised by the Austrian Platform for Research and Technology Policy Evaluation (fteval), Fraunhofer ISI, IFRIS, Joanneum Research and the COST Action PROFEEDBACK in Vienna in December 2024, a group of scholars and evaluators discussed the different evaluation practices in the so-called COST Inclusiveness Target Countries. These include candidate[1] countries for EU membership, as well as the EU Member States with less developed R&I systems. The workshop on '*R&I Evaluation Systems in COST Inclusiveness Target Countries'* was organised by PROFEEDBACK. The results of the presentations and the discussions are summarised in this communication. They were 'distilled' by the first author drawing on the contributions of the other workshop participants and endorsed by the contributors. It is not intended to provide a detailed assessment of the situation in the reviewed countries, but rather to appeal to all those responsible in the national governments and agencies devising and implementing R&I policies to pay more attention to evaluation.

1 https://european-union.europa.eu/principles-countries-history/eu-enlargement_en; accessed on 6th February, 2025.

The following countries were represented at the discussion: Albania, Bulgaria, Georgia, Greece, Hungary, Slovakia, Türkiye and Ukraine. The aim of the workshop, and hence this paper, was to develop recommendations for improving evaluation practices based on the common problems identified by evaluation practitioners in the field of research and innovation.

2. BEING IN THE EU (OR NOT) MAKES A DIFFERENCE

GOVERNANCE STANDARDS

Although the results of the discussion clearly showed that there are problems almost all countries have to struggle with, there is one fundamental difference: whether or not a country is a member of the EU. EU membership goes hand in hand with a certain degree of *'new public management'* focusing on efficiency, efficacy and impact, which includes also the accountability of public interventions (and spending) in terms of good governance practice or, rather, 'good governance requirement'. However, good governance is not a uniformly implemented standard.

FUNDING AND RELATED EVALUATION PRACTICES

In terms of evaluation practice, the use of EU funds and the utilisation of EU programmes have led to a clear push towards mandatory and more frequent evaluations, often applying similar evaluation criteria. By contrast, evaluations of purely national interventions still tend to be rather infrequent in most reviewed COST Inclusiveness Target Countries. This is, of course, also related to the number of policy interventions in the R&I area. In countries with small portfolios of support measures for research, technological development and innovation, there is generally less need for evaluation, especially if one relies on established measures (which should, however, also be subject to regular reviews). All of the COST Inclusiveness Target Countries are characterised by below-average R&D spending as a share of GDP (compared to the EU average).

The situation is even harsher for non-EU countries. Although they also benefit from a few EU programmes, these interventions are less frequent and often less comprehensive in terms of scope and scale. In addition, evaluations of national measures are rare.

Overall, this means that in the EU Member States, especially due to the prevalence of projects funded by the Recovery and Resilience Facility (RRF), the European Social Fund (ESF) and the European Regional Development Fund (ERDF), mandatory evaluations have to be undertaken regularly, while in the candidate countries this recurring regularity is lacking and evaluations demanded by EU funding are not always connected to the respective national or local governments or administrations. This also has implications for the trickle-down effect on national arenas of action and administrative competences. While EU practices come across as a modernisation and reform agenda with normative power in the EU[2], in the candidate countries it remains in essence an external matter that may be sometimes inspiring, but as a rule does not change regular administrative practices.

CHALLENGES IN NON-EU COUNTRIES

When it comes to EU expenditures on projects carried out in candidate countries, the evaluations made are sometimes seen as inappropriate and exaggerated. The efforts required to conform to EU evaluation standards seem sometimes particularly demanding by the local authorities. This is also caused by a lack of professional evaluation departments in the national or sectoral administrations, in particular of qualified evaluation administrators, and unevenly distributed knowledge about the purpose, the added-value, and the 'rules of the game' of evaluations. Moreover, external capacities that could carry out evaluations are difficult to identify, which increases transaction costs because often there is no national 'evaluation market', expressed by a lacking database of 'certified/qualified' evaluators, in which customers and providers of evaluations can find each other easily. Sometimes *Voluntary Organisations for Professional Evaluation* (VOPEs) can act as intermediaries, if they are known to the occasional evaluation enquirers.

2 The requirements of the RRF (Recovery and Resilience Facility) are considered to become a change maker in the EU Member States due to their focus on evaluating performance. In terms of inspiring practices, the ex-ante impact assessments of policy interventions regularly exercised at European level or the European Commission's consideration of broader social impacts have also been mentioned.

3. THE CONSEQUENCES OF A MISSING OR FRAGMENTED EVALUATION MARKET

THE 'IDEAL' EVALUATION MARKET

An ideal evaluation market would be characterised by sufficient demand for evaluations that can be met by a sufficiently large number of evaluation providers. What is 'sufficient' can be approximated through several factors. These include, among others, that evaluation providers are actually in competition with each other, do not enter into market-distorting agreements, are economically, legally and personally independent of the respective clients and can quickly make the necessary capacity and expertise available for the various evaluation requirements in a specific policy area at a competitive market price.

CONSEQUENCES OF NON-IDEAL EVALUATION MARKETS

Unfortunately, 'ideal' evaluation markets are encountered only rarely. National evaluation markets are often too small and fragmented across different policy areas, which is why evaluation expertise from abroad is sometimes also drawn upon. However, since evaluations require a high level of contextual knowledge, e.g. about the national actors and their characteristics and relationships to each other, in a particular sectoral R&I system, the choice of international evaluation providers is not always expedient. Further, information and data sources that need to be considered are usually only available in the respective national language(s). On top of these, in some EU Member States with less developed R&I systems and especially in the candidate countries, prices determined by national living standards are insufficient to be able to pay international providers. Access to international evaluators works better where information and data sources are available in English or in a larger cross-border common language area, such as the German-speaking area in the EU (Germany, Austria, Switzerland). This is also the reason why there are more evaluation providers in the R&I sector in Austria available than the small country itself can provide as a market.[3]

3 Streicher, J., Polt, W. and Unger, M. (2020). Eine Untersuchung der Marktsituation im Bereich der FTI-Evaluierung in Österreich. fteval Journal for Research and Technology Policy Evaluation (50). pp. 72-81. ISSN 1726-6629; DOI: 10.22163/fteval.2020.472

ALTERNATIVE APPROACHES AND THEIR LIMITATIONS

Buying in foreign evaluation expertise may solve the ad hoc challenge of commissioning and getting a good evaluation, but it does not necessarily contribute to national or local capacity and competence building. Therefore, mixed evaluation teams that involve national and international evaluators could be prioritised when the national market is insufficient. This is easier for large consulting firms that are based in different countries to comply with, but they are not always sufficiently experienced in evaluating R&I activities.

Another approach sometimes is, to not source via an evaluation market but rely on in-house capacity inside the government (ministries, agencies) to carry out evaluations themselves. However, this has several disadvantages, such as a higher risk of compromised independence and underutilisation or overly narrow specialisation. It is costly to stockpile evaluation expertise without being able to guarantee sufficient demand for evaluations from the public sector. An advantage of such approach is a potentially better integration of the evaluation results in subsequent policy-design decisions.

To minimise or shift the cost pressure to some extent, a basic supply of evaluation expertise can also be built up via universities (e.g. by means of evaluation courses that should be offered on a recurring basis). Although evaluation is part of the scientific canon and policy evaluations in particular make use of empirical social and economic research methods, expertise built up in this way without sufficient practical experience runs the risk of remaining too theoretical and of not being able to provide the necessary contextual knowledge.

THE SELF-REINFORCING CYCLE OF EVALUATION PRACTICE

A final major shortcoming of a poorly developed evaluation market to be highlighted here is the lack of established good practices. Evaluation as an ad hoc business implies that the risk of procedural errors increases on both, the client and contractor side. On the client side, this can affect the identification and commissioning of external evaluation providers by making formal errors such as disregarding the relevant publicity and deadline requirements for public procurement. Furthermore, problems with the creation of meaningful and clear Terms of Reference (ToR) and the estimation of realistic price-quantity structures for the requested services are becoming more frequent. The creation of good ToR, the estimation of realistic budgets for requested services and the implementation of efficient and correct procurement are skills that must be learnt.

The problem of creating realistic and competitive price-quantity structures is also encountered on the provider side of evaluations. This is due to inexperience with the evaluation object and the context in which an evaluation takes place or misalignment with the client's expectations (especially in the case of unclear ToR). The design of the requested products itself, i.e. the format and focus of meaningful evaluation reports and interim presentations, also differs from conventional scientific work. Here too, tacit knowledge that is accumulated in the course of practice is vital.

Both clients and providers also face shared uncertainties, such as expectations dealing with critical points and ensuring the use and usability of evaluation results. Ethical issues, in particular, can quickly become a problem in underdeveloped evaluation markets if appropriate agreements and standards of good evaluation practice are not already in place and known and need to be negotiated from scratch. Evaluation practice creates an evaluation culture, which in turn reflects back on evaluation practice. This can become a vicious or a virtuous cycle.

4. MAIN CHALLENGES

Insufficient evaluation markets in the field of R&I policy evaluation affect all countries. As mentioned above, being part of the EU or not makes an initial profound difference. But other factors, such as the size of a country and the importance of R&I as a policy field in the national or regional system, play a role too. All the countries, presented at the workshop, are at different levels. Nevertheless, problem areas were identified that played a role for most of the evaluators from the participating countries. These include, for example, the following:

STRUCTURAL CHALLENGES

- **Underfinancing and lack of regular evaluations:** Lack of financing and infrequent evaluation tenders can cause underdeveloped evaluation markets and a lack of available qualified evaluation experts. Moreover, the price specifications for evaluation contracts are often too low, which leads to under-budgeting. This is often due to a lack of a clearly defined range of services, a lack of understanding on the part of clients about methodological efforts and the challenges of data collection, which is why price-quantity frameworks are often significantly underestimated.

- **Conceptualisation and enforcement:** In many countries, strategies, concepts, programmes and even laws developed for certain interventions in the field of R&I are significantly more ambitious than their actual implementation and enforcement. This gap between aspiration and realisation is often a consequence of inadequately secured budgets and lack of human resources that quickly render the best intentions obsolete, while political changes can further disrupt implementation. Therefore, even the measures that sound best on paper must be critically scrutinised through evaluations ('reality check').

- **Data availability and accessibility:** If evaluations can make use of existing data, this is a golden opportunity that must be seized. In particular, monitoring data and longitudinal data on specific R&I policy intentions are of great advantage in this regard. Unfortunately, this is not often the case. Access to administrative data should be ensured in any case.

EVALUATION PROCESS CHALLENGES AND POLITICAL BARRIERS

- **Planning:** Evaluations should fit into the policy cycle, which, however, is often not the case. Thus, their normative power is limited and the results rarely used.

- **Political pressure:** In some countries, constructive feedback is confused with politically motivated criticism, which puts evaluators under stress. However, the only concern of evaluators should be to produce a good and useful evaluation and to act otherwise free of vested interests (including their own) and political influence.

- **Neglecting evaluation results:** Evaluations are a central element of 'policy intelligence'. Failure to use the results of good evaluations is a failure of policy. Follow-up steps and feedback loops for redesigning funding instruments are sometimes missing.

5. RECOMMENDATIONS

The participants of the Workshop call on those responsible in politics and the administration to improve evaluation practices. The following recommendations are offered to them at the end.

RECOMMENDATIONS FOR GOVERNMENTS AND ADMINISTRATIONS OF R&I POLICIES

INSTITUTIONAL FRAMEWORK AND DEVELOPMENT

1. Evaluation should be a central, mandatory component of any substantial policy intervention. This should be enshrined in law or at least have the character of administrative instructions, which can be included in call regulations. Sometimes, however, the problem is not the legal obligation, but enforcement. More frequent evaluation requirements help developing a competent national supplier market. But ensure competition and avoid a dominant provider, otherwise the market runs danger to become disrupted in the mid- to long-term.

2. The establishment of a functional evaluation practice and culture needs time.

3. Nominate one person in each of the ministries and funding agencies responsible for R&I policy making and R&I delivery to drive forward and centrally manage the evaluation agendas for the R&I activities. This person should also be the internal contact person for other colleagues when evaluation issues arise.

4. At a general policy level, establish jointly elaborated evaluation guidelines that apply to several fields of policy (and across departments). Of course, these must be adapted and supplemented for each ministry. Especially with regard to terminology, tendering procedures, assessment standards for the offers received, ethics, transparency, processing and accountability, a jointly elaborated guiding framework helps reducing uncertainties and transaction costs. As far as terminology is concerned, it is best to use the OECD Nomenclature as a starting point.[4]

4 OECD (2023), Glossary of Key Terms in Evaluation and Results Based Management for Sustainable Development (Second edition), OECD Publishing, Paris .https://doi.org/10.1787/632da462-en-fr

INTEGRATING EVALUATION INTO POLICY CYCLES

5. Every substantial R&I policy measure should be evaluated in terms of its relevance, efficiency, effectiveness, coherence, impact and sustainability. This applies to R&I programmes, R&I instruments, R&I organisations, agencies, systems (such as a country's entire research funding system), regulatory frameworks, etc. Start by addressing institutional weaknesses, because these usually consume a large amount of public spending and focus on organisations that can have an impact and could act as agents of change for the whole system, such as persistent innovators.

6. Evaluations are not only scientific work but require also judgment. Thus, take them seriously. Good evaluations come at a price, while under-funded evaluation projects generate inferior results and are therefore unsuitable for political legitimisation and guidance. A specific evaluation project is only being carried out once by the commissioned team and it should be provided with the best possible working conditions. This includes not only a sufficient budget, but also sufficient time and access to data to do the job well, as well as the opportunity for consultation with the responsible persons in the commissioning authority regarding any professional questions that may arise.

7. Make use of evaluation results. Integrate evaluations into the policy cycles so that their results are available when needed. Timely preparation of relevant calls for tenders for external evaluation is crucial in this regard, so that evaluations can start in good time and also work long enough to deliver useful results. An evaluation does not necessarily have to end with the delivery or acceptance of the evaluation report. Continue to use the knowledge gained by the evaluators, formally or informally. Clients should provide so-called *'management responses'* to the recommendations presented by the evaluators, stating whether and how they intend to proceed with the recommendation.

8. Create transparency. This applies to both, planned or tendered evaluation procedures and the results of evaluations. Publishing evaluation reports generally strengthens credibility and fosters dialogue within the affected community, but it also forces evaluators to deliver higher quality evaluations and evaluation reports, as the public nature of the process ensures accountability and minimizes the risk of embarrassment.

9. Gradually venture into more systemic evaluations or portfolio evaluations to avoid losing sight of the big picture when evaluating individual measures only.

EMPIRICAL EVIDENCE IS NEEDED

10. For the most important public R&I interventions monitoring systems should be established, which regularly collect the requested information[5], such as number of students or number of staff, number and type of publications and patents etc. disaggregated by meaningful categories like fields of research, organisations, departments etc. or number of granted projects disaggregated by fields of research and so on. Monitoring data are important sources for R&I policy evaluations to build on, but they do not replace evaluations.

11. When creating price-quantity frameworks, both clients and evaluators should be aware that empirical research incurs costs. Sometimes, however, relevant data are actually collected by government departments (quite possibly in the course of work not directly related to the policy intervention under review). Such databases or repositories should be made available to evaluators for aggregated analyses, subject to confidentiality clauses. Sometimes databases are also commercially curated (e.g. company data, publication data). The corresponding costs for using these databases must be budgeted for. Already existing licences, e.g. in ministries or agencies, should be made available to the evaluators for the time of the evaluation and the evaluation purpose. If necessary, specific adjustments in official statistics regulation should be made to ensure the availability of disaggregated data, if this is not prohibited for other more relevant reasons (such as martial law in Ukraine).

12. Overall, the value of having functional science, research and innovation statistics that are based on OECD and EUROSTAT standards should not be underestimated. They provide valuable data for capturing national, sectoral or regional research and/or innovation systems, even though they are usually too aggregated for specific R&I policy evaluation purposes. To obtain additional and more specific data, governments should also endeavour to participate in international or European surveys, such as the European Innovation Survey or SheFigures or use

5 An inspiring example is TUBITAK's new grant management platform that uses the advantages and functionalities of advancing digitalisation and can be used for monitoring purposes.

international support offers such as the Policy Support Facility of DG Research and Innovation for more ambitious R&I policy reviews.

CAPACITY BUILDING AND TRAINING

13. Utilise the evaluation requirements arising from EU interventions and programmes (e.g. RRF) in order to gradually generate learning effects for national evaluation practices. Do not see this primarily as an additional burden, but as an opportunity for your own reform efforts and the development of national evaluation expertise. There are already many useful guidelines and training programmes, which are also available in English.

14. A possible avenue for building up national evaluation capacity is to create a database of evaluators with practice in the evaluation of international programmes. Their registration should be on a voluntary basis, but the experience of these experts, especially those with long-term practice, is a good basis for building a national pool of experts, profiled by thematic areas.

15. To counteract the lack of suitable evaluation, personnel in the long term, evaluations as a subfield of empirical economics and social science research should be more strongly promoted in academic education or in specific trainings[6]. In addition, foreign evaluation providers should be encouraged to include local staff in their teams (local content policy). Creativity in developing public-private partnerships that link the academic and consulting sectors is called for.

RECOMMENDATIONS FOR EVALUATORS

16. Evaluation is neither rocket science, nor an easy scientific endeavour, because it requires profound context-related knowledge as well as robust methodological and social skills. As regards advanced evaluation, global networks such as OECD or UNESCO and certain academic organisations provide pertinent materials and tools to learn from (e.g. bibliometrics, patent analysis). In addition, being well versed in the methods of empirical social sciences and economic research is essential. Additional support

6 To give an example: https://www.mioir.manchester.ac.uk/study/short-courses/evaluation-of-science-and-innovation-policies/; accessed on 10 February 2025.

and knowledge sharing of best evaluation practices can also be used within specialized networks and platforms.[7]

17. Choose your performance indicators wisely. Sometimes R&I policy interventions expect too much and overburden the performance or overestimate the incentive mechanisms of their measures. Occasionally, they are simply poorly designed. Unfortunately, the evaluation questions often have excessive expectations, which is why it seems useful to supplement them with more finely grained or alternative indicators. Additionally, exploring alternative data sources, such as mining websites, can be more effective than of solely relying on self-declared data from funded projects.

18. On the part of the evaluators, it is recommended to expand the occasionally dominant self-image of their role as critical assessors and to stronger promote, in addition to accountability aspects, learning and steering effects. However, this does not mean that evaluations should be uncritical. Particularly in smaller countries with few evaluators, there is a risk of conformism or of being co-opted by the commissioning bodies to obtain favourable evaluation results. To avoid possible dependencies and to allow fresh perspectives, it can make sense to involve colleagues from abroad who are unaffected by local networks and cliques.

19. Try to express your evaluation results, including the recommendations made, as clearly as possible in order to be able to guide action. Avoid vague and superficial statements or recommendations that may not even address anyone in particular.

20. Explore opportunities beyond government contracts, if this is possible. Take advantage of the – admittedly not always numerous – demand for evaluations from the private or civil society sectors as well. This will enable you to apply and develop your methodological skills, giving you a better sense of what learning from evaluation or steering through evaluation means. Additionally, it will provide insights into the practical usability of evaluation and how you can contribute to this process.

7 Examples: Austrian Platform for Research and Technology Policy Evaluation (https://fteval.at/en/; accessed on 10 February 2025). It runs R&I policy evaluation conferences held every three years. Another example is DeGEval (https://www.degeval.org/en/working-groups/research-technology-and-innovation-policy/; accessed on 10 February 2025), which runs a dedicated working group for research, technology and innovation policy.

21. Insist on being allowed to publish your evaluation reports. It benefits your professional portfolio and CV. More importantly, it fosters an evidence-informed dialogue on science, research and innovation policy, contributing to the advancement of the field.

AUTHORS

KLAUS SCHUCH
ZSI – Centre for Social Innovation
ORCID: 0000-0002-7603-4384

NINO GOGADZE
Centre for Science and Technology Studies (CWTS), Leiden University
ORCID: 0009-0006-4498-9481

VITALII GRYGA
Institute for economics and forecasting of the NAS of Ukraine
ORCID: 0000-0003-0026-3591

ATTILA HAVAS
Institute of Economics, Hungarian Academy of Sciences Centre of Excellence
ORCID: 0000-0002-8571-8817

RENATA ANNA JAKSA
HÉTFA Research Institute
ORCID: 0009-0007-9299-5402

OLHA KRASOVSKA
Ukrainian Evaluation association
ORCID: 0009-0006-3230-684X

ARZU KURŞUN
Giresun University
ORCID: 0000-0001-9560-7689

SOFIA LIARTI
OPIX P.C.
ORCID: 0000-0002-2525-6411

ERIONA SHTËMBARI
University of New York Tirana
ORCID: 0000-0003-1390-8301

MIROSLAV ŠTEFÁNIK
Institute of Economic Research, Slovak Academy of Sciences
ORCID: 0000-0002-9993-7885

LENA TSIPOURI
National and Kapodistrian University of Athens & OPIX P.C.
ORCID: 0000-0002-7454-6271

ALBENA VUTSOVA
Sofia University " St. Kliment Ohridski"
ORCID : 0000-0001-8223-6727