

The monitoring and evaluation of the 2014-20 EU Cohesion policy programmes



IQ-Net Thematic Paper 36(2)

Laura Polverari



May 2015

European Policies Research Centre
School of Government & Public Policy
University of Strathclyde
40 George Street
Glasgow G1 1QE
United Kingdom

Tel: +44 (0) 141 548 3672

Fax: +44 (0) 141 548 4898

Email: eprc@strath.ac.uk

The place of useful learning

The University of Strathclyde is a charitable body, registered in Scotland, number SC015263

The research for this paper was undertaken by EPRC in preparation for the 38th IQ-Net meeting held in Vila do Conde, Portugal on 27-29 May 2015. The paper was written by Laura Polverari.

The paper is the product of desk research and fieldwork visits during Spring 2015 to national and regional authorities in EU Member States (notably partners in the IQ-Net Consortium). The field research team comprised:

- Stefan Kah (Austria, Slovakia, Slovenia)
- Dr Arno van der Zwet (Belgium)
- Dr Lucie Jungwirthová (Czech Republic)
- Lise Smed Olsen (Denmark)
- Heidi Vironen (Finland)
- Timothée Lehuraux (France)
- Dr Sara Davies (Germany)
- Dr Eleftherios Antonopoulos (Greece)
- Dr Carlos Mendez (Portugal, Spain)
- Rona Michie (United Kingdom)
- Dr Laura Polverari (European Commission)

EPRC thanks all those respondents from EU, national and regional authorities and other organisations who participated in the research. EPRC also gratefully acknowledges the financial support provided by participating Member States and regions, whose contributions are co-financed by technical assistance from the European Structural Funds. The author is grateful for helpful advice and feedback on earlier drafts of this paper from the IQ-Net research team, and for support with the formatting of the paper from Alyson Ross. The report is, however, the responsibility of the author alone.

The partners in the IQ-Net network are as follows:

Austria

- ÖROK Secretariat

Belgium

- Enterprise Agency Flanders

Czech Republic

- Ministry of Regional Development

Denmark

- Danish Business Authority

Finland

- South and West Finland (Etelä- ja Länsi-Suomi)

France

- Commissariat Général à l'Égalité des Territoires (CGET)

Germany

- Nordrhein-Westfalen (North Rhine-Westphalia), Ministry for Business, Energy, Industry, SMEs and the Craft Sector, EU Affairs Unit

Greece

- Management Organisation Unit of Development Programmes S.A.

Portugal

- Agency for Development and Cohesion (ADC)

Slovakia

- Central Coordination Body, Government Office

Slovenia

- Government Office for Development and European Cohesion Policy

Spain

- País Vasco (Basque Country), Provincial Council of Bizkaia, Department of Economy and Finance

United Kingdom

- Department of Communities and Local Government
- Scottish Government
- Welsh European Funding Office

For further information about IQ-Net, and access to the full series of IQ-Net Papers, please visit the IQ-Net website at: <http://www.eprc.strath.ac.uk/ignet/>

To cite this paper, please use the following: Polverari L (2015) 'The monitoring and evaluation of the 2014-20 Cohesion policy programmes' *IQ-Net Thematic Paper 36(2)*, European Policies Research Centre, University of Strathclyde, Glasgow.

Disclaimer

It should be noted that the content and conclusions of this paper do not necessarily represent the views of individual members of the IQ-Net Consortium.

THE MONITORING AND EVALUATION OF THE 2014-20 EU COHESION POLICY PROGRAMMES

1. INTRODUCTION	12
2. MONITORING AND EVALUATION IN THE NEW PROGRAMMING FRAMEWORK	13
2.1 Evaluation as a performance tool	13
2.2 Monitoring obligations	15
2.3 Performance framework obligations	16
2.4 Evaluation obligations	17
2.5 A new approach to evaluation.....	17
3. NEW MONITORING ARRANGEMENTS IN IQ-NET PROGRAMMES	21
3.1 Main changes to monitoring systems.....	21
3.1.1 Improvement of indicators	21
3.1.2 More systemic reorganisation	23
3.1.3 Introduction of new systems.....	23
3.1.4 The rationales for the changes introduced.....	25
3.2 Open challenges	26
3.3 Identifying results indicators.....	28
3.4 Enhanced cooperation with National Statistical Offices and additional data-gathering arrangements	31
4. THE 2014-20 EVALUATION PLANS.....	33
4.1 State of play	33
4.2 Objectives and timing.....	37
4.3 Scope and focus	38
4.3.1 Thematic approaches.....	39
4.3.2 Programme-wide approaches	41
4.3.3 Mixed approaches: thematic and programme-wide	41
4.4 Resources and responsibilities	42
5. FOCUS AND METHODOLOGICAL APPROACHES	46
5.1 Balance between implementation and impact evaluation	46
5.2 Methods – balance between theory-based impact evaluation and counterfactual impact evaluation	47
6. CONCLUSIONS.....	50

7. ANNEX I – MINI-CASE STUDIES ON IMPACT EVALUATIONS UNDERTAKEN IN 2007-13 and evaluation plans	53
7.1 Austria – Impact evaluation of ERDF-funded federal environmental measures in 2007-13.53	
7.1.1 Rationale	53
7.1.2 Delivery	53
7.1.3 Challenges	56
7.1.4 Assessment.....	56
7.2 Czech Republic – Pilot counterfactual evaluation of human resources and employment OP, support area 1.1	57
7.2.1 Rationale	57
7.2.2 Delivery	58
7.2.3 Challenges	58
7.2.4 Assessment.....	59
7.3 Denmark – The register-based approach to counterfactual impact evaluation	60
7.3.1 Rationale	60
7.3.2 Delivery	61
7.3.3 Challenges	62
7.3.4 Assessment.....	63
7.4 Portugal – Counterfactual impact evaluation of business incentive schemes in Portugal	64
7.4.1 Rationale	64
7.4.2 Delivery	64
7.4.3 Challenges	64
7.4.4 Assessment.....	65
7.5 Poland – Determining the value of selected economic indicators for beneficiaries of Regional Operational Programmes and for selected control groups	66
7.5.1 Rationale	66
7.5.2 Delivery	67
7.5.3 Challenges	68
7.5.4 Assessment.....	68
7.6 Wales – 2012 ESF leavers survey	70
7.6.1 Rationale	70
7.6.2 Delivery	70
7.6.3 Challenges	71
7.6.4 Assessment.....	71
8. ANNEX II – Evaluations foreseen in the Czech Partnership Agreement evaluation plan	72

THE MONITORING AND EVALUATION OF THE 2014-20 EU COHESION POLICY PROGRAMMES

EXECUTIVE SUMMARY

The legislative framework for Cohesion policy 2014-20 places unprecedented emphasis on the results-orientation of programmes. Results-orientation is being sought through at least three main components: ex ante conditionalities; programme design according to a logical framework; and, more emphasis on monitoring, reporting and evaluation. New rules on monitoring and evaluation were thus introduced, including a shift in emphasis in evaluation activities from implementation evaluation to impact evaluation, through two recommended approaches: theory-based and counterfactual.

Monitoring

In response to the new regulatory framework, IQ-Net partners have been introducing minor (e.g. Austria, Slovenia, Vlaanderen, Wales) or more significant changes to their existing monitoring systems (e.g. Denmark, Finland, Pomorskie, Portugal, Spain), or launched new monitoring systems altogether (France, Greece, England, Scotland and the Czech Republic, where the 2007-13 system went through a major overhaul). Changes have related to: the design, interpretation and tracking of indicators (e.g. Denmark, Greece, Portugal, Spain); the integration with domestic monitoring systems (e.g. Finland); introduction of interoperability across different platforms (e.g. Czech Republic); pursuit of the e-Cohesion agenda (transforming the monitoring system in full management tools for the e-administration of projects, e.g. Czech Republic, France, England); simplification of forms and procedures to lighten the load on project beneficiaries (e.g. Finland, Nordrhein-Westfalen); reassignment of monitoring responsibilities (Nordrhein-Westfalen); territorial reorganisation (e.g. in England); integration of all ESI Funds in a single platform (Czech Republic); and the introduction of improved accountability and quality control procedures (e.g. in Spain, Scotland).

As well as to fulfil regulatory obligations, the changes introduced can be linked to a number of different goals which include the simplification and reduction of the onus on project beneficiaries and programme management; the improvement of the timeliness and reliability of data submission; the improvement of data quality; harmonisation, integration and interoperability across programmes, funds and domestic systems to allow aggregation of information and the elaboration of summative views; increased linkages with management and accounting systems, to ensure timeliness and accuracy and a reduction of error rates in the processing of projects; and, improved ownership and management efficiency.

In addition to serving the needs of evaluation, the new or improved monitoring systems are intended to enable ongoing learning about the performance of programmes, funds, interventions or projects; to provide MAs with timely information to allow them to introduce adjustments to the programmes where needed; and, in a number of cases, to act as integrated management platforms for projects and programme management alike.

There are nevertheless both unresolved challenges and new tensions, linked to the rolling out and the complexity of the new systems, to the strengthened checks on target setting and on the data submitted by projects, and to privacy legislation and the treatment of sensitive data.

Over a third of IQ-Net partners have met difficulties with the selection of results indicators and the identification of the related baselines and targets. These difficulties have related to: the appropriate definition of objectives and intervention logic (e.g. in Greece and Slovenia); capacity issues, i.e. the difficulty for actors more accustomed to dealing with output and financial indicators to deal with a new type of indicator (e.g. in Greece, France); the uniform interpretation of indicators across different actors, intermediate bodies and applicants within a region or country (e.g. in Finland, England); the measurement of impact (and thus target values) for programmes whose funding represents only a small portion of public policy in the region (e.g. Austria, Nordrhein-Westfalen, Wales); and the different guidance on baseline and target-setting provided by different Commission DGs (a problem for multi-fund OPs, e.g. the Czech Republic, Pomorskie).

IQ-Net partner views on whether the data gathered by the monitoring systems will be adequate to fulfil the needs of evaluations are mixed, as is the confidence in the quality of the new monitoring systems. On the whole, however, IQ-Net partners appear positive regarding the fact that the new monitoring systems will represent an improvement compared to the past and that the new results-orientation will be useful in terms of allowing programmes to gain a better overview of what is being achieved.

A trend, across a number of IQ-Net partners, is the increased use of administrative data and strengthened cooperation with national Statistical Offices (e.g. in the Czech Republic, Denmark, Greece, Portugal, Slovenia, Spain and others). This applies particularly to results indicators and to the data on individuals and firms, which could be used to allow counterfactual impact evaluation of projects and interventions. In some cases, investments are being planned to strengthen national statistical databases (e.g. in Greece) and/or undertake additional data collection exercises, tailor-made to the information needs of programmes, through ad hoc surveys (e.g. Austria, Slovenia, France).

Evaluation

IQ-Net partners have also been working on the preparation of the Evaluation Plans required by the regulations. These have to be submitted for approval to the Programme Monitoring Committee (PMC) within one year from the adoption of the programme (arts. 114.1 and 110.2 CPR) and can have different scope (a single OP or more OPs, one or more funds etc.).

The state of play amongst IQ-Net partners with the elaboration of Evaluation Plans (EPs) is mixed, partly reflecting the varied state of OP approval. Most EPs are currently still being drafted, with the aim of submission to the PMCs between June and December 2015. Where progress is not advanced, this has sometimes been due to the necessity to focus more on the strategic drafting and negotiation of the programmes and/or their operational launch (e.g. in Greece, Spain, Vlaanderen).

The Evaluation Plans have different scope, largely reflecting the institutional settings of a country or strategic or operational choices (e.g. to allow exploiting economies of scale or the achievement of whole-country overviews etc.). Evaluation plans cover: a single programme, whether single-fund (e.g. in Nordrhein-Westfalen, England and in the French regions) or pluri-fund (as in Pomorskie and in

Slovenia); a single fund across an entire country (as in Austria); more than one programme within a region (in Wales, where a single EP will cover five ESI Funds programmes (two ERDF OPs and two ESF OPs, as well as the rural development plan); a single, national EP for all ESI Funds (as in Portugal, where there will be a single national EP for ERDF, ESF, EAFRD and EMFF co-funded programmes), or a single, national EP for ERDF and ESF (as in Denmark); or an EP for the whole Partnership Agreement, plus individual EPs for the OPs – as in the Czech Republic and Slovakia. In some cases, decisions on the coverage of the EPs are still pending.

In many cases – but not all, e.g. not in Greece – IQ-Net Managing Authorities already had Evaluation Plans in 2000-06 and 2007-13. These experiences were helpful to the MAs in terms of allowing them to draw lessons for the current Plans.

The guidelines on EP drafting provided by the Commission were deemed helpful by various IQ-Net partners (for instance Denmark, Portugal, Slovenia); there are however concerns that the Commission's guidance documents and appraisal of the EPs might go beyond the regulations (Nordrhein-Westfalen). Some areas of concern were also highlighted in relation to specific issues, for instance the need to include evaluation questions (which is felt to be premature at this stage, e.g. by Czech authorities), discrepancies in the guidance provided for ERDF and ESF (Czech Republic, Pomorskie), and the requirement to appraise impact on macro-indicators for small programmes (Denmark).

The objectives associated with the EPs and the evaluations therein contained are pluri-fold: from identifying evaluation (and thus data) needs in a timely manner (e.g. Portugal and Slovenia); to feeding the 2017 and 2019 Strategic Reports (e.g. Austria); to contributing to learning on 'what works' and on whether programmes require adjustments (e.g. Wales, Vlaanderen, Greece, Nordrhein-Westfalen); to appraising the added value of projects and types of interventions, and their contribution to programme goals (e.g. Denmark, Nordrhein-Westfalen); to gauging the effectiveness, efficiency and impact of programmes and/or individual priorities (e.g. Vlaanderen, Greece, Spain), improving the evaluation environment and the know-how for the management and delivery of evaluation (e.g. in the Czech Republic), and ensuring that the evaluations undertaken are relevant and of good quality (e.g. in the Czech Republic, Portugal, Slovenia and Scotland).

In a majority of IQ-Net partners, implementation evaluation will dominate the agenda during the first half of the programme period; impact evaluation in the second half (e.g. Czech Republic, France, Slovakia, Spain and others). Some IQ-Net partners, on the other hand, will adopt staggered (Nordrhein-Westfalen) or flexible approaches (Wales), or concentrate evaluation activity in the later part of the programming cycle.

The focus of evaluation activities will vary considerably across IQ-Net partners, and is largely still under development. Mostly, thematic – rather than programme-wide approaches – will be favoured.

Most IQ-Net partners have not yet finalised their budgets dedicated to evaluation activities (e.g. Austria, England, Greece, Nordrhein-Westfalen, Pomorskie, Portugal, Spain, Scotland). Where resources have been set aside (indicatively) for evaluation, they vary considerably, ranging from c. €600,000 in Vlaanderen, to c. €38 million in France – with various measures in-between. Allocations may be higher than in the 2007-13 period, where sums set aside for evaluation were not fully spent in all cases (e.g. in Slovenia where the 2007-13 evaluation budget of c. €5 million - 5 percent of the

Slovenian TA resources - was used only by about half). It is interesting to note that where evaluation responsibilities are devolved to programmes or regions, such as in France, a considerable degree of variation can be expected in the financial weight assigned to evaluation activities which is not correlated to the financial scale of programmes.

IQ-Net partners display different preferences in relation to whether evaluation work will predominantly be carried out in-house or outsourced to external evaluators through tendering procedures. In some cases, both approaches will be pursued (for instance in Austria, Czech Republic, Denmark and Spain). However, the use of external evaluators appears to be the favoured approach, in continuity with past practices. In at least one case (England) procurement will be done through a framework contract, along the model similarly followed by EU institutions (European Commission, European Parliament). In some cases, the evaluation market is perceived to be still immature and in need for further development (e.g. Greece, Portugal).

The coordination of evaluation activities is generally undertaken by national coordinating authorities or, where these differ, the Managing Authorities, generally supported by Evaluation Units, Evaluation Steering Groups or Evaluation Committees (e.g. in the Czech Republic, Nordrhein-Westfalen, Pomorskie, Slovenia, Spain).

In a number of cases the balance between implementation and impact evaluation is yet to be defined and will emerge as the EPs are finalised or even during implementation. For instance, in Austria and Slovenia the choice will depend on the evaluation questions that will be established and will be defined on a case by case basis. In some cases – notably in the Czech Republic, Slovakia and Spain – the intention is to realise a balanced approach covering both implementation and impact. In others, on the contrary, a shift towards impact evaluation (compared to 2007-13) appears quite pronounced (in Denmark, Finland, France, Pomorskie, Slovenia and Spain). Implementation evaluation will continue to dominate in Vlaanderen, largely due to the small financial scale of the programme. Nordrhein-Westfalen will pursue a more individual approach, i.e. a mix of implementation and project-specific evaluations (geared towards added value and results rather than impact on the regional economy). The view is that impact evaluations would be scarcely meaningful, given the financial size of the programme.

Only a few IQ-Net partner programmes have a degree of clarity in relation to the approaches that they'll adopt to evaluate impacts and most are not yet clear about the relative weight between counterfactual impact evaluation and theory-based impact evaluation. In some cases, this balance will emerge as the evaluation work unfolds, and the choice of methods will emerge from the tendering procedures (e.g. in Nordrhein-Westfalen and Slovenia). Where there is some more clarity on methods – such as in the Czech Republic, Wales and Denmark – preferences vary: theory-based and case study approaches are expected to dominate in the Czech Republic, CIE in Wales, and a mix of the two in Denmark. MAs are gearing up towards evaluation and are in some cases acquiring the necessary skills and information to make such decision in an informed manner (with dedicated training, as in Finland).

Concerns about the ability to carry out impact evaluation have emerged and relate to a number of factors:

- the measurability of impact on macro-aggregates when the programmes are a only small portion of wider public policy spending (e.g. in Austria, Vlaanderen, Nordrhein-Westfalen);
- the timetable according to which impacts can be measured, which is difficult to reconcile with the short-term interests of politicians (e.g. in Austria);
- the operational and methodological difficulties entailed by CIE (e.g. in Greece);
- the limited or even absent experience of impact evaluation in 2007-13, with regard to both CIE and to theory-based impact evaluation. Even when impact evaluations were attempted in 2007-13, these were sometimes found to have been of relatively low quality or improvable, due to the limited expertise and a lack of quality data;
- a lack of generalised evaluation culture (in Greece) that inhibits the use of more sophisticated evaluation methods that require technical expertise;
- the resource intensity of impact evaluation and the difficulty, for small administrations (such as in Vlaanderen), to supply the human resources that are necessary for the sourcing of necessary data and deal with evaluators; and
- the weak resilience of evaluation planning against institutional changes and budget cuts.

Capacity issues are being tackled with improvements to monitoring systems and collaboration with data providers and with dedicated training initiatives.

Conclusions

By and large, IQ-Net partners view with favour the results-orientation focus of the new regulations and, related, the new evaluation requirements and the shift in emphasis from implementation to impact. These changes are considered to represent a move forward compared to the past programme period when financial absorption, partly linked to the economic crisis, dominated the agenda.

The expectation amongst IQ-Net partners is that the new requirements will result in many cases in more and better reporting and evaluation, and in an improved understanding of the achievements of Cohesion policy in the regions and Members States, and its impact upon target groups. However, there is awareness that the availability of the right data and adequate data quality have been an issue in the past, and that this has hindered the ability of programme authorities to undertake good impact evaluation in 2007-13. This is why IQ-Net partners are investing resources and efforts towards improving programmes' monitoring systems and the availability and usability of administrative data, for example with enhanced collaboration with national statistics offices.

Capacities – within MAs, implementing bodies, project holders and the evaluation community – continue to be a problem in some countries, and new challenges are also emerging in relation to the necessity to carry out evaluations that require different skills than those employed in 2007-13. Dedicated training initiatives (both on monitoring and evaluation) have been put in place in some countries. However, only time will tell if these measures will be sufficient to enable the realisation of quality evaluation, particularly impact evaluation.

In designing the new programmes, the identification of results indicators has been challenging for a number of IQ-Net programmes and there are still various open issues. Nevertheless, views gathered from IQ-Net partners and DG Regio Evaluation Unit indicate that the experience has been helpful and that this work has helped making the programmes more focussed and better targeted towards achievements.

The Evaluation Plans are seen by and large as useful by IQ-Net partners, for example because they allowed for a more systematic reflection on the goals of evaluation. In most cases IQ-Net partners have learnt from past experiences of designing and implementing Evaluation Plans. However, some have voiced concerns about some components of the EPs, as requested by Commission guidelines, such as the need to specify evaluation questions and methodologies at a too early stage, and about the approach that might be taken by the Commission in their appraisals of the EPs, which is feared to go beyond what is foreseen by the regulations.

Insofar as it is possible to provide an overview at this stage, IQ-Net partners are taking different approaches with regard to the coverage, objectives and focus of their EPs. Choices regarding who undertakes evaluation work – whether the MAs or other bodies within the public administration, or external consultants selected via public procurements – also vary. Such decisions often reflect the financial resources and skills available and the expertise required for different types of evaluation. Whatever the choices, emphasis is intended to be placed on the quality control of evaluation outputs.

Whilst impact evaluation will gain prominence on the whole, as required by the new regulatory framework, it is still too early to appraise how such type of evaluation will be carried out in practice, and whether the skills and data in place will be adequate to enable it. Based on the information available at this stage, CIE appears to be preferred over theory-based impact evaluation; however, a number of concerns have been expressed by IQ-Net partners about the practical applicability of both these approaches.

Two themes that did not emerge particularly strongly from the field-research are the dissemination and follow-up of evaluation activities, and the involvement of stakeholders in the design and delivery of evaluation, which nevertheless are important aspects of the evaluation cycle (as well as required by the CPR).

The research undertaken leads to a number of questions:

- first, data quality continues to be an issue and some partners are implementing Action Plans to this purpose:
 - Where successful measures have been put in place to improve the availability of administrative data and/or the quality of monitoring data, can any lessons be drawn from these (for the benefit of other IQ-Net partners)?
- second, with regard to the methods for impact evaluation IQ-Net partners have expressed different views about their preference amongst CIE and theory-based evaluation. In many cases, IQ-Net partners are still unsure about the extent to which such approaches will be used:
 - What are partners' views about the practical benefits and drawbacks of these approaches?
 - Do partners think that CIE and theory-based evaluation are sufficient methods to appraise impacts?
 - What are the barriers to the successful realisation of impact evaluation?
 - Would there be merit in shifting the focus of impact evaluation, so that it transcends the co-funded programmes and encompasses instead also other domestic policies (e.g. where programmes, like in Austria, Vlaanderen, Denmark, Nordrhein-Westfalen,

are too small to affect macro-indicators) or the broader framework of EU investments towards the achievement of Europe 2020 targets?

- Third, with regards to the planning of evaluation activities:
 - What would improve the utility of the Evaluation Plans?
 - Have stakeholders been involved in the planning and design of evaluation activities? Where this has been the case, are there good practices that IQ-Net partners can share in this respect?
- Fourth with regard to capacity. After four cycles of Cohesion policy implementation, and despite sizeable investments dedicated to this, there are still perceived weaknesses in a few countries about evaluation capacities and lack of evaluation culture:
 - What has failed in past programme periods in this respect?
 - What would be useful measures to address these shortcomings?
 - The domestic context, as illustrated amongst others in the Sixth Cohesion Report, plays an important part in determining a country's or region's ability to effectively utilise Cohesion policy resources. The same can presumably be argued about evaluation capacity. Are the ex ante conditionalities considered sufficient to address the shortcomings faced by some Member States? And, what is IQ-Net partners experience with these?

THE MONITORING AND EVALUATION OF THE 2014-20 EU COHESION POLICY PROGRAMMES

1. INTRODUCTION

Against the background of the new regulatory framework, this paper reviews the varied experiences of IQ-Net partners with the fulfilment of the new monitoring and evaluation requirements. It is based on secondary source research and fieldwork interviews carried out in all the IQ-Net partner countries and regions with Managing Authorities and Intermediate Bodies, as well as an interview with the European Commission (DG Regio).

The paper starts with a review of the new monitoring and evaluation obligations, their rationale and the main characteristics of the new evaluation approach, characterised by an intended shift from evaluating implementation to appraising impact (Section 2).

It then goes on to discuss the new monitoring arrangements in IQ-Net partner programmes, including the changes introduced to monitoring indicators and systems, the selection of results indicators, the main issues with the design and operation of the 2014-20 monitoring systems, and the challenges that remain to be addressed, as well as the additional efforts undertaken to supplement monitoring data with administrative and survey data, to better serve the needs of evaluation (Section 3).

Section 4 provides a review of the state of play and of the content of IQ-Net partners' Evaluation Plans and IQ-Net partners' views of the guidelines provided by the European Commission. Section 5, to follow describes the methodological approaches favoured by IQ-Net partners, particularly in relation to impact evaluation, and the challenges being faced.

Some conclusions and questions for discussion are presented in Section 6. As most IQ-Net partners are still at early stages with their EP preparations, the aim of the report is to take stock of current experiences, and stimulate and inform discussion at the IQ-Net meeting, with a view of facilitating an effective exchange of good and bad practices, and support mutual learning.

Two annexes, at the end of the paper, provide examples of impact evaluations undertaken by IQ-Net partners in 2007-13 (Annex I) and the evaluation activities foreseen by the Evaluation Plan for the Czech Partnership Agreement (Annex II).

2. MONITORING AND EVALUATION IN THE NEW PROGRAMMING FRAMEWORK

2.1 Evaluation as a performance tool

The legislative framework for Cohesion policy 2014-20 places unprecedented emphasis on the results-orientation of programmes. This is arguably an outcome of a variety of disparate impulses, including: the adoption within the European Commission of public management paradigms advanced by international bodies such as the OECD and UNDP, which over time have placed increased emphasis on performance;¹ the effect of the recent recession on public spending programmes and the related imperative, for public administrations, to deliver more results with reduced resources, and to provide evidence on the effectiveness and efficiency of public spending;² and the mixed evidence on the achievements of previous Cohesion policy cycles, despite significant volumes of expenditure (historically equivalent to c. one-third of the EU budget) and the amount of resources invested in evaluation activities.³

Results-orientation can be achieved through 'results-based management' (RBM), which is an approach to the management of public policies intended to:

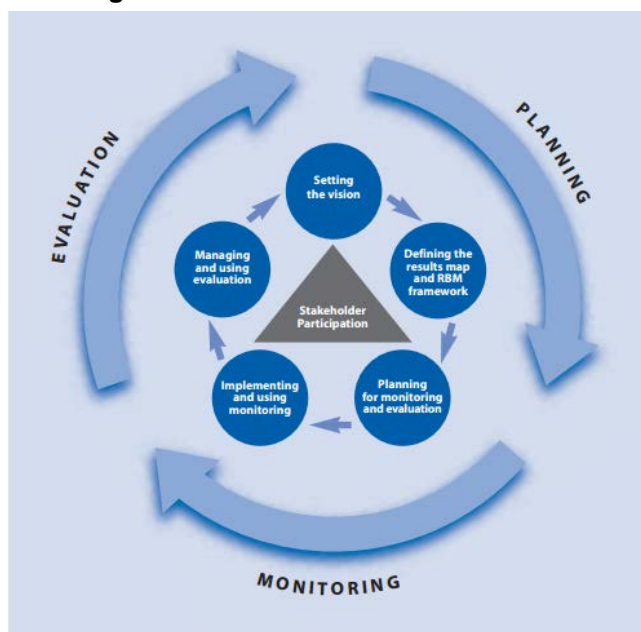
- 'support substantive accountability ...
- prompt corrective action
- ensure informed decision-making
- promote risk management
- enhance organizational and individual learning

linked together in a continuous process' (UNDP 2009, p. 11), as illustrated in Figure 1 below. This is the logic underlying the 'programming' and 'results-orientation' principles of Cohesion policy.

¹ OECD 2011, 2008 and 2005 quoted in Van Ongevalle and Huyse H (2014) 'Dealing with complexity through actor-focused planning, monitoring and evaluation (PME)', *Evaluation*, 20.4, 447-466, and UNDP (2009) *Handbook on Planning, Monitoring and Evaluating for Development Results*.

² Dahler-Larsen P (2007) 'Evaluation and Public Management', in Ferlie E, Lynn L E and Pollitt C (2007) *The Oxford Handbook of Public Management*, p. 618.

³ Bachtler J, Begg I, Polverari L and Charles D (2013) *Evaluation of the Main Achievements of Cohesion Policy Programmes and Projects over the Longer Term in 15 Selected Regions (from 1989-1993 Programme Period to the Present)*, *Final Report to the European Commission (DG Regio)*, European Policies Research Centre, University of Strathclyde (Glasgow) and London School of Economics.

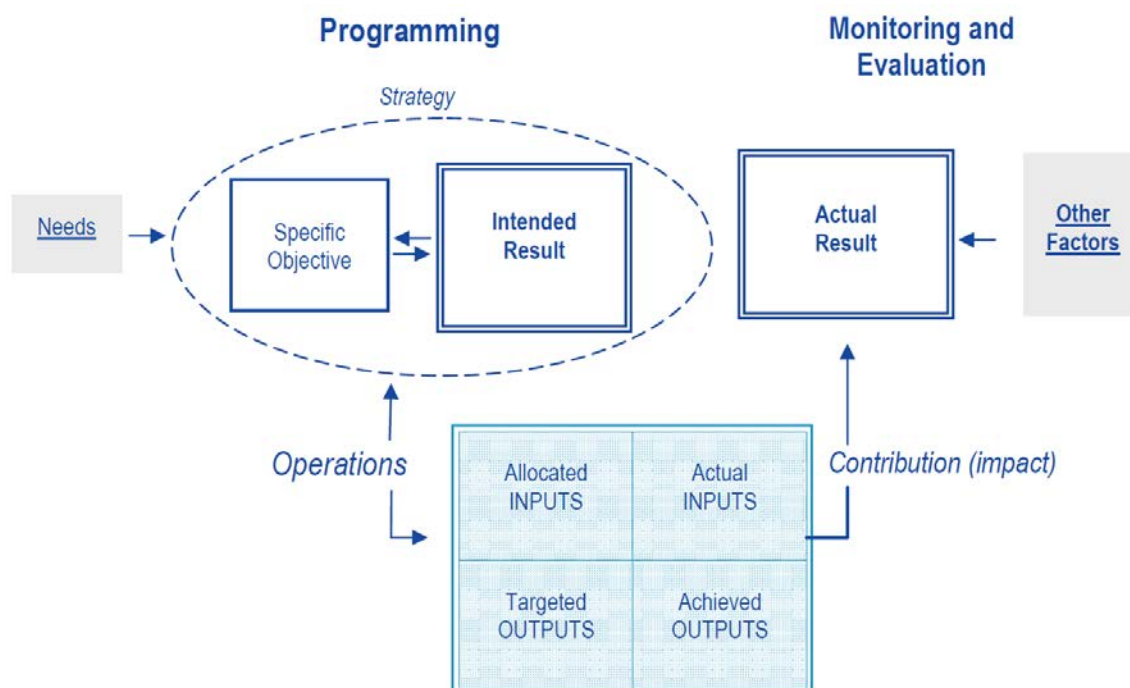
Figure 1: Results-based management

Source: UNDP (2009) *Handbook on Planning, Monitoring and Evaluating for Development Results*, p. 10.

Within Cohesion policy, results-orientation is being sought through at least three main components.

- First, **ex ante conditionalities** have been negotiated to ensure that the necessary conditions for successful programme delivery are in place before programmes are launched or that they are established soon after their adoption.
- Second, more explicit efforts have been made to design programmes according to a **logical framework** (see the two examples provided in Figure 2 and Figure 3 below), whereby programmers start with the appraisal of the need to be addressed and the results to be achieved through the interventions to fulfil such need, rather than from the resources available (essentially a reversal of the traditional input-driven logic that has dominated the programming approach so far). This approach should lead to a '**clearer articulation of the policy objectives** [which] is key to implement a results-oriented policy and moving away from an excessive focus on the absorption of funding';⁴
- Third, as a consequence of the above, **more emphasis is being placed on monitoring, reporting and evaluation** during and after programme implementation to: (i) test *in itinere* whether the anticipated results are being obtained or, conversely, whether efforts should be re-targeted (monitoring, reporting and evaluation), and (ii) appraise ex post what has been achieved thanks to the programme – in other words the contribution of the programme to addressing the initial need or programme impact (evaluation).

⁴ European Commission (2014) *The programming period 2014-2020. Guidance document on monitoring and evaluation, European Cohesion Fund, European Regional Development Funds, Concepts and Recommendations*, March 2014, available online at http://ec.europa.eu/regional_policy/sources/docoffic/2014/working/wd_2014_en.pdf (last downloaded 3 March 2015).

Figure 2: Example of a simplified logical framework

Source: European Commission (2014) *The programming period 2014-2020. Guidance document on monitoring and evaluation, European Cohesion Fund, European Regional Development Funds, Concepts and Recommendations*, March 2014, p. 5.

Timely and focused monitoring during implementation, and robust evaluation before, during and after programme implementation, are essential components of any results-based management system and of the performance framework of Cohesion policy. Within Cohesion policy, this is enshrined in a number of rules, which include a variety of tasks for Member State authorities.

2.2 Monitoring obligations

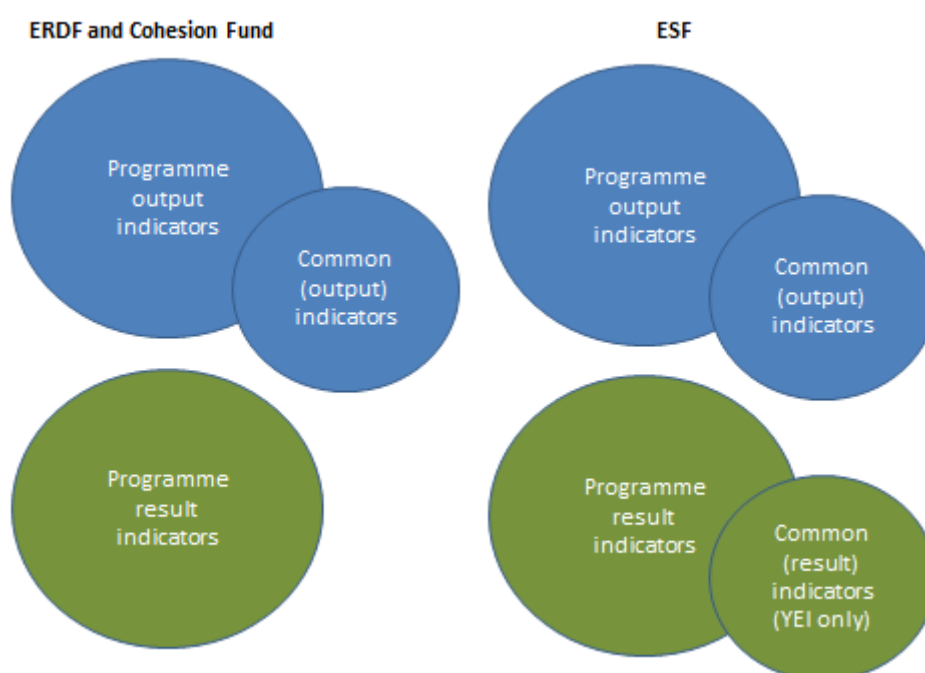
The main tasks in relation to monitoring include setting up a monitoring system that tracks progression towards established target values (to be reached in 2023) in relation to:

- **‘Output indicators’** linked to the actions funded by the programmes (art. 27.4(b) and art. 96.3(b) CPR), for example the number of recipients (e.g. firms, trainees) supported, the number of additional jobs created through the investment, the km of new metro/rail lines built, etc.⁵
- **‘Result indicators’** linked to the specific objectives of the programmes (art. 27 and art. 96 CPR). Results indicators measure the change determined by the investments realised for the target user-groups. Results indicators are not compulsory for programmes in relation to the performance framework, as this refers to the output indicators. However, with their focus on

⁵ For results indicators, the programme documents have to provide starting baseline values (art. 6 ERDF regulation). Where programmes are not able to provide baseline values, they are required to do so soon after adoption, according to a timetable specified in a dedicated action plan (art. 19.2 of ex ante conditionalities). For output indicators, on the other hand, no baseline values are foreseen, given that output indicators track the outputs of the investments made, and their starting point is therefore equal to zero (e.g. number of firms assisted by a programme measure).

change, they are crucial for the realisation of the results-orientation approach, which is the real cornerstone of this programme period.

- A selection of '**common indicators**' linked to the actions of programmes (for ERDF) and also to the specific objectives of programmes in relation to the Youth Employment Initiative (for ESF). These are indicators that programmes have to select from a list provided by the European Commission in an annex to the Fund-specific regulations and submitted to the Commission annually through electronic data transmission.⁶ These common indicators are meant to enable EU-wide aggregation of selected types of outputs across programmes and Member States.



Programme authorities must also prepare **annual reports** whose mandatory content is specified in art. 50 and art. 111 CPR.

2.3 Performance framework obligations

The 2014-20 programme period has also seen the introduction of a new performance framework, whereby programmes are required to establish measurable milestones (to be achieved by 2018) and targets (to be achieved by the end of the programme period, in 2023), linked to the allocation of a performance reserve of between 5 and 7 percent (articles 18-20 and Annex II CPR).

⁶ Art. 6 ERDF regulation; art. 5 of the ESF regulation, whereas 13 of the Cohesion Fund regulation, and Annex I in all three regulations and Annex II of the ESF regulation.

2.4 Evaluation obligations

The main tasks in relation to evaluation during the programme period (art. 54 and art. 56 CPR)⁷ include the following.

- **Drafting Evaluation Plans** (for each programme or for more than one programme), to be submitted for approval to the Programme Monitoring Committee (PMC) within one year from the adoption of the programme (arts. 114.1 and 110.2 CPR). Content and suggested structure of the Evaluation Plans is illustrated in detail in the European Commission's 'Guidance Document on Evaluation Plans'.⁸ The Plans can be reviewed and amended by the PMG during the entire lifecycle of programmes.
- **Implementing the evaluations foreseen in the Evaluation Plans.** These evaluations can be of at least two types:
 - *implementation evaluations*, focused on how a programme is being managed and delivered;
 - *impact evaluations*, assessing the effectiveness, efficiency and impact of programmes. These impact evaluations should be carried out so as to ensure that 'at least once during the programming period, an evaluation shall assess how support from the ESI Funds has contributed to the objectives of each priority' (art. 56.3 CPR).
- **Discussing evaluation results within the Programme Monitoring Committee** and drawing and implementing the relevant recommendations (art. 56.3 CPR).
- **Disseminating evaluation reports** via programme websites, through the website of the European Commission (to whom programme Managing Authorities have the obligation to send all evaluation reports realised) and, potentially, through other means (art. 54.4 and art. 56.3 CPR).

2.5 A new approach to evaluation

The emphasis on impact evaluation in the regulations and guidance document for the 2014-20 period is a key innovation compared to previous programme periods:

'To date Cohesion Policy evaluations have tended to focus more on implementation issues than capturing the effects of interventions. *For the 2014+ period, the Commission wishes to redress this balance and encourage more evaluations at EU, national and regional level, which explore the impact of Cohesion Policy interventions on the well-being of citizens, be it economic, social or environmental or a combination of the three.* This is an essential element of the strengthened result-focus of the policy.'⁹ (emphasis added)

⁷ Ex ante evaluation is left out of this present note as it falls beyond the scope of the thematic paper.

⁸ European Commission (2015) *Guidance document on Evaluation Plans. Terms of Reference for Impact Evaluations. Guidance on Quality Management of External Evaluation*, February 2015 available online at http://ec.europa.eu/regional_policy/sources/docoffic/2014/working/evaluation_plan_guidance_en.pdf (last downloaded 3 March 2015).

⁹ European Commission (2014) *op. cit.*, p. 8.

The European Commission's guidance note on monitoring and evaluation mentions two main approaches for impact evaluation:

- theory-based impact evaluation, and
- counterfactual impact evaluation.

The emphasis on theory-based impact evaluation is new within Cohesion policy, given that counterfactual evaluations were also undertaken in previous programme periods. Traditionally, the evaluation of the impact of Cohesion policy programmes has been carried out with a view to establishing the achievements realised and the effectiveness of programmes (whether goals and targets were met). In other words, evaluation has been traditionally 'goal-oriented', with programme activities 'evaluated on the bases of whether they help[ed] achieve formally stated goals' whose meaningfulness was 'taken for granted'.¹⁰

Goal-oriented approaches to evaluation, also beyond Cohesion policy, have generally been characterised by a summative function (in line with the ideological/cultural setting from which they emerged, the New Public Management agenda). Their main aim has been to draw conclusions about the achievements of a programme, i.e. what it has delivered. However, it has often had more limited value in terms of learning and lesson-drawing.¹¹ Furthermore, in their practical application, goal-driven evaluations have been frequently hampered by a neglect of causality. They have also sometimes led to goal displacement and to what van Thiel and Leeuw (2002) have called a 'performance paradox' (where outcomes have become goals in their own rights, and a focus on outcomes leads to 'everything but an improvement in performance').¹²

By contrast, **theory-based approaches** to evaluation have a more marked formative ambition. They focus 'how and why the program works' (Weiss 1998, 60). It is this formative function that, according to the European Commission, makes theory-based impact evaluation particularly useful for Cohesion policy. With its emphasis on the 'mechanisms' of change, theory-based evaluation 'pinpoints at which link in a larger causal chain repair work is needed',¹³ providing room for the learning that the European Commission views as 'an overarching objective of all evaluations'.¹⁴ A further important contribution of theory-based evaluation is that it can help reveal important assumptions about a programme that may have been implicit,¹⁵ and in this way it can support the understanding of reasons for under-performance, where this is the case.

¹⁰ Dahler-Larsen P (2007) 'Evaluation and Public Management', in Ferlie E, Lynn L E and Pollitt C (2007) *The Oxford Handbook of Public Management*, p. 624.

¹¹ *Ibidem*, p. 626.

¹² *Ibidem*, p. 627.

¹³ Dahler-Larsen (2007) *op. cit.*, p. 629.

¹⁴ European Commission (2014) *op. cit.*, p. 6.

¹⁵ Marra M (2015) 'Cooperating for a more egalitarian society: Complexity theory to evaluate gender equity', *Evaluation*, 21.1, pp. 32-46.

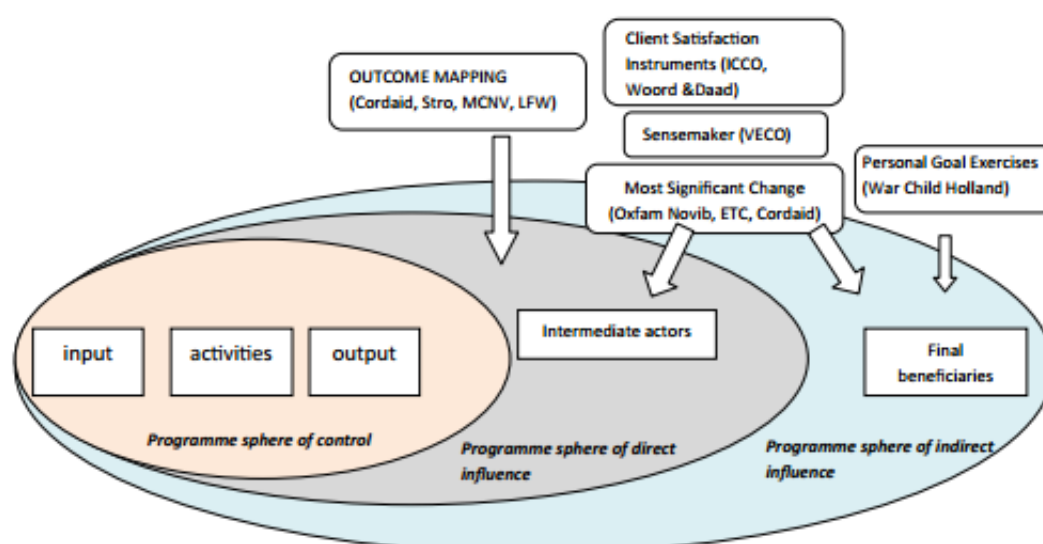
Box 1: Theory-based evaluation in summary

A useful definition of the key characteristics of theory-based evaluation has been proposed by Leew (2012), one of the main advocates of this approach:

- TBE focuses on assumptions called the program/intervention theory;
- TBE has two vital components: the first is conceptual; the second empirical. Conceptually, theory-based evaluations explicate a program theory or model. Empirically, theory-based evaluations seek to investigate how programs cause intended or observed outcomes;
- Central elements of these theories are mechanisms and their links with context and outcomes (Pawson, 2006; Pawson and Tilley, 1997). Mechanisms are the 'cogs and wheels that have brought the relationship [between policies and outcomes] into existence' (Elster, 1989: 3; 2007: 36). Hedström (2005) defines a mechanism as 'a constellation of entities and activities that are linked to one another in such a way that they regularly bring about a particular type of outcome';
- TBE applies different sets of methods in order to find ('reconstruct') and test assumptions and/or to improve them (Leeuw, 2003, in press; McLaughlin and Jordan, 2004).¹⁶

These methods typically comprise literature review and documental analysis, analysis of administrative data, interviews, case studies, surveys and facilitated discussions with stakeholders, with a view to reconstructing the programme logic through a process that enables triangulation. Specific techniques are also used, e.g. contribution analysis and general elimination methodology, to disentangle the effects of the intervention/programme under observation from wider, exogenous developments. However, theory-based evaluation is not able to fully disentangle the outcome of a programme or intervention from the wider developments that take place beyond and beside the policy intervention. One of its advantages, according to its advocates, is exactly its ability to bring such wider developments to the fore as contributing factors (what Mayne, 2012, has called the 'causal package' i.e. the interventions, plus other supporting or confounding factors, and the relationship amongst them).¹⁷

Figure 3: Further example of a logical framework



¹⁶ Leew F (2012) 'Linking theory-based evaluation and contribution analysis: Three problems and a few solutions', *Evaluation*, 18.3, pp. 348-63.

¹⁷ Mayne J (2012) Contribution analysis: Coming of age?, *Evaluation*, 18(3), 270-280, p. 276, quoted in Nacrošis V (2014) Theory-based evaluation of capacity-building interventions, *Evaluation*, 20(1), 134-50.

Van Ongevalle and Huyse H (2014) 'Dealing with complexity through actor-focused planning, monitoring and evaluation (PME)', *Evaluation*, 20.4, 447-66, p. 454.

The second approach, **counterfactual impact evaluation (CIE)**, has been utilised for many years = although not widely - within Cohesion policy. It is an approach that appraises the impact of an intervention by establishing the situation that would have occurred had the intervention not taken place. This is done by comparing a group of subjects exposed to an intervention with a comparison group which was not (e.g. a group of firms that received support from a certain scheme, with a group of firms that did not, or a group of trained individuals with a group of similar individuals who have not been trained) or examining beneficiaries of an intervention before and after exposure.¹⁸ There are various techniques for counterfactual evaluation (illustrated in the Commission's guidance document on monitoring and evaluation in 2014-20) and different methods that can be followed for the selection of control groups. This approach is recommended for some types of impact evaluation, notably 'relatively homogeneous interventions with a high number of beneficiaries'. A key feature of counterfactual evaluation is that it aims to appraise what has changed as a result of the intervention, but not the causality of this, i.e. the how and why questions. The formative potential of this approach is therefore considered to be more limited in scope than in TBE (and qualitative or mixed-methods case study research).

Implementing both types of impact evaluation can be challenging. Evaluation methods need to be tailored to the evaluation questions to be answered. This requires programme managers to be able to perform a balancing act between: (i) selecting the most meaningful evaluation questions; (ii) adopting the most suitable evaluation approaches, which, combined together, can answer those questions; (iii) applying the evaluation approaches in a way that is consistent not just with the goals of the evaluation exercises but also with the resources and capacities available; and, lastly, (iv) deal with the processes of managing the evaluations and their follow-up, processes that are rarely linear and straightforward.

This IQ-Net paper aims to provide information on these topics in order to support fruitful discussions amongst IQ-Net partners during the forthcoming IQ-Net meeting on the challenges related to the evaluation of 2014-20 programmes and the lessons that can be learnt from the past.

¹⁸ Evalsed: The resource for the evaluation of Socio-Economic Development, September 2013, p. 97.

3. NEW MONITORING ARRANGEMENTS IN IQ-NET PROGRAMMES

3.1 Main changes to monitoring systems

The research conducted for this paper indicates that, in many cases, IQ-Net partners will continue to operate the monitoring systems used in 2007-13. In some programmes, these are regarded as requiring only **minor adjustments** to suit the new regulatory requirements and e-Cohesion agenda (e.g. Austria, Vlaanderen, Wales) or link up the monitoring of financial data with the domestic accounting system (in Slovenia, so that data are uploaded only once). In other cases, more **substantial revisions** are required to meet the new regulatory requirements and improve the functioning of systems, such as simplification of monitoring for project beneficiaries and Managing Authorities (e.g. Finland), addressing weaknesses that emerged during 2007-13 (e.g. in Denmark, Finland, Pomorskie, Portugal, Spain), and harmonisation of monitoring across programmes, funds and with domestic policies (e.g. in France). There is also a group of countries or programmes where **wholly new monitoring systems** are being introduced, as in France, Greece, England and Scotland or the previous system has undergone a major overhaul (Czech Republic). A summary of the different approaches to the 2014-20 monitoring systems is provided in Table 1 below.

Table 1: 2014-20 monitoring systems

No change except for the adjustments necessary to fulfil new regulatory requirements	Improvement on previous systems	New monitoring systems or major overhauls of previous ones
Austria (ATMOS) Belgium Vlaanderen Slovenia Wales	Denmark Finland Poland, Pomorskie Portugal Spain	Czech Republic France Germany, Nordrhein-Westfalen Greece England Scotland

Source: IQ-Net country research.

3.1.1 Improvement of indicators

Among the countries that have made changes to existing monitoring systems, Denmark, Portugal and Spain have placed particular emphasis on the improvement of the design of monitoring indicators and the way they are interpreted and tracked by project applicants and implementing authorities.

- In **Denmark**, this was in response to the fact that in 2007-13 indicators were not always selected with the necessary reflection and targets were often overly ambitious. These efforts have entailed work by the Danish Business Authority (DBA) to select more appropriate indicators, the production of 'indicator guidelines' for applicants, sensitisation work aimed at project applicants by the growth fora secretariats on how to select and interpret indicators and how to establish realistic target values, and enhanced checks ex ante, whereby the DBA can send applications back to applicants where the guidelines have not been adequately followed. Project applicants are also required to describe the chain of effects, which is intended to provide a good basis for evaluation. According to the DBA, the monitoring system now features 'significantly improved indicators'.

- In **Portugal**, the main driver for developing a more comprehensive and robust monitoring system of indicators was the need to gear monitoring (and management) towards results. Major efforts were focused on developing indicators, because this was perceived to be the main challenge of the new results-oriented programming logic. To this aim, the Portuguese authorities asked the OECD to carry out a full and independent analysis of the indicators included in the Portuguese OPs, including their appropriateness for the needs of implementation (e.g. for project selection), monitoring ('scoreboard' indicators) and evaluation (results indicators). The review concluded that the system of indicators was overall of very good quality, but also provided practical recommendations regarding specific weaknesses (e.g. the fact that a few indicators labelled as 'outcomes' measured outputs, some heterogeneity across programmes, the 'narrowness' of some indicators, and others). It also identified areas for future work, for example the recommendation to take advantage of the work done and extend the indicators to the wider framework of domestic public policies, that the results-orientation is maintained during the implementation phase through the project selection process and the design of contracts with beneficiaries, and that a framework for monitoring and evaluation is set up to enable policy-learning and the communication of achievements to the broader public.
- Similar reasons informed the work done to improve the monitoring indicators in **Spain**, where the role of the monitoring system will be reinforced to ensure that the programmes are able to deliver the new results-orientation and comply with the performance framework requirements. The strengthening of the monitoring system engaged the national officials in the DG for EU funds (Ministry of Finance and Public Administration) for over a year, resulting in a number of changes to the monitoring system in use in 2007-13, aimed primarily at ensuring that achievements across all ERDF programmes, also across OPs, can be measured and that the programmes' contribution can be identified. The main changes entail:
 - **Streamlining the number of indicators across OPs** – in 2007-13 there were too many indicators from the different intermediate bodies and that some indicators were too general and not relevant.
 - **Standardisation of the interpretation of the indicators and their calculation** to enable aggregation of data across programmes.
 - **Enhanced accountability, through the assignment of management responsibilities for indicators** to ensure that the key requirements are addressed and that the indicator values entered into the system are reliable and accurate.
 - **Synchronisation between financial implementation and physical performance.**
 - **Obligation for programme managers to update physical indicator values at the same time as spending certification.**

3.1.2 More systemic reorganisation

In **Finland**, on the other hand, the improvements introduced to the Cohesion policy monitoring system (now renamed EURA 2014) were more systemic and related to a number of perceived needs. First, better links between Cohesion policy and domestic business aids monitoring systems were developed (TUKI2014), to avoid duplication of effort and the risk of inaccuracies/mismatches. Second, e-Cohesion was embraced fully, to make the monitoring system an integrated platform for all communication and documentation exchange between applicants/beneficiaries and the MA. Third, the monitoring platform was utilised in order to reduce error rates and improve the efficiency of implementing bodies in managing applications. The key novelties introduced, compared to the previous EURA 2007 system, include the following.

- **Full electronic provision** – EURA 2014 is implemented electronically. All data-processing (application and other forms) is done electronically. There is no need to sign forms and send them by post (as previously). This is intended to allow for significantly faster processing of applications, particularly where additional information or corrections are needed.
- **No separate monitoring forms** – Monitoring information is provided together with the payment application. Hence, there are no longer separate monitoring forms, and this reduces the risks of delays or inaccuracies.
- **No mid-term reporting for projects** – Mid-term reporting is no longer required. Instead, description of implementation progress is provided as part of the periodic monitoring information.
- **Provision of key information in English** – The name of the project and a summary are also provided in English, subsequently published through the Structural Funds information services for wider, international dissemination.

EURA 2014 enables the delivery and processing of project applications, payment applications, monitoring forms and final reports as well as the receipt and preparation of project decisions. All applications, except ERDF business aids and ERDF applications addressed to Tekes (the Finnish Funding Agency for Innovation), are submitted to EURA 2014. The new monitoring system EURA 2014 will thus not only have increased functionality, as it will effectively serve as a management as well as a monitoring tool, but also match more adequately the changed organisational arrangements that support the delivery of the programme, whilst at the same time reducing the number of errors to the minimum.

3.1.3 Introduction of new systems

In other countries, wholly new monitoring systems are being introduced. This is the case in the Czech Republic, France, Greece, England and Scotland.

In the **Czech Republic**, a new monitoring system was designed and tested during the three years prior to the programme's finalisation, and a final version was being tested by the MAs in March 2014 (when IQ-Net fieldwork took place). The new system is considered to be rather complex and quite cumbersome for programme managers (given its comprehensive nature – discussed below), but it fulfils the requirements of the new performance framework, and it is considered to be an improvement on past practice particularly from the point of view of applicants and beneficiaries. The main characteristics of the new system are:

- **a single system for all Operational Programmes and all ESI Funds**, which means less onus for project beneficiaries (applicants/beneficiaries have to learn the system only once and can use it for all other potential applications), more harmonisation across funds, and more oversight/summative potential for the purposes of reporting and evaluation;
- **fully integrated electronic platform** for data transfer across actors and levels of implementation, **effectively serving as an e-management tool** through which applications and monitoring reports which be submitted to the MA solely electronically (whereas until now they were printed and submitted physically);
- **interoperability with other platforms**, i.e. some parts of an application form can be copied to another application, or information relevant to applications can be downloaded directly into the forms from applicable Czech registers or databases; and
- tracking of the **information required by the new performance framework**, such as results indicators and milestones.

The drawback is that combining all funds (ERDF, ESF, CF as well as EFF and EAFRD) and all the information required by the related regulations (thematic objectives, performance framework, categories of regions etc.) has made the system quite cumbersome for programme administrators.

In **France**, the joint effect of the changes in the requirements contained in the ESI Funds regulations and the decentralisation of ERDF management to the reformed regions (and of ESF, to a much more limited degree) has led to the introduction of a new information system for the ERDF called **SYNERGIE**. This new system will replace the former Présage. In addition to tracking the indicators required by the new regulatory framework, the new system will allow the electronic administration of projects throughout the full project cycle, from submission of an application to closure (this, however, will be optional rather than compulsory). It will also communicate with other information systems and databases. For the moment, the implementation of the new system is uneven across regions, the ESF and EAFRD will run their own systems, for the regions the integration in the new platform is not compulsory and some will continue to run their own monitoring systems. However, the ambition is eventually to create a common platform for the monitoring and implementation of the French PA and OPs.

New online monitoring systems are also being introduced in Nordrhein-Westfalen, England and Scotland. In **Nordrhein-Westfalen**, a new IT platform, revised monitoring forms (much more streamlined, shorter and simpler to fill in) and new procedures for entering data into the system are being introduced to simplify monitoring requirements for programme beneficiaries, reduce the length of time necessary to input information into the system, and devolve the responsibility for uploading monitoring information in the IT system to implementing bodies (taking advantage of a major streamlining of the delivery arrangements of the programme, i.e. the reduction of the number of intermediate bodies from the previous 108 to the current nine). Whilst in 2007-13 monitoring data was entered into the monitoring system by the Managing Authority, in 2014-20 the Implementing Bodies will upload data directly into the system, which should allow the Managing Authority to focus more on strategic management.

In **England**, the new online monitoring system has been devised to embrace the e-Cohesion agenda as well as to unify all monitoring data in a single platform across England, in line with the rationalisation of Cohesion policy support in a single England-wide (national) programme (against the 10 regional ERDF programmes of 2007-13). Whereas there were previously some differences between the English regions in indicator definition (where in addition to the standard core indicators the regional programmes also added their own indicators), the system will now track the same indicators across all regions. Indicator definitions will also be standardised, although the process of coming up with indicator definitions is proving to be quite problematic.

A new monitoring system is also being rolled out in **Scotland**, particularly with a view to obtaining better insights on what is being achieved through the programmes than had been possible in 2007-13. The previous monitoring system focused mostly on financial progress and is not considered to have been adequate to track outputs and results, particularly because of the lack of satisfactory quality control procedures on the data submitted by projects, which was not checked and verified on an ongoing basis. The Managing Authority is thus not satisfied with the reporting of results and outcomes during 2007-13 and is now investing extensive effort to ensure that the new monitoring system allows the appraisal of what operations are achieving throughout implementation and that the data submitted by projects are accurate, through checks undertaken prior to payments being made to the projects. The anticipated move to simplified costs should facilitate moving monitoring away from purely financial aspects to focus instead on achievements, but the MA is conscious that this may be challenging given the need for a change in the mind-set of all actors involved, since in 2007-13 target-setting was often poor and reporting on physical indicators was neither followed up nor verified.

Lastly, the monitoring system in **Greece** also received a major overhaul, in this case to address the criticisms levied about the efficacy and accuracy of the system in place during 2007-13, particularly in relation to the quality of indicator data and major gaps or weaknesses in the monitoring of key fields, such as R&D and ICT. The Greek authorities thus upgraded the existing Integrated Information System with better-defined and more customised indicators, guidelines for field completion and compatibility with SFC2014 and electronic exchange protocols, and an improved role for beneficiaries in providing data, linked with efforts to enhance the administrative capacity of all stakeholders (MAs, IBs, beneficiaries), on measuring and tracking indicators, through training.

3.1.4 The rationales for the changes introduced

Whether through changes introduced to previous monitoring systems, or via the creation of new systems altogether, the reforms of monitoring systems respond to a series of common aims and rationales, beyond the need to respond to the regulatory requirements, namely:

- **simplification and reduction of the onus upon project beneficiaries and/or programme administrators**, also with a view to improving data quality, efficiency and timeliness;
- **streamlining, harmonisation and interoperability of monitoring platforms across programmes, funds and between EU and domestic systems**, to allow aggregation of information and the elaboration of summative views on achievements beyond the single fund or programme, and to provide evaluation with the data that are necessary to isolate, insofar as possible, a programme's contribution;

- **increased linkages or full-out integration with management and accounting systems**, to ensure the timeliness and comprehensiveness of monitoring submissions (e.g. by linking the inputting of data and/or the satisfactory pass of checks on data quality to certification of expenditure and payment, as in Scotland), faster processing of projects and reduction of error rates;
- **introduction of enhanced checks and accountability**, to increase the quality of data entered into the systems, also with a view to adequately accommodating the needs of evaluation;
- in some cases, **sharing the load**, i.e. direct inputting by beneficiaries or implementing bodies of data into the system, in order to improve ownership, quality and management efficiency.

The new or improved monitoring systems are intended to enable ongoing learning about the performance of a programme, fund, intervention or project; to provide MAs with timely information to allow them to introduce adjustments to the programmes where needed; to serve the data needs of evaluation and, in a number of cases, to act as integrated management platforms for projects and programme management alike. These are ambitious goals that are linked to the new results-orientation of programmes, but they go beyond the regulatory requirements illustrated in Section 2.2.

3.2 Open challenges

A number of open challenges or unsolved issues in relation to the design and operation of the new or revised monitoring systems remain. As in the previous programming cycles, a period of adaptation and adjustment will be necessary to iron out unforeseen weaknesses or address new issues that might emerge as implementation progresses. The main issues relate to the rolling-out of new systems and procedures, to ensuring that the data tracked by the monitoring systems is accurate and reliable, and to the treatment of sensitive data, particularly in relation to ESF. The selection of results indicators and the setting of baselines and target values were also problematic in some cases (discussed in the next section).

Rolling out of new systems. In **Finland**, the EURA 2014 monitoring system was launched on 5 May 2014, although various components are still in the process of being finalised. Even though the system builds on the experience of implementing the previous monitoring system, EURA 2007, the many new features (illustrated in Section 3.1.2) have meant that the launch has not taken place without problems. The exceptionally large (simultaneous) user group at the end of August 2014 resulted in specific technical problems. For instance, the completion of applications and submission of applications to the authorities did not work in all cases, and in some applications information had gone missing. There are also capability issues, and a training programme has been organised for the key users of the system (e.g. the regional councils and intermediate bodies, such as the ELY-Centres), particularly on the new components introduced. There have been both monitoring-specific training events and general training events that also covered issues relating to EURA 2014. Five training events aimed at intermediate bodies were held between May and September 2014 in Helsinki, Joensuu, Tampere and Oulu, involving more than 350 people, and more are in the pipeline.

The problems with the new monitoring system in the **Czech Republic** lie in its complexity. Despite efforts to simplify and reduce the administrative burden for all involved, finding a balance between comprehensiveness and manageability is proving to be an issue. Combining all the ESI Funds and all the regulatory requirements into a single system – that includes thematic objectives, performance framework, categories of regions etc., and the fact that the requirements for different funds can vary – has rendered the new system particularly complex and increased the burden on programme administrators.

Quality control. Data quality was a concern in a number of cases in 2007-13, and the procedures to control the quality of the values inputted for output and result indicators are still considered to be in need of improvement in a number of cases.

- In **Vlaanderen**, for example, the MA considers that the monitoring system is on the whole effective in terms of data collection; however, it is now introducing more systematic controls on the data submitted by projects in their four-monthly reports, which were not always accurate in the past, with problems often emerging at project closure.
- The quality of physical monitoring data has been a particularly contentious point in **Scotland** during 2007-13. The Managing Authority was not satisfied with the reporting of outputs and results during 2007-13, which it viewed as highly inaccurate, and it is now pursuing an improvement in the quality of data reported by making checks as a precondition for payment.
- In **Denmark** too, it was felt that projects often selected indicators almost randomly, as well as setting targets that were too high. There is thus a degree of realism regarding the degree and pace at which improvements will be made, recognising that indicators and target values may be adjusted during the life-cycle of projects.
- In **Portugal**, agreements with public bodies outside the MAs are being explored as a cross-check for data gathered through the monitoring systems.

Sensitive data and privacy legislation. The procedures in place to deal with privacy legislation and dealing with sensitive data in relation to project participants has emerged as an issue in **Wales**, where project holders are not always forthcoming in handing out information about firms financial accounts or have concerns about collecting personal information about participants (e.g. where there is ESF support for particular groups such as people dealing with substance misuse problems) or for which consent has to be provided by the involved party. The sensitivity of this data then has implications for the whole data set.. Although there is now a legal basis for collecting participant-level data, there are still problems related to data protection issues. For instance, projects can support participants who do not provide data on data protection grounds (data provision does not determine eligibility), but then cannot claim for them (as one cannot claim for a participant unless the monitoring system includes the complete set of data). This can lead to a discrepancy in reporting, where the AIR can report a higher figure than the monitoring system, which can undermine the monitoring system and create problems in relation to financial claims and programme expenditure. The treatment of sensitive data is also an issue in the **Czech Republic**. A new group of laws, set to enter into force on 1 July, is imposing some changes on the way such data can be collected and treated. This will require changes to the way ESF

data on persons is collected and the introduction of a new system at the end of 2015 whereby data on project recipients, before and after the realisation of a project, will be entered directly by beneficiaries.

3.3 Identifying results indicators

The selection of indicators, particularly of results indicators, appears to have been largely straightforward (or at least without major problems) for most IQ-Net partners. However, for over a third of IQ-Net partners the selection of results indicators and identification of baselines and targets has entailed difficulties relating mainly to:

- **appropriate definition of objectives and intervention logic** (e.g. in Greece and Slovenia);
- **capacity issues**, i.e. the difficulty for actors more accustomed to dealing with output and financial indicators to deal with a new type of indicator (e.g. in France);
- **uniform interpretation of indicators** across different actors, intermediate bodies and applicants (e.g. in Finland, England – examples of interpretation difficulties are illustrated in Box 2 to follow);
- **measurement of impact** (and thus indication of target values) for programmes whose funding represents only a small portion of public policy in the region and which can thus only contribute to a certain trajectory of change, rather than lead to a quantifiable impact (e.g. Austria, Nordrhein-Westfalen, Wales); and
- **different approaches to baseline and target-setting by different Commission DGs** (a problem for multi-fund OPs, e.g. the Czech Republic).

An overview of the experience of IQ-Net partners with the identification and quantification of results indicators is provided in

Table 2 below.

Box 2: Examples of doubts regarding indicator interpretation

In Finland, discussions on the indicators have been held in the coordination group comprising representatives of the regional councils of Länsi-Suomi. The key message is that there is a need for guidance on the interpretation of the indicators, as it is important that the project applicant and the authorities have a common understanding on the content of the indicators when the funding decision is made.

- New research or R&D jobs
 - What is the difference between research and R&D jobs?
- New innovation platforms and other innovation openings
 - How are innovation platforms and innovation openings defined? Examples would be helpful.
 - How long can an innovation platform be new, particularly in a context where there will be a follow-up project?
- Firms which as a result of support produce a new or a developed project or a material which promotes low carbon
 - Does this need to take place during project implementation or soon after?
- Solutions of communities which promote low carbon
 - What is included as communities here?
 - What could be counted as promotion?
- Demonstrations implemented in firms which promote low carbon
 - What is the difference between a demonstration and a pilot?
- Firms participating in projects led by R&D institutes
 - What is counted as participation?
- Saved energy
 - It is difficult for an authority to take a position on how realistic the figures given by the applicant are.
- Knowledge- intensive start-up firms, which have significant preparedness for the development of a new product, service or production methods as a result of the support
 - Several indicators use the term 'significant' (i.e. significant increase, significant improvement), but what does significant mean?

Source: IQ-Net country research – Finland.

Table 2: Assessment of identification of core and results indicators and related target values

Assessment	Country/region
No particular problems or minor issues	<p>Vlaanderen, Scotland, Slovakia – no particular problems reported with the identification of indicators.</p> <p>Denmark – no particular problems with identifying indicators. For results indicators, one has been identified for each priority.</p> <p>Germany (NRW) – there were discussions between the NRW MA and the Commission on output and result indicators during the negotiations. Issues related to the scope of indicators, i.e. whether they should aim to measure the impact of the programme on the Land as a whole (what Commission wanted and NRW resisted). However, these issues were resolved and are not seen to have been significant. Programme authorities agreed to replace some of their original indicators with indicators proposed by the Commission, and they feel that the indicators agreed are acceptable.</p> <p>Portugal – the new logic has been internalised by Portuguese authorities since the outset, involving a concerted effort to choose indicators. In some fields, it was more difficult, but overall challenges were overcome.</p> <p>Spain – overall, the indicator framework is considered to be more robust than in the previous period. An online forum has been created for the different implementing bodies to identify any challenges with indicators (measurement, interpretation etc.) and to resolve them. No major problems so far.</p>
Significant challenges	<p>Austria – identifying results indicators has been very challenging. In the light of the relatively low level of funding, Austrian authorities did not feel that they could identify the programme's impact (e.g. additional jobs in public sector research), but rather indicate the contribution of the programme/trajectory of change.</p> <p>Czech Republic, Pomorskie – the different understandings by different DGs within the European Commission of how the baseline values of context indicators should be populated and their scope (i.e. whether they should be linked to target values and come from statistical sources and thus linked with the interventions only indirectly) created problems for multi-fund OPs.</p> <p>England – harmonising indicator definitions across England has not been straightforward.</p> <p>Finland – monitoring indicators a key concern in Finland, particularly with regard to the uniform interpretation of indicators across project applicants and funding authorities. There remain doubts and inconsistencies for which guidance is needed (see Box 1 for examples).</p> <p>France – identifying result indicators, especially in certain fields where none had existed before, and setting their accurate initial, intermediate and final target values have been very challenging. The results-orientation of 2014-20 has led MAs to work on some of these indicators and on the logic framework for the first time. Annual Activity Reports already existed but reporting was mainly focused on financial and output indicators.</p> <p>Greece – the identification of results indicators was challenging due to the need to improve the intervention logic through clearer identification of specific objectives. The ex ante conditionality on statistical systems and results indicators has only partially been fulfilled and the OPs contain the required Action Plan.</p> <p>Slovenia – serious problems with the identification of results indicators. Objectives have had to be reformulated and target values re-set.</p> <p>Wales – a positive experience was reported with regard to the core indicators: the Commission's core indicator definitions were not in place at the start of the 2007-13 period, so the definitions provided have been more useful this time, and the Commission's enhanced guidance has been considered helpful. However, identifying results indicators has been challenging, especially for the ERDF. The difficulties related to trying to find indicators that the programmes have a realistic chance of affecting and identifying appropriate targets.</p>

Source: IQ-Net country research.

Amongst the group of IQ-Net partners for whom the identification of results indicators has been challenging is **Wales**. Issues have emerged particularly on the ERDF side, specifically in relation to identification of targets and trying to find indicators that programmes have a realistic chance of affecting. However, the work done in these areas is considered in some ways to have had a positive effect, in terms of making the programmes more focused. The Welsh shadow PMC for 2014-20 set up a sub-group (Measuring Success) to consider alternative approaches to the measurement of success and impact within the ERDF, ESF and RDP programmes. The remit of the group was to consider what were the most important transformational changes that should result from the 2014-20 programmes and how the MA was going to measure the degree of success in achieving them. The key conclusions and recommendations of the group are illustrated in the box below.

Box 3: Results-orientation in Wales. Monitoring recommendations by the ‘Measuring Success’ sub-group of the PMC

The main area where beneficial economic change would be observed would be within the business community, and impact measures typically used within the business environment could be appropriate measures.

Key economic issues within Wales are issues of productivity, as opposed to employment, and impact measures selected should reflect this.

Many businesses see their social and environmental impact as a key element of economic performance, and social/environmental parameters could sit comfortably alongside economic parameters.

It was impossible for the sub-group to recommend any specific impact measures to the PMC as there is not enough historic longitudinal data on business-related measures to allow a retrospective baseline to be constructed to allow progress to be measured.

There should be support for ongoing work on improving data on business need, examining options for the creation of a business-centric model of economic development. Specifically, this could involve development of databases more typically seen in private than public sector (e.g. for supermarket loyalty cards or financial comparison websites), designed to continually update the type of information they hold and over time becoming valuable monitoring tools.

Source: WEFO Papers to the February 2015 PMC,
<http://gov.wales/docs/wefo/publications/150415papersfebruary2015meeting.pdf>.

3.4 Enhanced cooperation with National Statistical Offices and additional data-gathering arrangements

IQ-Net partner views on whether the data gathered by the monitoring systems will be adequate to fulfil the needs of evaluations are mixed, as is the confidence in the quality of the new monitoring systems. On the whole, however, IQ-Net partners appear positive regarding the fact that the new monitoring systems and procedures put in place for the 2014-20 period will represent an improvement compared to the past and that the new results-orientation will be useful in terms of allowing programmes to gain a better overview of what is being achieved.

A trend, across a number of IQ-Net partners, is the increased use of administrative data and strengthened cooperation with national Statistical Offices (e.g. in the **Czech Republic, Denmark, Greece, Portugal, Slovenia, Spain** and others). This applies particularly to results indicators and to the data on individuals and firms, which could be used to allow counterfactual impact evaluation of projects and interventions.

In some cases, investments are being planned to strengthen national statistical databases and/or undertake additional data collection exercises, tailor-made to the information needs of programmes, through ad hoc surveys. In **Greece**, for instance, there are concerns about the availability of administrative data. Particular issues relate to the timely publication of statistical data from the national statistics office (EL.STAT), the shortage of data at regional and sub-regional levels, and the statistical validation of data sources. One of the proposals currently under consideration is the conclusion of a cooperation agreement with the bodies providing statistical and administrative data, including EL.STAT, possibly to be funded with Technical Assistance resources.

Further, additional data-collection exercises, for example through survey research, are anticipated or have been undertaken in a number of countries, including **Austria**, **Slovenia** and **France**,

- In **Austria**, ÖROK commissioned factsheets in 2013, one for each of the nine *Länder* and a national summary, to acquire baseline data specifically geared to the OP. These factsheets have been produced by the Joanneum Research Policies Institute, using a variety of sources, including data from the Austrian Statistics Office, labour market data and data produced at the European level, as well as own surveys and calculations. The data collection exercise will be repeated when needed (e.g. for the Strategic Reports in 2017 and 2019) and a budget has been set aside for this. It is considered that this exercise will play an essential role in any future evaluation activity to appraise the impact of the programme.
- In **Slovenia**, the MA has decided to draft its evaluation plan as soon as possible and at the same time to convince those Slovenian authorities dealing with the data relevant to evaluation to identify missing data and how to collect missing data. This applies especially to the national Statistical Office of Slovenia and the Institute of Macroeconomic Analysis and Development (*Urad RS Slovenije za makroekonomske analize in razvoj*). Additional data-gathering exercises will be funded, especially surveys of the general public, though no specific budget has been earmarked for this as yet (beyond the budget earmarked for evaluation activities).
- In **France**, three regions (Corsica, Languedoc-Roussillon and Midi-Pyrenees) intend to use data sources that are additional to the available administrative and monitoring data, such as surveys, field research and studies. It should be noted that the decisions by these French regions do not appear to be correlated with the financial scale of programmes or their overall evaluation budgets (Corsica, for example, is the smallest French programme).

In **Finland**, **Nordrhein-Westfalen**, **Vlaanderen** and the three **UK** nations (England, Scotland, Wales), on the other hand, there are no plans at present to gather additional information beyond what is generated by the monitoring system. In Nordrhein-Westfalen, the expectation is that, in general, the monitoring systems should generate the information needed for programme management and evaluation, and there is reluctance to impose further demands on project beneficiaries. Thus, if it proves necessary to collect additional information, strong efforts will be made to ensure that this is minimal (though the Nordrhein-Westfalen MA has noted that, in the past, the most useful additional information has been qualitative and collected via interviews – which have allowed for in-depth understanding, for example of specific difficulties with aspects of implementation).

4. THE 2014-20 EVALUATION PLANS

4.1 State of play

The state of play amongst IQ-Net partners with the elaboration of Evaluation Plans (EPs) is mixed, partly reflecting the varied state of OP approval.

Most EPs are currently still being drafted, with the aim of submitting plans to the Programme Monitoring Committees between June and December 2015. Where progress is not advanced, this has sometimes been due to the necessity to focus more on the strategic drafting and negotiation of the programmes and/or their operational launch (e.g. in **Greece, Vlaanderen, Spain**). For instance in **Spain**, less than half of the OPs have been approved (9/22 OPs) to date. The preparation of the evaluation plan is therefore at a very early stage because, until the negotiations are closed, there is some uncertainty on the typology of actions that will be co-funded. A different approach was taken in the **Czech Republic**, on the other hand, where the preparation of the EPs for the single Operational Programmes is well underway, even though the OPs are still pending approval.

Table 3: State of play with Evaluation Plan preparation

Not started	Initial stages	Under development	Advanced	Approved
Austria (ERDF OP)	Vlaanderen	some Czech OPs (even though OPs not yet approved)	some Czech OPs (even though OPs not yet approved)	Czech PA
Greece	Portugal			
Finland (to be ready autumn 2015)	Slovenia	Pomorskie	Denmark (to be presented to PMC June 2015)	
England (but to be submitted to PMC by mid-June)	Spain		Some French OPs	
	Scotland		Nordrhein-Westfalen (full draft by summer)	
			Slovakia (full draft by 30 June)	
			Wales (full draft ready, to be submitted to PMC in June)	

Source: IQ-Net country research. **Note:** the information relates to the time of fieldwork (March/April 2015).

The Evaluation Plans have different scope, largely reflecting the institutional settings of a country or strategic or operational choices (e.g. to allow exploiting economies of scale or the achievement of whole-country overviews etc.). Evaluation plans can cover:

- **a single programme**, whether single-fund (e.g. in **Nordrhein-Westfalen, England** and in the French regions) or pluri-fund (as in **Pomorskie** and in **Slovenia**, where the EP will cover the multi-Fund OPs for 2014-20, which involve ERDF, ESF and the Cohesion Fund);

- **a single fund across an entire country** – as in **Austria**, where there will be two national EPs, one for ERDF, managed and coordinated by ÖROK, and one for the ESF, under the responsibility of the Federal Ministry of Labour, Social Affairs and Consumer Protection;
- **more than one programme within a region** – e.g. in **Wales**, where a single EP will cover five ESI Funds programmes (two ERDF OPs and two ESF OPs, as well as the rural development plan);
- **a single, national EP for all ESI Funds** – as in **Portugal**, where there will be a single national EP for ERDF, ESF, EAFRD and EMFF co-funded programmes, or a **single, national EP for ERDF and ESF**, as in **Denmark**;
- **an EP for the whole Partnership Agreement, plus individual EPs for the OPs** – as in the **Czech Republic** and **Slovakia**.

In some cases, decisions on the coverage of the EPs are still pending. For instance, in **Spain** it is still being considered whether there would be: one plan per OP, or a national plan for ERDF and one for the ESF, or a joint ESF/ERDF plan. National authorities are leaning towards the national fund-specific option, although this will have to be agreed with the regions in the monitoring and evaluation committee and is thus not yet defined. Part of the reason for this preference concerns the distinctive character of the interventions funded under each Fund, which contribute to separate thematic objectives, as well as with the pragmatic need to keep the plans relatively short. Given the expected increase in the evaluation work under both Funds, a single EP with both funds would be rather lengthy.

Even where evaluation responsibilities are largely devolved to the programme level, and EPs cover single regional programmes, synergy and cooperation across more than one programme may be foreseen. In **France**, for example, the EP of Aquitaine anticipates collaboration with other regions to realise evaluations dealing with impacts on common themes, in order to share costs and improve the comprehensiveness and quality of the evaluations. This approach, which has already been the focus of an experiment in an ex ante evaluation on Financial Instruments, may be partly due to the recent reorganisation of regional boundaries, which is set to come into force in 2016.¹⁹

The EPs are intended to be flexible documents, to allow accommodation of emerging evaluation needs as they emerge and re-calibrating work where deemed necessary. In some cases, annual EPs are also foreseen. For instance, in **Slovenia** the EP is conceived as a 'living document'. It is therefore not very detailed regarding the kind of evaluation that will be done in each year, as the details of each single evaluation in a specific year will be provided in annual evaluation plans, decided annually in dedicated PMC meetings.

In many cases – but not all, e.g. not in Greece – IQ-Net Managing Authorities already had Evaluation Plans in 2000-06 and 2007-13. These experiences were helpful to the MAs in terms of allowing them to draw lessons for the current Plans. For example, in **Wales**, the Evaluation Plan for 2007-13 was helpful, as it focused on both process and impact and made a clear separation between them. For

¹⁹ Due to this territorial reform, Aquitaine, for example, will be merged with two other regions, Limousin and Poitou-Charantes.

2014-20, the MA has thought about other areas that are going to be evaluated beyond these two issues and has drawn a number of lessons from evaluation activity carried out under the 2007-13 Evaluation Plan, summarised in Box 4, below.

Box 4: Lessons from the 2007-13 Evaluation Plan in Wales

The 'research summaries' published by the Welsh Funding Office (WEFO) have been a success. These are summaries of published research that are directly relevant to Structural Funds programmes. The summaries generated 'an accessible and comprehensive evidence base to inform the management of the programmes and development of future programmes'.²⁰ These will be continued in 2014-20 and extended to cover the RDP, and circulated to a wider audience and published online.

There has been successful application of counterfactual impact evaluation methods (see case study) in evaluating the impact of ESF interventions on unemployed and economically inactive individuals, enabling the Managing Authority to be more confident about the impact of ESF. This approach will be continued in 2014-20 and may be extended to ERDF.

The use of ex ante evaluations to build evidence into programme design for 2014-20 has also been considered successful. A useful first step was for evaluators to produce synthesis reports summarising all evidence relevant to programme priorities. These were shared with those drafting the programmes; the success means that the MA will be able to use the existing specifications and contract management arrangements as best practice templates for any future ex ante evaluations.

The introduction of mandatory project-level evaluation in 2007-13 has built evaluation capacity in Wales and contributed to the evidence base. This will remain mandatory in 2014-20.

The regular collection of ESF participant-level data has greatly assisted in the evaluation of ESF programmes.

The parallel data collection system for businesses supported by ERDF encountered more difficulty encouraging ERDF projects to submit business-level data; as a result, evaluation of ERDF was much more limited than evaluation of ESF programmes in 2007-13.

In terms of what has not worked so well, some of the thematic evaluations failed to generate useful evidence to inform programme management. In hindsight, the evaluations were commissioned too early, when insufficient activity had taken place. To avoid this happening again, it is proposed that a full mid-term evaluation will take place once 40 percent of programme funds have been spent (e.g. on a staggered basis, so that evaluation starts when the first priority reaches 40 percent or once every priority has reached this percentage).

The commissioning route for the ongoing evaluation of the RDP will be revised and a specialist Framework Lot for Rural Research will be established to facilitate increased flexibility to undertake ad hoc evaluation.

Source: WEFO draft Monitoring and Evaluation Strategy (March 2015).

Lessons from previous experience with the evaluation plans were also drawn by other IQ-Net partners (e.g. Denmark, England, Portugal) and relate to the need for EPs to be flexible, relevant (focused on needs) and realistic in ambitions. For **Portugal**, for example, the 2007-13 period was a very positive experience, giving evaluation a more prominent role, enhancing coordination among institutions responsible for the implementation/coordination of the Funds, and ensuring a more coherent approach to evaluation. This experience was useful in informing the development of an Evaluation Plan for the 2014-20 programme period, particularly in terms of

- identifying evaluation needs, taking into account the usefulness of the policy for the main stakeholders;
- ensuring that IT systems collect relevant information for evaluation purposes;

²⁰ WEFO draft Monitoring and Evaluation Strategy (March 2015) p. 3.

- coordinating the time to collect additional information for the evaluation process with a more timely availability of statistical data; and crucially
- including in the Evaluation Plan the strategy for communicating the results of evaluation exercises (a theme, that by and large, has not emerged particularly strongly from the IQ-Net country research).

The guidelines on EP drafting provided by the Commission were deemed helpful by various IQ-Net partners (for instance **Denmark, Portugal, Slovenia**).²¹ Slovenian interviewees praised the balance between precision and flexibility in facilitating different approaches and were appreciative of the fact that guidelines are drawn up to help Member States. This view is not unanimous, however. An opposite assessment and a general concern were recorded in Germany. **Nordrhein-Westfalen** and the other German *Länder* are basing their evaluation plans on the text of the Regulations and regard the Commission's guidance document simply as guidance rather than as a binding document. However, the Nordrhein-Westfalen MA fears that the Commission may not in fact accept this approach. Their experience, for example with the communication plan, has been that the Commission has tended to generate additional demands above and beyond what is set out in the regulations.

Some areas of concern were also highlighted in relation to specific issues, often emerging in the Commission services' process of appraisal of the EP drafts. With the **Czech** authorities, for example, the European Commission picked up on the fact that evaluation questions were missing, but Czech authorities consider this information to be too detailed and impossible to provide and plan in detail when the OPs are not yet finalised. Another area of concern relates to the differences in the guidance documents for the ERDF and ESF programmes and the other ESI Funds. This can be challenging for multi-fund OPs. For **Denmark**, a critical issue is the requirement to carry out at least one evaluation on how the Structural Funds have affected macro-indicators. The Structural Funds' support in Denmark comprises just about 0.03 per cent of the Danish GDP and it is therefore more difficult to disentangle impacts on macro-indicators.

Interestingly, in **Poland**, the national coordinating authority has provided additional detailed guidance and a prescriptive template to the Polish regions as they drew up their plans. This is part of a broader approach where the level of funding allocated to regionally-managed programmes has increased in 2014-20 but the regulation of different aspects of management and implementation by national government has remained strong. A problem with this is the additional time required of MAs to respond to these inputs. In **France**, CGET issued a guide for monitoring, management and control of the OPs (which was essentially a shortened, translated version of the guidance document produced by the European Commission on 22 December); whilst in **Slovakia**, the Central Coordinating Authority is developing guidance on how to prepare a good quality EP. The guidance capitalises on the experience of drafting a national EP (to be published on 30 June) and will be disseminated to the MAs on 1 June.

²¹ Austrian, Flemish and Greek authorities, for example, did not have a particular view on the guidelines provided by the Commission, given the relatively little work done on the EPs. They expressed a general view that the guidelines look reasonable but that problems may emerge once authorities engage fully with the drafting of the Plans and the Commission examines the EP drafts. For Austria, key issues are flexibility and proportionality.

4.2 Objectives and timing

The main objectives and timing of the evaluation activities foreseen in the EPs also vary across IQ-Net partners.

The main objectives associated with the Evaluation Plans range from:

- ensuring that evaluations provide inputs for the two **Strategic Reports of 2017 and 2019**. In **Austria**, for example, evaluation is intended to become part of the strategic process at PA level, i.e. a national-level strategic discussion similar to the STRAT.AT+ process during 2007-13, and in continuation with past evaluation arrangements, notably STRAT.EVA (in 2007-13) and KAP-EVA (in 2000-06);²²
- **identifying evaluation needs in a timely manner**, taking into account the requirements of various types of stakeholders, and through this, allowing for **data requirements** to be mapped at the outset of the programme period, so that information systems are planned to collect the relevant information, and additional data requirements that cannot be obtained from available statistical sources are collected through additional activities (e.g. in **Portugal** and **Slovenia**);
- **learning**
 - **about ‘what works’ and ‘what doesn’t work’** i.e. answering applied questions that are useful for projects and programmes (e.g. in **Wales**),
 - **about whether the programme requires adjustments and implementation lessons** (as in **Vlaanderen, Greece, Nordrhein-Westfalen**);
- appraising **whether projects have contributed to achieving their and the programme’s goals** (e.g. in **Denmark**) and **gauge the effectiveness, efficiency and impact of programmes and individual priorities**, in line with the regulations (e.g. in **Vlaanderen, Greece, Spain**);
- assessing the **added value of particular types of intervention**, based on project-level data (e.g. in **Nordrhein-Westfalen**);
- **improving the evaluation environment and the know-how for the management and delivery of evaluation** via coordination of evaluation processes and education of the MAs (e.g. in the **Czech Republic**); and
- **ensuring that the evaluations planned and undertaken are relevant and of good quality** (e.g. in the **Czech Republic, Portugal, Slovenia** and **Scotland**). In Scotland, it is felt that in 2007-13 evaluation activities were not planned so that they would feed into policy implementation and that to an extent there was ‘evaluating for the sake of evaluating’. The

²² The Austrian federal platform for evaluation KAP-EVA (*Koordinierungs- und Arbeitsplattform Evaluierung*) was set up by ÖROK in 2002. Its scope was to coordinate and support the evaluation of the Structural Funds programmes. It was re-launched for the 2007-13 programme period as STRAT.EVA, but only one event has been organised (in 2011) as part of the so-called 'STRAT.AT plus process', the Austrian NSRF-based partnership platform.

MA now intends to be more thoughtful about how evaluation activity is structured in 2014-20, to ensure that it is useful. In Portugal, the main objective will be to ensure that evaluations can serve as a basis for evidence-based policy-making, namely in relation to the efficiency, effectiveness and utility of supported interventions, as well as in support of good management and implementation of the ESI Funds.

The timings of evaluation activities foreseen in the EPs also vary:

- **Implementation first, impact later** – A general trend, amongst IQ-Net partners who are planning to carry out both implementation and impact evaluation, is for the former to be undertaken in the first half of the programme period and the latter in the second half (e.g. **Czech Republic, France, Slovakia, Spain** and others). In **France**, for example, evaluation activities will take place in two stages: a first stage, during 2014-18, will focus on data collection, setting out methodologies for future impact evaluations, work on the performance framework (to ensure that mid-term objectives are met), and an implementation evaluation on the YEI; and, a second stage, during 2019-22, will entail impact evaluation for every priority of each OP (each to take maximum six months and a value of €100,000) and the 2019 strategic report.
- **Staggered approaches to ensure management oversight** – In **Nordrhein-Westfalen**, the choice will fall onto a staggered approach, whereby evaluations are scheduled to take place at different times during the programme period, to allow the MA to actively oversee evaluations, meeting regularly with evaluators and keeping an eye on their work.
- **Flexible timing to reflect programme spend** – In **Wales**, the timing of evaluations, particularly those relating to the ERDF, will be adjusted according to progress with programme spend. The MA intends to tie evaluations in to where they think they have spent certain proportions of the Funds, so that there is enough to evaluate. There will be less flexibility on timing for ESF, however, where the regular ‘Leavers Survey’ is used to collect data on ESF longer-term indicators. There is only a requirement to report this on two occasions, but the MA prefers to carry out the survey more frequently so results are not skewed to particular reporting years. The MA will share the dataset with projects so they can use it in their own project evaluations. Perhaps two surveys per year will focus on data collection, which could be analysed and reported on only every two years.
- **Concentration in the last years of the programme period** – Where impact evaluation will dominate the agenda, on the other hand, as in **Denmark**, there is an expectation that evaluation activities will concentrate particularly in the later part of the programming cycle.

4.3 Scope and focus

The core of the evaluation activities will consist of the compulsory evaluations foreseen by the 2014-20 CPR, i.e. at least one for each Priority Axis and evaluations on the YEI in conjunction with the ESF. Nevertheless, and not surprisingly given the many and varied goals assigned to evaluation, additional evaluations will also be undertaken in a number of cases (based on the information available for those IQ-Net partners that already have quite developed EPs). By way of example, the list of evaluation activities foreseen under the Czech Partnership Agreement is provided in Annex II.

4.3.1 Thematic approaches

Thematic approaches, sometimes cutting across programmes and Funds, appear to be favoured in a number of cases. Sometimes, this choice is dictated by previous evaluation experiences. In **Finland**, for example, the evaluations in 2007-13 were thematic rather than covering specific priorities or programmes. They focused on: the promotion of business activities (theme 1); the promotion of innovation activities and networking and strengthening of knowledge structures (theme 2); improving regional accessibility and business environment (theme 3); and environmental impacts and sustainable development (theme 4). A similar thematic approach is likely to be continued in the evaluation plan for 2014-20. Future themes have not yet been formally discussed and there is as yet no formal indication on the content of the evaluations; however, given the emphasis on low-carbon themes, it is likely that this theme will be included. The possibility to carry out joint evaluations for the activities supported by various ESI Funds is also being contemplated, in order to appraise results as a whole, illustrating the complementarity of the funds. However, as noted, individual themes for the evaluations are yet to be decided (and will be set out in the Evaluation Plan).

The focus of evaluation was also under development in **Denmark** at the time of fieldwork. However, some elements of the future evaluation plans have already been decided, notably that one cross-cutting evaluation will be carried out with a focus on the effects of the Structural Funds, as well as other public funds managed by the growth fora, in the rural areas of Denmark. This has been decided because all growth fora are required to develop strategies for initiatives that will benefit rural areas. Further, there will be evaluations of all ERDF and ESF projects with an allocation of EU funding above DKK 2 million (c. €268,000). There may be also similar evaluations for projects below this investment threshold (a decision that will be taken by the individual growth fora) and a range of thematic evaluations (yet to be defined).

In **Nordrhein-Westfalen**, the MA is currently considering a number of optional themes to be appraised under each OP Priority, as illustrated in Table 4.

Table 4: Possible themes to be evaluated under the Nordrhein-Westfalen OP

Priority	Option being explored
Priority 1: Strengthening research, technological development and innovation	<ul style="list-style-type: none"> An initial evaluation of procedures used for selecting projects under the competitive call approach, which will be undertaken once the first round of competitive calls has been completed (i.e. in autumn 2015), with a view to assessing whether any adjustments are needed, also drawing on experience from 2007-13. Later in the period, an evaluation of whether/how projects have contributed to the development of one or two of the sectoral/thematic fields (<i>Leitmärkte</i>) that the competitive calls are targeting. It is likely that this evaluation will also draw on data on projects funded in 2007-13.
Priority 2: Increasing SME competitiveness	<ul style="list-style-type: none"> The effects of SME funding in the context of tourism projects, also drawing on experience from 2007-13. The effects of start-up funding, also drawing on experience from 2007-13.
Priority 3: Supporting efforts to reduce CO₂ emissions	<ul style="list-style-type: none"> The effects of various instruments used to support combined heat and power generation, also drawing on experience from 2007-13.
Priority 4: Sustainable urban and urban districts development/ preventative measures	<ul style="list-style-type: none"> The effects of various instruments used to support combined heat and power generation, also drawing on experience from 2007-13.
Priority 5: Technical Assistance	<ul style="list-style-type: none"> Communication work Whether the MA's efforts to simplify procedures relating to applications, commitments and implementation have been effective.
Non Priority-specific	<ul style="list-style-type: none"> Communication work Whether the MA's efforts to simplify procedures relating to applications, commitments and implementation have been effective.

Source: IQ-Net country research – Germany.

Options are also being discussed in **Spain**. What is clear is that there will be an increase in the number of evaluations and in the themes and instruments covered. All thematic objectives, priorities and OPs will be evaluated. Some examples of evaluations being considered are:

- Regional innovation strategies
- PA mid-term evaluation
- Needs analysis in rural areas (digital economy) and coastal areas (water treatment)
- Cross-cutting themes such as gender equality and sustainable development
- Efficiency analysis (input-output methodology)
- Financial instruments
- Case studies of integrated urban strategies

The Spanish national authorities are currently consulting the regions/intermediate bodies on what types of evaluations they would like undertake and which ones they would do themselves, which will be included in the national evaluation plan.

In the case of other IQ-Net partners, particularly where the work on the EPs is still at very preliminary stages, there is no clarity about the scope and coverage of future evaluations. For instance, in **Austria**, there is an intention to evaluate, at least in a first phase, the way EU and domestic funding are interconnected, but the exact content of future evaluation activities has yet to be defined; whilst in **Greece**, the coordination authority is considering commissioning a strategic evaluation on liquid waste

or energy. However, the full picture of the evaluations that will be undertaken will become clear only at a later stage.

4.3.2 Programme-wide approaches

In a couple of IQ-Net programmes, programme-wide evaluations, rather than thematic evaluations, will be favoured. This is the case of **Vlaanderen** and **Slovenia**, where evaluation activities will cover the whole programme, including each Priority as requested by the regulations.

4.3.3 Mixed approaches: thematic and programme-wide

A mixed approach will be followed in **Portugal**, where evaluations will cover:

- entire programmes – with a view of re-evaluating the relevance of overall programme strategies, or appraise the efficacy of implementation aspects related to the programmes, or quantify results and impacts at programme level;
- the PA as a whole – with a view of re-evaluating the relevance of the global strategy of the PA (for instance in response to contextual changes), as well as to appraise the global impact of the PA and its contribution to Europe 2020 targets;
- selected themes – e.g. the relevance of Cohesion policy funded interventions to other national or European policies, the impacts on selected themes, the extent to which certain interventions, which were not much evaluated in the past, are being successful, the effectiveness and relevance of key instruments and others; and
- specific territorial aspects – for instance, the way territorial specificities have been dealt with in the programmes, or the impact of programmes at sub-regional scales and others.

A preliminary list of the evaluations that are currently being considered is provided in the Table 5 below.

Table 5: Evaluations currently being considered as part of the Portuguese ‘Global Evaluation Plan’ of the Portuguese Partnership Agreement (Portugal 2020)

Programme evaluations	Contribution of the ESIF Funds to OP objectives by priority, 2018 and 2022
Thematic evaluations	Competitiveness: <ul style="list-style-type: none"> • Impact of RIS3 Strategies on R&I systems, 2018 • PA contribution to internationalisation of the Portuguese economy, 2018 • Impact of 2007-13 Structural fund programmes on the performance of Portuguese firms, 2016/2017
	Social inclusion and employment: <ul style="list-style-type: none"> • Youth Employment Initiative, 2015 and 2018) • Entrepreneurship, 2018
	Human capital: <ul style="list-style-type: none"> • PhD Scholarships funded by the 2007-13 OPs • Contribution of the PA to improving the quality of education and training systems • Impact of the PA on early school leaving, 2018 and perhaps 2021
Territorial evaluations	Evaluation of the implementation \ contribution of CCLD (2016 \ 2021)
Global evaluations	<ul style="list-style-type: none"> • Contribution of the PA to the Europe 2020 strategy, 2018 • Contribution of the PA to employment and employability, 2019 • The macroeconomic impact of the PA, 2019 • Meta-evaluation of the PA, 2019 • Contribution of the PA to green growth, 2019/2020

Source: Leal C (2015) *Evaluation Global Plan of Portugal 2020 (PGA PT 2020)*, Presentation at the 38th IQ-Net conference, Vila do Conde, 28 May 2015.

4.4 Resources and responsibilities

There is no general rule as to whether evaluations will be carried out in-house or commissioned to external evaluators, consultants and academics. In some cases, both approaches will be pursued, based on the expertise required by the evaluations and available in-house and on the market.

The financial resources set aside for evaluation also vary, though it should be noted that figures, where available, are mostly preliminary at this stage and that various IQ-Net partners have not yet finalised their budgets dedicated to evaluation activities (e.g. **Austria, England, Greece, Nordrhein-Westfalen, Pomorskie, Portugal, Spain, Scotland** and others). Where resources have been set aside (indicatively) for evaluation, they vary considerably, ranging for example from:

- c. €600,000 in **Vlaanderen** (just under 10 percent of the TA of the OP) – subdivided between a main, general evaluation of the entire OP (on effectiveness and efficiency of delivery) with c. €250,000, and a number of ad hoc evaluations to be defined during implementation, with a budget of c. €350,000;
- to a total of €990,000 for all evaluation activities relating to ERDF and ESF interventions in **Finland** (representing 1.3 percent of the TA of the OP);

- to the c. €1.3 million+ (DKK 10 million) in **Denmark** – this is the sum earmarked for the external evaluations of projects that have been granted more than DKK 2 million; however, additional funds will also be dedicated to thematic evaluations (amounts yet to be decided);
- to c. €1.8 million+ (c. 50 million CZK²³) in the **Czech Republic** (see Annex II for more detail). This is a purely indicative figure, since some MAs do not have clear ideas on allocations as yet but will be responsible for additional funding (e.g. for the IROP, no precise estimations of resources have been calculated and the MA generally prefers internal evaluations, undertaken with no ad hoc budget);
- to c. €4.2-4.9 million (£3-3.5 million) in **Wales** (up from £2.5 million in 2007-13);
- to c. €6 million in **Slovenia**; and, lastly,
- to c. €49 million in **France** (as illustrated in Table 6 below).

These allocations are sometimes higher than in the 2007-13 period (e.g. in Wales and Slovenia), although in at least one case (**Slovenia**) the lower evaluation budget of 2007-13 (€5 million, in that case, c. five percent of the Slovenian TA resources) was used only in part (about 50 percent).

It is interesting to note that where evaluation responsibilities are devolved to programmes or regions, such as in **France**, a considerable degree of variation can be expected in the financial weight assigned to evaluation activities (in Table 6). This factor does not appear to be linked simply to the financial scale of programmes, but presumably also relates to strategic choices, perception of need, and sensitivity of regional leadership.

Sometimes evaluation budgets are shared amongst different bodies, e.g. in **England**, half of the Technical Assistance budget will be handed out to the Local Enterprise Partnerships for their own activities (which may include evaluation), and in **Denmark**, where some evaluation funding will be retained by the DBA and another part devolved to the Regional Growth Fora.

Table 6: Resources earmarked for evaluation in 2014-20 French OPs

Operational Programme	Total OP (EU contribution: ERDF, ESF, YEI - as applicable)	Evaluation & Studies (EU contribution: ERDF, ESF, YEI - as applicable)	%
OP ERDF Alsace	87,207,490	396,398	0.45
OP ERDF Réunion Conseil Régional	1,130,456,061	1,000,000	0.09
OP ERDF-ESF Aquitaine	459,274,123	1,500,000	0.33
OP ERDF-ESF Auvergne	255,979,483	356,212	0.14
OP ERDF-ESF Basse-Normandie	226,830,000	1,038,921	0.46
OP ERDF-ESF Bourgogne	223,730,000	1,100,000	0.49
OP ERDF-ESF Bretagne	369,500,000	2,269,638	0.61
OP ERDF-ESF Centre	255,091,336	870,000	0.34
OP ERDF-ESF Champagne-Ardenne	228,141,635	354,137	0.16
OP ERDF-ESF Corsica	115,850,000	294,391	0.25
OP ERDF-ESF Franche-Comté & Jura	184,510,000	700,000	0.38

²³ MRD (2015): *Evaluation Plan of the Partnership Agreement for programming period 2014-2020*, approved in February 2015, available at <http://www.strukturalni-fondy.cz/cs/Fondy-EU/Narodni-organ-pro-koordinaci/Evaluace/Evaluacni-plan> (last downloaded 30 March 2015).

Operational Programme	Total OP (EU contribution: ERDF, ESF, YEI - as applicable)	Evaluation & Studies (EU contribution: ERDF, ESF, YEI - as applicable)	%
OP ERDF-ESF Guadeloupe Conseil Régional	610,700,000	2,434,462	0.40
OP ERDF-ESF Guadeloupe & St Martin Etat	203,800,000	834,970	0.41
OP ERDF-ESF Guyane Conseil Régional	392,480,000	1,099,707	0.28
OP ERDF-ESF Haute-Normandie	289,630,881	744,000	0.26
OP ERDF-ESF Ile-de-France & Seine	482,482,139	1,500,000	0.31
OP ERDF-ESF Languedoc-Roussillon	424,744,633	1,403,092	0.33
OP ERDF-ESF Limousin	144,670,000	680,000	0.47
OP ERDF-ESF Lorraine et Vosges	408,540,000	1,312,649	0.32
OP ERDF-ESF Martinique Conseil Régional	520,951,695	3,000,000	0.58
OP ERDF-ESF Mayotte	214,400,000	1,607,078	0.75
OP ERDF-ESF Midi-Pyrénées & Garonne	460,087,007	1,631,124	0.35
OP ERDF-ESF Nord-Pas de Calais	847,956,471	3,000,000	0.35
OP ERDF-ESF Pays de la Loire	379,460,000	1,150,000	0.30
OP ERDF-ESF Picardie	299,202,678	1,500,000	0.50
OP ERDF-ESF Poitou Charentes	267,950,000	937,825	0.35
OP ERDF-ESF Provence Alpes Côte d'Azur	433,090,120	3,052,717	0.70
OP ERDF-ESF Rhône Alpes	509,400,000	1,349,825	0.26
OP ESF Alsace	46,252,510	70,000	0.15
OP ESF Guyane Etat	83,900,000	650,000	0.77
OP ESF Martinique Etat	124,700,000	-	-
OP ESF Réunion Etat	516,843,939	-	-
IOP Alpes	34,000,000	500,000	1.47
IOP Loire	33,000,000	224,000	0.68
IOP Pyrénées	25,000,000	87,500	0.35
IOP Rhône	33,000,000	260,370	0.79
OP Massif Central (POMAC)	40,000,000	200,000	0.50
NOP EUROPACT	72,600,631	-	-
NOP ESF	2,893,824,983	10,000,000	0.35
NOP YEI	433,938,640	-	-
TOTAL	14,763,176,455	49,109,016	0.33

Source: CGET. These are indicative, preliminary amounts, relating to EU contribution only (no national co-financing), and which exclude internal staff costs.

IQ-Net partners display different preferences in relation to whether evaluation work will predominantly be carried out in-house or outsourced to external evaluators contracted through tendering procedures (for example in **England**, where the MA is planning to commission a framework contract, along the

model similarly followed by EU institutions – European Commission, European Parliament). In some cases, both approaches will be pursued – for instance in **Austria, Czech Republic, Denmark** and **Spain**. The Czech Republic appears to be the IQ-Net partner that is planning to make the most extensive use of internal evaluations (see the table provided in Annex II). Spain is also planning to make greater use of evaluation expertise internal to the public administration (notably from the Institute of Fiscal Studies attached to the Ministry of Finance and Public Administration, which is the Managing Authority of all Spanish ERDF programmes). Evaluations will be undertaken both in-house, by the MA, and externally, by universities and evaluation consultancies selected through public tenders as in the past.

Nevertheless, the use of external evaluators is the most frequent approach, in continuity with past practices (Table 7). In **Greece** and **Portugal**, however, the 2007-13 experience has highlighted that the evaluation services market is still small and needs to be developed either by building capacity in other companies to better address the needs of evaluations related to Cohesion policy or by developing in-house capacities to carry out independent evaluations. This is something that is intended to be addressed during the programme period.

Table 7: Approaches to evaluation: in-house v. outsourced

Mixed	Predominantly or wholly external
Austria (balance to be decided) Czech Republic Denmark (internal by DBA and growth for a mainly for implementation; external for impact) Spain (balance to be decided)	Vlaanderen England Finland France Greece Nordrhein-Westfalen (with possibility of limited internal) Portugal Slovakia (with possibility of limited internal) Scotland Slovenia Wales

Source: IQ-Net country research.

The coordination of evaluation activities is generally undertaken by national coordinating authorities or, where these differ, the Managing Authorities, generally supported by Evaluation Units, Evaluation Steering Groups or Evaluation Committees (e.g. in the **Czech Republic, Nordrhein-Westfalen, Pomorskie, Slovenia, Spain**).

5. FOCUS AND METHODOLOGICAL APPROACHES

As discussed in Section 2, one of the main intended changes in the focus of evaluation activities, compared to past programme periods, is the increased emphasis to be placed on the evaluation of achievements and impacts. Nevertheless, what the exact balance between implementation evaluation and impact evaluation will be, and, for the latter, what the relative weight of counterfactual impact evaluation (CIE) and theory-based evaluation (TBE) is to be are not yet clear.

5.1 Balance between implementation and impact evaluation

In a number of cases the **balance between implementation and impact evaluation is yet to be defined and will emerge as the EPs are finalised or even during implementation**, for instance: in **Austria** and **Slovenia**, the choice will depend on the evaluation questions that will be established and will be established on a case by case basis; or in **Portugal** where, beyond a general preference for horizontal/broader evaluations, covering themes or territorial approaches across a range of programmes, the balance between implementation and impact evaluation has not yet been decided.

Notwithstanding the uncertainties at this stage, in some cases – notably in the **Czech Republic**, **Slovakia** and **Spain** – the intention is to realise a balanced approach covering both implementation and impact: implementation evaluation in the first years of the programme period and impact evaluation in the second half of the programme period. So, whereas in practice there will be a strengthening of impact evaluation – in line with the Regulations – the aim is to achieve a balance.

In other countries, instead, a shift towards impact evaluation (compared to 2007-13) appears quite pronounced, notably:

- in **Denmark** – where the main focus will be placed on impact evaluation. Some implementation evaluation may be undertaken where relevant during the programme period (e.g. where measures do not progress as planned, to identify the key constraints and how to address them); however, in many cases implementation evaluations would be carried out internally by DBA and the growth forum secretariats and these exercises may be viewed not as evaluations in the strict sense;
- in **Finland** – where past evaluation activities focussed mostly on implementation; however, for the current period, whilst implementation evaluation will continue to a degree, the emphasis is expected to shift towards impact evaluation.
- in **France**, where impact evaluations on each OP priority for French programmes will be carried out (though at the moment no insights are available on methodologies that might be followed);
- in **Pomorskie**, where the MA intends to evaluate impacts at different levels - each ROP Priority, specific themes and horizontal issues (e.g. of business support, sustainable development), as well as the impact ROP as a whole;

- in **Slovenia**, where the MA expects the majority of evaluations to be theory-based (though there will also be some counterfactual evaluations, maybe in the context of the YEI, later in the programme period); and
- in **Wales**, where the balance between implementation and impact evaluation will be weighted more towards impact evaluation (although ‘without under-commissioning implementation-type evaluations’). Indicatively, about a quarter of the evaluation budget may be devoted to implementation evaluation, a further quarter to theory-based impact evaluation, and the remaining half to CIE (more on this in the next section).

Where implementation evaluation will continue to dominate is Vlaanderen, largely due to small financial scale of programme, which leads to minimal impacts and measurability problems.

Lastly, **an individual approach will be taken in Nordrhein-Westfalen**, where what will be pursued is a mix of implementation and project-specific evaluations (geared towards added value and results rather than impact on the regional economy). The view is that impact evaluations would be scarcely meaningful, given the financial size of the programme, and that focusing on projects results and added value would be more sensible.

5.2 Methods – balance between theory-based impact evaluation and counterfactual impact evaluation

Only a few IQ-Net partner programmes have a degree of clarity in relation to the approaches that they’ll adopt to evaluate impacts and most are not yet clear about the relative weight between CIE and TBE is not yet known. MAs are gearing up towards evaluation and are in some cases acquiring the necessary skills and information to make such decision in an informed manner (e.g. in Finland where an academic evaluation expert was asked to provide training to the regional and structural policy delegation at the Ministry of Employment and Economy).

Further, IQ-Net partners do not plan to be prescriptive regarding the methods in all cases. **Nordrhein-Westfalen** and **Slovenia**, for example, anticipate that in most cases the choice of methods would be left to the consultants, rather than specified in the calls for tenders. In Nordrhein-Westfalen, this will be the case particularly for those evaluations that will aim to assess the added value of Cohesion policy funding through an analysis of project-level results: in these cases methodological questions will be complex and context-specific, and will require the technical skills and expertise of evaluators to be suitably operationalised. This is also why the Nordrhein-Westfalen MA objected to the Commission’s request for methods to be specified in the EPs as inappropriate.

Amongst the IQ-Net partners that have achieved more clarity about the methods that they will likely employ are Czech Republic, Denmark and Wales. In the **Czech Republic**, where both CIE and theory-based impact evaluation are anticipated, the expectation is that theory-based approaches and case study methods will likely dominate, given the fact that CIE is considered appropriate only for a limited range of studies (e.g. ESF projects and programmes), but not for wider evaluations at priority or programme level. Whilst in **Wales**, CIE will be generally favoured because it is viewed as the most rigorous approach. The MA aims to use this method where possible, but accepts this would not be sensible for some types of investments (e.g. infrastructure). The MA envisages that CIEs will be possible for ERDF (and RDP) business-level interventions aimed at business growth and productivity,

and for ESF (and RDP) interventions aimed at helping people into work. In the programme areas where it is not feasible to undertake CIEs, the MA will work with project sponsors, scheme managers and their Evaluation Advisory Group to develop alternative methods – ‘although it must be acknowledged that these will lack the rigour of CIEs’.²⁴ The MA hasn’t commissioned theory-based impact evaluations in 2007-13 and would like to try and do more in 2014-20. However, they consider that this may best be done at project level, perhaps to complement elements of the Leavers Survey, which can provide data for large samples on types of activity, and which could potentially inform theory-based evaluation. In **Denmark**, both theory-based and counterfactual impact evaluations will be carried out. The external evaluators that will be contracted for cross-cutting impact evaluation of ERDF and ESF projects are expected to adopt a theory-based approach, evaluating the chain of effects of such projects; whilst CIE will be carried out internally by the DBA using data provided by Statistics Denmark, and this will continue years after projects are completed in order to evaluate the long-term effects. This will build on the ‘register-based approach’ started in 2011, illustrated in detail in Annex II).

Many IQ-Net partners have expressed **concerns about their ability to carry out impact evaluation**. These concerns relate to:

- the **measurability of impact on macro-aggregates when the programmes are a only small** portion compared to wider public policy spending (e.g. in **Austria, Vlaanderen, Nordrhein-Westfalen**)
- the **practical challenges linked to the timetable of the measurability of impacts**, and the difficulty to reconcile this timetable with the **short-term interests of politicians** who want to understand what interventions achieve and what works on much shorter timeframes (e.g. **Austria**);
- the **operational difficulties entailed by CIE** – related for example to the techniques to be used for the selection of control groups – and the fact that this method is only suitable to certain types of interventions (where there are large number of recipients and non-treated recipients for example, i.e. where it is possible to identify a suitable control group) but not for others; and the suitability not just of CIE but also of theory-based evaluation to appraise the impact of certain types of investments, like transport infrastructure (a concern emerged in **Greece**)
- the **limited or even absent experience of impact evaluation in 2007-13**. This applies to both CIE (even in countries where (limited) CIE was undertaken in 2007-13, like **Austria, Czech Republic** and **Portugal**, as reported in the examples provided in Annex I) and to theory-based impact evaluation (hardly used in 2007-13 by any of the IQ-Net partners); even when impact evaluations were attempted in 2007-13, these were sometimes found to have been of relatively low quality (with some exceptions, some of which are discussed in Annex I) due to the limited expertise available on both sides of evaluation (commissioning bodies and evaluators) and to a lack of quality data;

²⁴ WEFO draft Monitoring and Evaluation Strategy (March 2015) p. 7.

- a **lack of generalised evaluation culture** (in **Greece**) that inhibits the use of more sophisticated evaluation methods that require technical expertise;
- the **resource intensity of impact evaluation** and the difficulty, for small administrations (such as in **Vlaanderen**), to supply human resources that are necessary for the sourcing of data that are needed for impact evaluation and to deal with evaluators in directing and coordinating the evaluation studies;
- the **weak resilience of evaluation planning against institutional changes and budget cuts**. As illustrated in preceding parts of this report, impact evaluations need to be planned well in advance, to ensure that the data required are in place. However, in 2007-13 institutional and financial changes have had a negative effect on evaluation activities. For instance, in England, following the abolition of the RDAs and the centralisation of the MA function, there has been little evaluation activity (and what there has been has not been made publically available); in Slovenia, economic crisis and austerity have led to shifting focus on absorption; in England, the ERDF MA, the Department for Communities and Local Government (DCLG), has seen its staff cut by 30 percent during the period and a reduction in external expenditure for external consultants.

Capacity issues are being tackled with improvements to monitoring systems and collaboration with data providers, as discussed in Section 3, and with dedicated training initiatives, in a number of countries; however only time will tell if these measures will be sufficient.

6. CONCLUSIONS

By and large, IQ-Net partners view with favour the results-orientation focus of the new regulations and, related, the new evaluation requirements and the shift in emphasis from implementation to impact. These changes are considered to represent a move forward compared to the past programme period where financial absorption, partly linked to the economic crisis, dominated the agenda.

The expectation is that the new requirements will result in many cases in more and better reporting and evaluation, and in an improved understanding of the achievements of Cohesion policy in the regions and Member States, and its impact upon target groups.

There is awareness amongst IQ-Net partners that the availability of the right data and adequate data quality have been an issue in the past, and that this has hindered the ability of programme authorities to undertake good impact evaluation in 2007-13. In the light of the new regulatory requirements, IQ-Net partners have invested resources and efforts towards improving programmes' monitoring systems, and the availability and usability of administrative data, for example with enhanced collaboration with national statistics offices.

Capacities – within MAs, implementing bodies, project holders and the evaluation community – continue to be a problem in some countries, and new challenges are also emerging in relation to the necessity to carry out evaluations that require different skills than those employed in 2007-13. Improvements to monitoring systems and collaboration with data providers, as discussed in Section 3, and dedicated training initiatives (both on monitoring and evaluation) have been put in place in a number of countries. However, only time will tell if these measures will be sufficient to enable the realisation of quality evaluation, particularly impact evaluation.

In designing the new programmes, the identification of results indicators has been challenging for a number of IQ-Net programmes and there are still various open issues. Nevertheless, views gathered from IQ-Net partners and DG Regio Evaluation Unit indicate that the experience has been helpful and that this work has helped making the programmes more focussed and better targeted towards achievements.²⁵

The Evaluation Plans are seen by and large as useful by IQ-Net partners, for example because they allowed for a more systematic reflection on needs for and goals of evaluation activities. In most cases, IQ-Net partners have learnt from past experiences of designing and implementing Evaluation Plans. However, some have voiced concerns about some components of the EPs, as requested by Commission guidelines, such as the need to specify evaluation questions and methodologies at a too early stage, and about the approach that might be taken by the Commission in their appraisals of the EPs, which is feared to go beyond what is foreseen by the regulations.

The progress of IQ-Net partners with the preparations of the EP is mixed, with only a few at an advanced stage and a few who have yet to start. The delays with programme approvals, and the necessity for programme authorities to focus on programme launch (for 2014-20 programmes) and

²⁵ DG Regio's Evaluation Unit is currently summarising what it is that the programmes are anticipating realising. The fact that programmes have identified clear results indicators and targets is enabling for the first time the production of a narrative about what the programmes are intending to achieve, which would not have been possible at the start of the last programme period. This is viewed by the Commission as 'a good start'.

closure (for 2007-13 programmes) appear to have played a role in the lagging status of EP preparations in various cases. A key challenge will be to ensure that the need to respond to mandatory EP templates (not just European, but sometimes also national), and to amend the EPs in response to the appraisal carried out by the European Commission's Evaluation Units, does not result in the transformation of EPs into administrative requirements, rather than tools put in place to improve the performance of programmes.

Insofar as it is possible to provide an overview at this stage, IQ-Net partners are taking different approaches with regard to the coverage of their EPs. Some EPs cover a single fund across a country, others only individual OPs, others again a multiplicity of OPs and others. The goals assigned to the EPs and the evaluation activities therein contained also vary, ranging from learning about 'what works', to appraising projects' contribution to programme goals or their added value, to the appraisal of programme effectiveness and impacts, to the assessment of delivery efficiency and the provision of information which is useful to improve programme delivery as it unfolds. Choices regarding who undertakes evaluation work – whether the MAs or other bodies within the public administration, or external consultants selected via public procurements – also vary. Such decisions often reflect the financial resources and skills available and the expertise required for different types of evaluation. Whatever the choices, emphasis is intended to be placed on the quality control of evaluation outputs.

Whilst impact evaluation will gain prominence on the whole, as required by the new regulatory framework, it is still too early to appraise how such type of evaluation will be carried out in practice, and whether the skills and data in place will be adequate to enable it. Based on the information available at this stage, CIE appears to be preferred over theory-based impact evaluation; however, a number of concerns have been expressed by IQ-Net partners about the practical applicability of both these approaches.

Two themes that did not emerge particularly strongly from the field-research are the dissemination and follow-up of evaluation activities, and the involvement of stakeholders in the design and delivery of evaluation, which nevertheless are important aspects of the evaluation cycle (as well as required by the CPR).

The research undertaken leads to a number of questions:

- first, **data quality** continues to be an issue and some partners are implementing Action Plans to this purpose:
 - Where successful measures have been put in place to improve the availability of administrative data and/or the quality of monitoring data, can any lessons be drawn from these (for the benefit of other IQ-Net partners)?
- second, with regard to the **methods** for impact evaluation IQ-Net partners have expressed different views about their preference amongst CIE and theory-based evaluation. In many cases, IQ-Net partners are still unsure about the extent to which such approaches will be used (to the extent that a few do not appear to be planning to use extensively neither approach):
 - What are partners' views about the practical benefits and drawbacks of these approaches?

- Do partners think that CIE and theory-based evaluation are sufficient methods to appraise impacts?
- What are the barriers to the successful realisation of impact evaluation?
- Would there be merit in shifting the focus of impact evaluation, so that it transcends the co-funded programmes and encompasses instead also other domestic policies (e.g. where programmes, like in Austria, Vlaanderen, Denmark, Nordrhein-Westfalen, are too small to affect macro-indicators) or the broader framework of EU investments towards the achievement of Europe 2020 targets?
- Third, with regards to the **planning of evaluation activities**:
 - What would improve the utility of the Evaluation Plans?
 - Have stakeholders been involved in the planning and design of evaluation activities? Where this has been the case, are there good practices that IQ-Net partners can share in this respect?
- Fourth with regard to **capacity**. After four cycles of Cohesion policy implementation, and despite sizeable investments dedicated to this, in a few countries there are still perceived weaknesses about evaluation capacities and lack of an evaluation culture:
 - What has failed in past programme periods in this respect?
 - What would be useful measures to address these shortcomings?
 - The domestic context, as illustrated amongst others in the Sixth Cohesion Report, plays an important part in determining a country's or region's ability to effectively utilise Cohesion policy resources. The same can presumably be argued about evaluation capacity. Are the ex ante conditionalities considered sufficient to address the shortcomings faced by some Member States? And, what is IQ-Net partners experience with these?

7. ANNEX I – MINI-CASE STUDIES ON IMPACT EVALUATIONS UNDERTAKEN IN 2007-13 AND EVALUATION PLANS

7.1 Austria – Impact evaluation of ERDF-funded federal environmental measures in 2007-13²⁶

Original title: “Wirkungsevaluierung – ein Praxistest am Beispiel der EFRE-geförderten Umweltmaßnahmen des Bundes in Österreich in der Periode 2007-2013”²⁷

7.1.1 Rationale

The study looked at measures implemented by the federal intermediate body KPC (Kommunalkredit Public Consulting). It was divided into two parts: an impact analysis and a system analysis. The reasons for carrying out this evaluation were the good measurability of impacts (quantitative indicators), the almost nationwide implementation of the measures (seven-out-of-nine Austrian OPs) and the interest and support of the responsible intermediate body (Federal Ministry of Agriculture, Forestry, Environment and Water).

7.1.2 Delivery

The study was carried out between July 2010 and May 2011 and published as part of the ÖROK publication series. The study was commissioned by ÖROK and the evaluation team consisted of a consultant from the ÖAR Regionalberatung,²⁸ who carried out the impact analysis part, and two consultants from RIMAS (Research Institute for Managing Sustainability) at the Vienna University of Economics and Business,²⁹ who carried out the system analysis part. The impact analysis was divided into two stages: definition of an impact model and measurement of impacts.

The evaluation focused on impacts of measures implemented by the intermediate body KPC, which is responsible for €31.7 million of ERDF funding across seven of the nine Austrian ERDF OPs of the 2007-13 programme period (all but Carinthia and Vienna).

The main quantitative data sources were ATMOS (the Austrian ERDF monitoring system) and KPC's own database. The evaluators included data from the start of the 2007-13 programme period (first project approvals took place in November 2008) until June 2010. Qualitative data was collected via interviews with 10 project owners and 18 key actors involved in programme management (e.g. MAs and IBs).

The study applied the so-called 'Process Monitoring of Impacts' (PMI) approach, which was developed by an Austrian consultancy in 2005. It has been tested in the context of INTERREG III A and B, as well as for an evaluation of the Objective 2 Programme Styria in 2004 and 2005.³⁰ PMI

²⁶ This case study has been drafted by Stefan Kah.

²⁷ Melidis K, Kopp U and Martinuzzi A (2011) *Wirkungsevaluierung - ein Praxistest am Beispiel der EFRE-geförderten Umweltmaßnahmen des Bundes in Österreich in der Periode 2007-2013*, ÖROK Schriftenreihe Nr. 186, Wien.

²⁸ One of the Austrian consultancies in the area of regional development (<http://www.oear.at/?lang=en>).

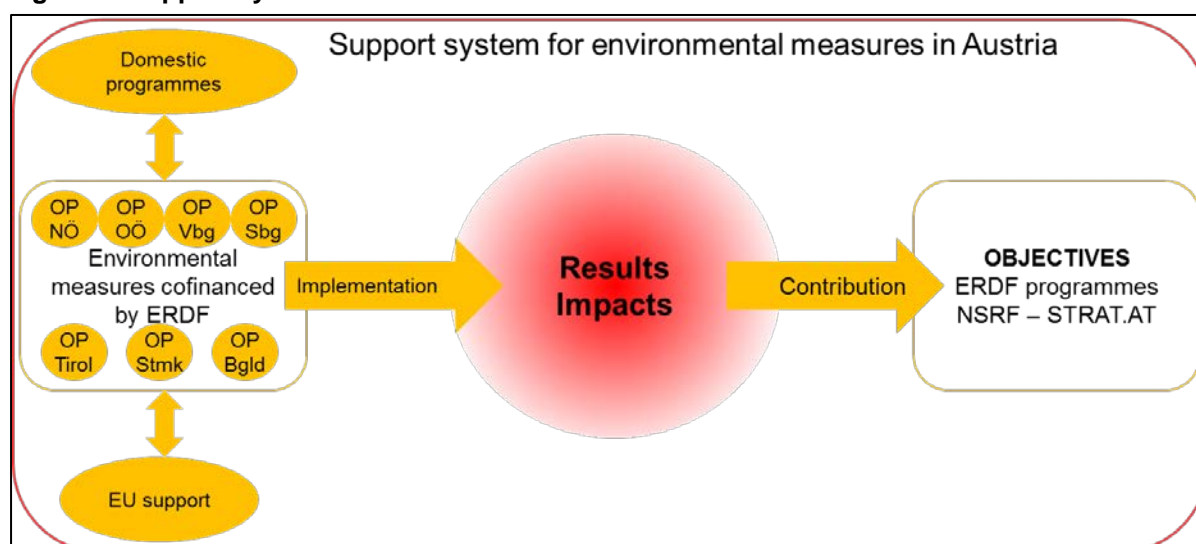
²⁹ <http://www.sustainability.eu/>.

³⁰ Hummelbrunner (2005) *Process Monitoring of Impacts. Towards a new approach to monitor the implementation of Structural Fund Programmes*, Vienna. Available at: <https://www.bka.gv.at/DocView.axd?CobId=14624>.

defines models of intervention (projects, fields of activities, programmes) and orientates the focus of monitoring activities towards the observation of these models and associated processes that should lead to the desired impacts. Processes are defined as impact-chains (output-result-impact), where outputs are linked to results by deriving assumptions about how/by whom outputs should be used in order to produce desired effects. Monitoring shifts the focus to factors that can be directly influenced by activities (projects), thus placing emphasis on results. PMI leads to the establishment of a comprehensive management information system, which structures programmes and their intervention logic along impact diagrams and adjusts existing monitoring elements and procedures in order to observe the implementation of these models.

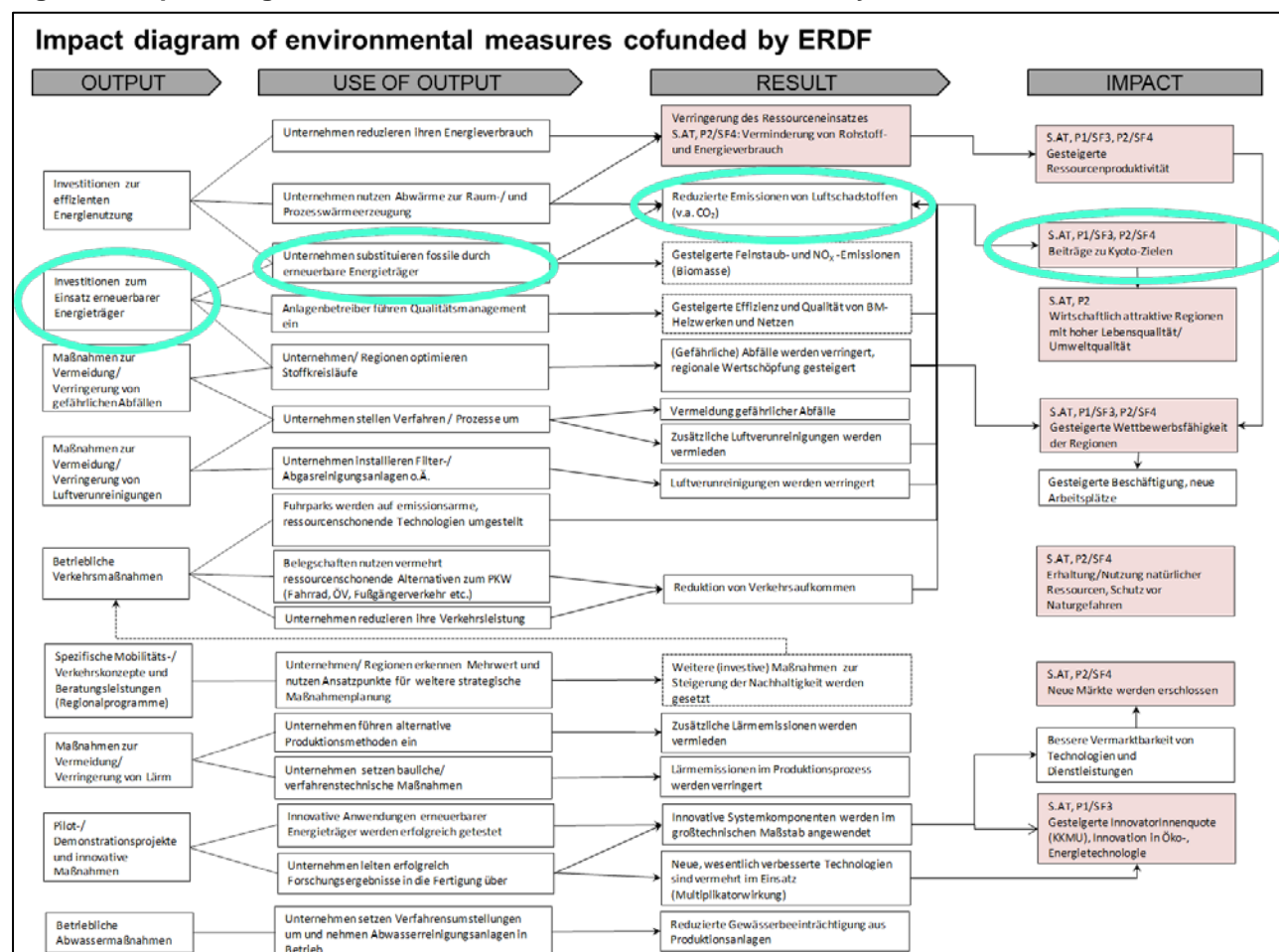
Figure 4, below, shows how the implementation of environmental measures co-financed by ERDF is supposed to create results and impacts that contribute to the objectives of ERDF programmes and the Austrian NSRF.

Figure 4: Support system for environmental measures in Austria



Source: Melidis K (2011) *Programmübergreifende Evaluierung der EFRE-kofinanzierten Umweltmaßnahmen der KPC*, presentation at the STRAT.EVA Workshop on 7 April 2011, Vienna, p. 2. Translated by the author.

The impact analysis identified nine different project types for environmental measures co-funded by the ERDF and implemented by KPC. Each of them displays distinct impact mechanisms and contributes differently to the achievement of programme objectives. Figure 5 shows how these nine different project types produce nine different outputs (y axis), which are used differently (second column) to create different results (third column), which can then be related to impacts (fourth column) defined in the OP.

Figure 5: Impact diagram of environmental measures co-financed by ERDF

Source: Melidis K (2011) *Programmübergreifende Evaluierung der EFRE-kofinanzierten Umweltmaßnahmen der KPC*, presentation at the STRAT.EVA Workshop on 7 April 2011, Vienna, p. 6. Partly translated by the author.

For one of the nine project types, Table 8 provides an example of how outputs link to results and then impacts (translation of the highlighted boxes above).

Table 8: Example of an output-result-impact chain

Output	Use of output	Result	Impact
Investments for the efficient use of energy	Firms substitute fossil by renewable energy sources	Reduced emission of air pollutants	<ul style="list-style-type: none"> • Priority 1, intervention area 3 (innovative investments in firms) • Priority 2, intervention area 4 (energy savings, renewable energy and environmental investments) • Contribution to Kyoto objectives

There were positive findings with regard to the achievement of target values for the defined result indicators. For instance, the additional capacity of renewable energy (123.9 MW) had already reached double its target, and the reduction of greenhouse gases stood at 46 percent (151.9 kt/a).

However, the main added value of the evaluation was to gain experience in connecting outputs with results via assumptions on the use of outcomes. It is acknowledged that this approach requires additional work, as assumptions have to be elaborated and also verified. It is possible to use

monitoring data for such verification, combined with additional data, collected ad hoc. When applied during ex ante or ongoing evaluations, the actual achievement of results cannot be captured. To do so requires a combination with other approaches.

7.1.3 Challenges

The main challenge was the definition and clear delimitation of the evaluation question, especially in consideration of available data. A methodological issue resulted from the fact that data had been linked to impact chains in an ex post process.

7.1.4 Assessment

The evaluation fulfilled ÖROK's expectations. It was commissioned and carried out at a time (2010/11) when there was a strong move towards result-orientation and impact evaluation. One of the main lessons was that evaluation already begins during the programme planning stage, as this is the time when the necessary preconditions must be set to allow for a meaningful programme design and evaluation of achievements at a later stage. The evaluation can be seen as useful as an exercise about connecting financial inputs and the related outputs, results and impacts. This was a fundamental skill for the design of the new programme.

7.2 Czech Republic – Pilot counterfactual evaluation of human resources and employment OP, support area 1.131

7.2.1 Rationale

The reasons for selecting the counterfactual impact assessment (CIA) as the method for this evaluation consisted of personal conviction by the then head of evaluation unit at the Ministry of Labour and Social Affairs, who had received training on this technique, on the meaningful applicability of the CIA method. It was one of the first applications of this method in the Czech Republic and is still assessed as one of the best evaluations based on the CIA method.

The head of the evaluation unit was strongly convinced of the applicability of the method for Operational Programme Human Resources and Employment (OPHRE) interventions. The original idea was to apply the CIA to the active policy of employment, but the evaluation was hindered by the lack of data. The focus was thus transferred to the evaluation of impacts of the OPHRE intervention supporting companies that implemented projects in support area 1.1, i.e. investment in human capital in companies and the development of systems that help to increase the flexibility of the workforce, its knowledge and skills, where the data were more readily available.

The main evaluation questions defined in the terms of references were:

- 1) What was the real causal effect/impact of the intervention 1.1 OP HRE (at the level of projects) in comparison with the counterfactual situation without intervention, on supported enterprises in terms of:
 - a. the competitiveness of enterprises (e.g. increase of volume of sales and profit),
 - b. companies' investments in training employees,
 - c. staff fluctuation in companies,
 - d. growth of enterprises expressed by newly created jobs.
- 2) Are there statistically significant differences of effects on some groups of supported enterprises?
- 3) What is the mechanism of the effect of the call for projects on supported enterprises? Why does the intervention work/not work?

The study is followed by "Evaluation of impacts of the OPHRE, support area 1.1, actualisation as a follow-up to the data from 2012-2013". The aim of the study is to provide a quantitative and qualitative analysis of the impacts of the OPHRR support on employers and employees in supported companies. The quantitative part of the study is elaborated by the internal capacity of the Ministry of Employment and Social Affairs, and the qualitative part is realised externally: "Evaluation of impacts of the support area 1.1 of the OPHRE with a usage of qualitative methods" as a project. The evaluation uses long-

³¹ This case study has been drafted by Dr Lucie Jungwiertova. For more information, see: MLSA (2015) Outputs from „Pilotní counterfactual impact evaluation OP LZZ, oblast podpory 1.1“, available at <http://www.esfcr.cz/file/8798/> (last downloaded 25 April 2015).

term time-series data that were previously unavailable and is enriched with qualitative aspects that should shed light on the intervention mechanisms. The results should be available by summer 2015.

7.2.2 Delivery

The evaluation was commissioned by the Ministry of Labour and Social Affairs, by its evaluation unit. The evaluator was selected via competitive tender, and the offer of the IREAS centre (in cooperation with the University of Economics, Prague) was selected as the best one. The team of seven persons was led by Oto Potluka and Jan Brůha from the University of Economics, Prague. According to the interviewee, the competitive advantage of the main authors was their compatible knowledge of the analysed issue – Oto Potluka knew the intervention well and Jan Brůha was responsible for the statistical part of the analysis.

The objectives of the evaluation were twofold. First, the project had to evaluate the impacts of the intervention on companies that implement projects in the support area 1.1, which is focused on investment in human capital in companies and on the development of systems that help to increase the flexibility, knowledge and skills of the workforce. The project also included an assessment of the effect of ESF support on various groups of companies. Companies were divided by size, region, sector, forms of support and types of training received. The possible mechanism of the effect of support from the ESF on supported companies was also identified. Second, the evaluation had to expand awareness among the evaluation community regarding the scope for the application of counterfactual impact assessment for the evaluation of public policies and particularly of the interventions provided from the EU Structural Funds in the Czech Republic. This is the reason why workshops and seminars were included in the required outputs.

The timetable for project realisation was given by the ToR, and the evaluation took place between 2011 (2Q) and 2013 (3Q). The following quantitative evaluation methods of counterfactual impact evaluation were used to evaluate the impact on supported companies:

- Differences in differences, using a propensity score
- Regression discontinuity
- Instrumental variables

The selected methods were applied to an anonymised data sample of 34,065 companies, which was provided by the Czech Statistical Office (CZSO) and the database Monit7+. The impact was evaluated on sales and profit, number of staff members, investment in training and fluctuation.

7.2.3 Challenges

First of all, it should be emphasised that this evaluation is considered as one of very few cases of a good application of the CIA method in the Czech Republic. However, the interviewee mentioned some challenges in hindsight that can be grouped into the following categories:

- Evaluation design – the interviewee and the main author of the evaluation design would have proposed applying a mix of methods in hindsight. He used to be persuaded that the CIA could be a stand-alone method, but now he considers it as a test of a good quality that should be

completed or (even better) preceded by well-developed theoretical assumptions, based on a good knowledge of the policy causality.

- Skills within the commissioning body and evaluation team – both sides had a relatively good knowledge of the CIA method. One member of the commissioning body discussed evaluation questions and design with one of the main European promoters of the method (Enrico Rettore) to obtain a strong theoretical basis for the study. In retrospect, however, the interviewee would have changed certain aspects (see below) linked to the state-of-the-art of their knowledge at that time.
- Data availability issues – data availability resulted in the modification of the focus of the study – from the active policy of employment to the evaluation of impacts of the OPHRE intervention on companies that implemented projects in support area 1.1. Another problem was linked to the available time-series given by the short period since the intervention had been carried out. This is the reason why another evaluation with access to new data is being elaborated. The data availability also influenced the methods used. It was difficult to aggregate some data, and thus only the propensity score-matching could have been applied, and consequently the results were not significant enough.
- Timetable of the evaluation – the evaluation was realised between 2011 and 2013, a relatively tight timetable with respect to data availability. The interviewee would have proposed a more flexible timetable.

7.2.4 Assessment

The expectations of the evaluation were twofold, mirroring the objectives. The aim to expand awareness among the evaluation community concerning the potential scope for the application of counterfactual impact assessment in the evaluation of public policies was successfully fulfilled. The real evaluation of the impacts of the intervention was influenced by the challenges described above.

The main strength of the method is the good quality test that it can provide. However, this method should be accompanied by other methods – a qualitative insight into the intervention should not be underestimated. The CIA method is emphasised by the EU, but it is very sensitive to many factors – data availability and the suitability of interventions for this form of analysis.

If the clock could be turned back, the interviewee would elaborate more hypotheses for the evaluation and would not adopt the hypotheses from the policy-makers. He would extend the timetable to gain enough time to obtain data. He would also work with the process-tracing method to gain a better understanding of the mechanism of the intervention.

The experience influenced the evaluation plans for 2014-20 at the level of the OPs but not at the level of the PA. CIA requires homogenous interventions that cannot be ensured at the level of the PA.

7.3 Denmark – The register-based approach to counterfactual impact evaluation³²

7.3.1 *Rationale*

The introduction of the register-based approach to impact evaluation, which has been used since 2011, was initially based on discussions with and recommendations from Professor David Storey. He visited the Danish Business Authority (DBA) to share his experience of conducting evaluations and referred to the challenges of measuring effects with quantitative data. He encouraged the DBA to use the unique individual and company-level statistics that are available in Denmark to monitor the impact of the Structural Funds more closely. This was at a time when there were also requests from the regional growth fora to better document the effects of publicly funded initiatives in the regions. The DBA contacted Statistics Denmark to discuss how they might collaborate, and that was how the ongoing counterfactual impact evaluation was initiated. The approach was inspired by methods used by the Ministry of Higher Education and Science, the DBA and the regional Business Links (In Dan: Væksthuse) to measure the effects of other publicly funded programmes. The approach differs between ERDF and ESF projects.

For the ERDF, data is generated from Statistics Denmark's register of Danish companies in terms of their employment, turnover, export, etc. It is not possible to receive data from Statistics Denmark about individual companies; in using this approach, projects within the same priority areas are grouped. In order to evaluate the effects of Structural Funds projects, the development of the participating companies within a certain group of projects is compared to the development of a control group of similar companies that have not participated in the project. If the companies that participated in the project perform better, e.g. in terms of job creation, this indicates that the projects have had positive effects.

Initially, Statistics Denmark data has only been used to measure effects in projects with company participation, mainly under the ERDF. In the 2014-20 programme period, monitoring of individuals in projects is initiated using data gathered by Statistics Denmark. The results of ESF projects will be evaluated on the basis of personal data, which indicates the development of individuals that have participated in projects. For example, it is possible to identify whether individuals who did not have any qualification before the project started have received an education by the end of the project, and if individuals who were outside the labour market when the project started have come closer to inclusion when the project ends, e.g. whether they receive salary, etc. In some ESF projects, company data can be used as a supplement, e.g. in cases when already employed individuals undergo training in order to improve the results of the company. As with the ERDF, the individual data used for measuring results of ESF projects will be carried out by grouping projects and comparing with a control group. Representatives from the DBA believe that assessing the results of ESF projects will generally be easier, because the results will be evident when the project is completed or shortly thereafter, while ERDF projects often have more long-term effects based on macroeconomic indicators.

A benefit of the method which is highlighted by the DBA is that it is possible to monitor the effects of projects during and also years after the projects have ended, without requiring any input from the

³² This case study has been drafted by Lise Smed Olsen.

project beneficiaries. Furthermore, this monitoring of companies has provided an overview of types of companies that participate in ERDF projects. The participation of companies and individuals was previously not registered systematically.³³ As a counterfactual evaluation, the approach is considered useful by the DBA and the growth fora, but it does not explain why projects have facilitated results. For this purpose, it is also important to implement theory-based evaluations.

7.3.2 Delivery

The DBA, the five regions, and the regional municipality of Bornholm decided to develop the register-based approach in collaboration with Statistics Denmark. Three employees at the DBA are involved with impact evaluation using the data received from Statistics Denmark. One-to-two employees from each of the growth forum secretariats were involved from the beginning in discussing the development of the method and how to present the analysis of results. A steering group has now been set up, including the representatives from the growth forum secretariat that meets twice a year to discuss the methodological approach and results of the effects monitoring. Initially, a cooperation agreement was made between the parties for 2011-13, and a new agreement has been decided for 2014-16. It is planned to continue with three-year agreements, where resources are allocated to register-based evaluation. The funds from the DBA are granted from the government's annual budget, and the regions use regional development funds. The register-based approach is not financed with EU Structural Funds, because the regions also have the opportunity to measure the effects of other publicly funded initiatives using the same approach.

All ERDF (and, when relevant, ESF) projects that have involved company participation were included in the evaluation. In the current programme period, individual participants are also included in the evaluation. In 2014, a cross-regional register-based impact evaluation with a focus on projects initiated during 2007-10 was made public by the DBA.³⁴ In the autumn of 2015, the same evaluation will be performed again, this time including projects that were initiated in 2011.

The growth fora have been able to use the statistics in order to document the effects of their regional development strategies. For example, the Region of Southern Denmark based its evaluation on areas prioritised for regional development during 2007-13. Two evaluations have been published. In 2013, the results of the evaluation based on projects initiated in 2010-13 show significant outcomes within three of the prioritised areas in terms of job creation, which has increased compared to the control groups. For the tourism-related projects, it was not possible to identify clear labour market effects from the ERDF projects.³⁵ An updated version published in 2014 focusing only on projects initiated in 2007-10 shows particularly good results for job creation in tourism.³⁶ Evaluation of seven prioritised areas for regional development in the Mid Denmark region, including projects initiated in the period 2007-10, show positive results within five focus areas, but tourism and the focus area of welfare

³³ Olsen L S (2014) IQ-Net thematic paper, Denmark, spring 2014.

³⁴ Erhvervsstyrelsen (2014) *Faktabaseret monitorering og effektvurdering af strukturfondsindsatsen i 2007-2013*. http://regionalt.erhvervsstyrelsen.dk/file/540641/effektvurdering_strukturfondsindsats.pdf

³⁵ Syddansk Vækstforum (2013) *Monitering og effektmåling: Strukturfondsprojekter i Syddanmark*. http://regionalt.erhvervsstyrelsen.dk/file/423519/effektmaaling_strukturfonde_syddanmark.pdf

³⁶ Syddansk Vækstforum (2014) *Monitering og effektmåling: Strukturfondsprojekter i Syddanmark 2007-2010*. <http://detgodeliv.regionsyddanmark.dk/images/Media/Publikationer/Tema/PDF/Effektmaaling%202014.pdf>

innovation do not demonstrate positive labour market effects from the ERDF projects.³⁷ Each of the six growth fora decided whether to publish the impact evaluations from their regions; only evaluations from Mid Denmark and Southern Denmark are available on the DBA's website for effects monitoring (<http://regionalt.erhvervsstyrelsen.dk/effektmaaling>).

The DBA is currently carrying out focus evaluations of selected larger projects implemented during the previous programme period in the different regions. The criteria for project selection are that they have had enough participants that it makes sense to measure the development and employment of participating companies. The technical approach called indirect standardisation, as recommended by Statistics Denmark, is used for the evaluation.

7.3.3 Challenges

The DBA has experienced some challenges with the registered-based approach to impact evaluation. First, it was challenging to introduce the tool for indirect standardisation at the DBA, e.g. it involved extensive programming. Second, it has in some cases been a challenge to match the CPR (personal/social security) and P (company registration) numbers, because they were not registered correctly by the projects. In the reporting system that is under development for the 2014-20 period, there is a control system that links with Statistics Denmark to ensure that CPR and P numbers are registered correctly from the start. Third, the way in which company participation has been registered has been a challenge in some cases, e.g. in cases where high-growth companies have barely participated in a project except for one meeting or event, but have still been registered as a participant, the effects evaluations have been misleading. In the current programme period, it has been specified that only companies that participate as partners or state aid recipients can be registered as participants. Finally, an overall challenge in using the counterfactual evaluation approach is that it is only possible to control what can be observed from the figures received from Statistics Denmark, not the explanatory factors that cannot be observed and which can be very significant. It is therefore important to use the evaluation with some provisos regarding the method used.

The presentation of the cross-regional impact evaluation by the DBA briefly mentions weaknesses in the approach: "The effects of individual projects are not evaluated, but the entire group of participating companies of all Structural Fund projects. The evaluation therefore does not take account of the content of individual projects."³⁸ The evaluation for Mid Denmark includes the same explanation. The evaluations for Southern Denmark provide a more elaborate explanation: "The results indicate that the Structural Fund projects have had an effect. Yet they should be interpreted with caution. The impact evaluation ensures that the participant groups and control groups are similar regarding company size, industry, location and growth prior to the impact evaluation period. In terms of other firm characteristics such as motivation and commitment, growth orientation, development needs etc. the participating companies may differ from the control group."³⁹

³⁷ Vækstforum Midtjylland (2014) *Faktabaseret monitoring og effektvurdering af Vækstforum Midtjyllands regionale indsatsområder i 2007-2013*.

http://regionalt.erhvervsstyrelsen.dk/file/546021/effektvurdering_region_midtjylland.pdf

³⁸ Erhvervsstyrelsen (2014:16).

³⁹ Syddansk Vækstforum (2013:3; 2014:3).

7.3.4 Assessment

According to the DBA, there has generally been approval of the register-based approach for impact evaluation by the DBA and the growth fora. It is considered important to have gained an overview of the participants and their development after participation in Structural Funds projects, which may be considered the greatest strength of the approach. Further, the DBA has been contacted by the European Commission and other Member States that are interested in learning about the approach. Most recently, the Managing Authorities of ERDF and ESF in Sweden visited the DBA in order to learn about the method. Sweden and the Netherlands, with their data availability, are among the few countries that have the possibility to carry out the same type of register-based impact evaluation.

The main weaknesses as mentioned above involve the way in which company participation has been registered differently, and that the figures do not explain why an initiative has been successful or not. At the start-up meeting that is organised at the DBA for each project before it is initiated, the specifications for how to report participation, etc. are presented and discussed with the applicant. For the Evaluation Plan, it is currently being considered whether the external evaluator (for projects granted more than DKK 2 million) should also participate in the start-up meeting.

The experience with the register-based approach that was initiated in 2011 has resulted in a desire for its continuation in the Evaluation Plan for 2014-20. The DBA is under the impression that the Monitoring Committee will back this proposal. In the Plan, it will be described how the counterfactual evaluation will interlink with the theory-based evaluations.

7.4 Portugal – Counterfactual impact evaluation of business incentive schemes in Portugal⁴⁰

7.4.1 Rationale

Evaluations of business aid schemes in many European countries have often been based on qualitative research methods and the analysis of the evolution of indicators linked to the goals of the interventions (investment, job creation, spending on innovation, etc.). These approaches are useful for monitoring and management purposes, but they do not capture the actual impact of incentive schemes on business performance. In this context, the European Commission has urged Member States to employ more rigorous counterfactual approaches to assess the impact of business aid schemes by comparing the performance of assisted firms with non-assisted firms. Similarly, the Memorandum of Understanding accompanying Portugal's financial assistance programme highlighted the need for impact studies of SME support instruments in Portugal. This study aimed to address these knowledge gaps, taking advantage of available data from the previous programme period of the Structural Funds in 2000-06.

7.4.2 Delivery

The study was conducted in-house by staff of the NSRF Observatory (the public body responsible for programming and evaluation coordination) with technical support from Prof Daniele Bondonio (an external expert of DG Regio and Professor at the University of Eastern Piedmont in Italy), the Research Unit on Complexity and Economics (Technical University of Lisbon) and with the collaboration of the Strategy and Planning Office of the Ministry of Solidarity and Social Security (GEP / MSSS), the Managing Authority of the Competitiveness OP and the National Statistics Institute to collect the necessary data.

The study evaluated the impact of a range of business aid schemes (SIME, SIPIE, SIUPI, NITEC, SIED, QUA-DROS, URBCOM, PIFC, SIVETUR e MAPE) within the framework of the 2000-06 OP 'Incentives for the Modernisation of the Economy'. The counterfactual technique employed was "coarsened exact matching", based on a comparison of performance in groups of assisted firms and non-assisted firms with similar characteristics (notably in terms of size, maturity, sector, level of skilled employment, geographical location). The sample included 3,904 assisted firms and 221,258 non-assisted firms, and the main impact variables were the business survival rate and job creation (including skilled jobs). The study was launched at the end of 2011 and completed in 2012 at a cost of €9,900.

7.4.3 Challenges

A design weakness of the study is that it examined the impact on a limited number of performance indicators relating to business survival and job creation, excluding the impact on other core objectives of the schemes (productivity, innovation and internationalisation) due to data availability limitations. There were also data gaps in relation to the financial capacity/autonomy ratios of firms (before 2004), although robustness checks for different time periods found that this did not bias or impact on the results of the analysis.

⁴⁰ This case study has been drafted by Dr Carlos Mendez.

7.4.4 Assessment

The study provided interesting conclusions that support and challenge government thinking. First, the positive results suggest that the aid schemes are well designed. Second, the incentive schemes are particularly effective in weaker firms (e.g. new companies with limited financial capacity). However, support for these companies brings additional risk, which explains the government's decision to establish financial capacity as a key criterion for eligibility. Third, the average cost of each job created is lower in large firms than in SMEs and more jobs are created, on average, in large firms. The implication is that large firms should remain eligible for support in the future, contrary to European Commission recommendations. Fourth, the schemes are more effective in creating jobs when targeting companies whose activity is directed primarily at national/external markets, supporting government decisions in the 2007-13 and 2014-20 periods to prioritise companies most exposed to international competition.

Reference

Mamede R, Fernandes T, Alexandrino da Silva A (2013) *Análise contrafactual dos impactos dos incentivos do POE/PRIME na sobrevivência e no crescimento das empresas*, Coleção e cadernos do Observatório do QREN.

7.5 Poland – Determining the value of selected economic indicators for beneficiaries of Regional Operational Programmes and for selected control groups⁴¹

7.5.1 Rationale

This was a pilot evaluation study, carried out in 2014, designed to test the potential for using data from Poland's Main Statistical Office (Główny Urząd Statystyczny – GUS) for the purpose of quasi-experimental, counterfactual approaches. The evaluation focused on changes in the functions and characteristics of enterprises in receipt of EU funds through Regional Operational Programmes (ROPs) from the 2007-13 period. The counterfactual approach was taken for several reasons:

- First, from the perspective of public funding, the evaluation wanted to discover what would have happened had the firms not received support.
- Second, evaluation has been given increasing prominence in Poland over the past decade and there is a desire to build on existing capacity and experience, incorporating techniques such as counterfactual research, which are increasingly prominent in public policy evaluation.
- As part of this, the European Commission, including DG Regio, places increasing emphasis on impact assessments, including through counterfactual evaluations. Poland, as the largest beneficiary of Cohesion policy funding in the 2014-20 period, is keen to develop rigorous methodological evaluation techniques.
- The counterfactual approach was chosen for this specific case of business support because this technique facilitates the assessment of relations between different socio-economic factors. The socio-economic environment is crucial in understanding the impact of business support.
- Finally, it is important to note that the study is based on cooperation between several public bodies: the Ministry for Infrastructure and Development (MiIR, the main ministry for coordination of Cohesion policy), GUS and the Marshal Offices of four regional self-governments which are Managing Authorities for ROPs (Pomorskie, Śląskie, Zachodniopomorskie and Świętokrzyskie were the four cases selected). It is anticipated that this pilot study will be extended to all 16 regions, demonstrate the utility of GUS data for counterfactual studies, build capacity and prompt future counterfactual evaluations in subsequent years in GUS, MiIR and the regions.

The main aim of the evaluation was to establish values for selected economic indicators (relating to competitiveness, employment, R&D activities) in enterprises to meet the needs of the Ministry of Infrastructure and Development in its analyses of the effects of enterprise support from Cohesion policy in the 2007-13 period. Specific aims were:

- To develop a methodology for analysing data available from GUS suitable for counterfactual studies.
- To conduct four pilot studies in Regional Operational Programmes from 2007-13.

⁴¹ This case study was drafted by Dr Martin Ferry.

The key evaluation questions were:

- Which method and statistical techniques, related to counterfactual analysis, can be employed in analysing information from the GUS database?
- What is the situation of ROP beneficiary firms in comparison to a selected counterfactual control groups in terms of: employment level, competitiveness, innovativeness?
- What is the situation of ROP beneficiary firms in comparison to the control group, taking into account sub-groups of firms according to size, branch and region, local area, intensity of support, type of support?

This study represents a vital contribution to MilR work on impact evaluation, particularly theory-based impact evaluation. By producing consistent, regional-level data on a range of socio-economic indicators, this research will form the statistical basis of future studies. MilR will use the values established for selected economic indicators (among beneficiaries and in the control groups) to carry out theory-based evaluation of the impact of business support in the ROPs.

7.5.2 Delivery

Responsible actors – The responsible body was the Ministry of Infrastructure and Development, in conjunction with ROP Managing Authorities in the Marshal Offices of four regional self-governments.

Focus – The research focused on specific business support measures in the four case study ROPs 2007-13. These measures were funded under Priority 1 of the ROPs which cover technological research and development (R&D), innovation and entrepreneurship. Specific measures included in the research were: support for SMEs and micro-enterprises, including strengthening innovative activities, support for business infrastructure and support for businesses through financial instruments such as JEREMIE.

Timetable – The pilot case studies were conducted in 2014. Work to extend the study to the remaining regions, some national Operational Programmes and to begin theory-based evaluations that draw on the statistical base will commence in 2015.

Methods and techniques – The pilot evaluations involved an iterative, reciprocal exchange of information and data between the evaluation units in MilR, individual Managing Authorities in the four case study regions and GUS. The Managing Authorities identified four main stages. First, individual Managing Authorities would request data from GUS to meet the requirements of their proposed counterfactual evaluations of business support measures in specific ROPs, including potential indicators for assessment among beneficiaries and for the establishment of control groups consisting of equivalent firms that had not received support. Next, the Managing Authorities checked the utility of the data received from GUS and highlighted specific gaps and problems. GUS would respond to these queries and a set of indicators and associated statistical data were produced for counterfactual analyses. Counterfactual analyses were then carried out in each region. These were based on a standard approach that combined two statistical techniques: Propensity Score Matching and 'difference in difference'. Integrating both techniques allowed a detailed assessment of indicators and revealed additional differences between the selected populations of firms.

7.5.3 Challenges

The main challenges of the research related to:

- Lack of experience in carrying out counterfactual analyses among public authorities and more generally in the Polish evaluation community.
- Limitations in the data available. Although GUS has a comprehensive database, there were some definitional issues. For instance, the GUS classification of SMEs comprises those with fewer than 9 employees. This definition limits the potential population of firms that can be included.
- The scope of the research was also limited by the relatively early stage of ROP implementation when the study was carried out. The evaluation, conducted in 2014, was based on projects completed by 2012 and GUS data from 2013. In this respect, there were specific problems with indicators relating to innovative activity. An evaluation of the full impact of these ROP business support measures will not be possible until 2017.

7.5.4 Assessment

The pilot evaluation studies are regarded by MiIR as a success. The evaluations demonstrate significant variation in selected economic indicators between beneficiaries and control groups. The analysis used 11 common indicators related to the functioning of firms, including: employment, financial scale of operations, profitability, internationalisation, material and non-material assets etc.

The initial results and lessons learnt from the pilot cases included:

- The most significant differences between supported and non-supported firms concerned changes in level of employment (ROP beneficiary firms on average employed 9.6 more FTE positions than non-beneficiaries over the period). It should be noted that this increased employment meant that indicators of productivity (measured by output per head) did not record a strong performance among beneficiaries.
- There was also variation between beneficiaries and non-beneficiaries in the level of net receipts from sales. Beneficiaries recorded a rise in receipts that was modest but statistically significant in comparison to increase of receipts among non-beneficiaries.
- There was a clear tendency for ROP beneficiaries to initiate or strengthen export activities in comparison to non-beneficiaries.
- For several indicators, the research revealed significant differences not just between beneficiaries and control groups but between the regions.

One of the main strengths of this approach was that it demonstrated the value of the GUS database as a source for counterfactual analyses. This model also facilitated the comparative assessment of support measures in terms of different kinds of impact (e.g. on employment, productivity, internationalisation).

One weakness of this evaluation related to limitations in the capacity of organisations to carry out a new, challenging approach. There were also specific gaps in the data available, which limited the scope of the information produced. The strongest conclusions are drawn from simple measurements of employment, productivity and scope of operations. Gaps and inconsistencies in the data and issues related to the timing of the evaluation meant that measurement of important but less tangible impacts, such as innovative activity, were unfeasible in the short term. Definitional issues limit consistency of the data under specific headings.

The aim is to extend the pilot evaluation to all ROPs and to some national OPs. The study also served as a capacity-building exercise for counterfactual evaluation techniques, and it is anticipated that these techniques will be used in other studies launched by the national and regional bodies involved. Moreover, the data provided by this study will form the statistical bases for MiLR theory-based impact evaluations of Cohesion policy programmes.

7.6 Wales – 2012 ESF leavers survey⁴²

7.6.1 Rationale

The impact evaluation approach was trialled in 2010 to see if it would work; it did, and therefore every Leavers Survey since has included it. The aim of the 2012 ESF Leavers Survey was to assist in assessing the effectiveness of labour market interventions delivered under ESF. The two ESF OPs benefiting Wales in 2007-13 were together providing c. £1.4 billion of investment, and when initially agreed, the programmes were expected to provide support to almost 300,000 individual participants – 267,500 under the Convergence programme and 26,600 under the Competitiveness programme – i.e. around 10 percent of the Welsh population. Given the scale of the investment, it was seen as essential to evaluate the impact of interventions supported by the ESF OPs. The overarching objective of the survey was to understand the characteristics and outcomes for those who have participated in ESF projects.

Chapter 6 of the report presents the results of Counterfactual Impact Evaluation (CIE) analysis techniques based on data pooled from four Leavers Surveys (2009, 2010, 2011, 2012). The intention was to benefit from the increased statistical power that can be gained from a larger sample size – “This is particularly important in CIE analysis where relatively restrictive criteria have to be used for respondents to be incorporated in the analysis (namely the ability to observe transitions in economic activity over a period of at least 12 months) and the reduced sample sizes that can result.” (p. 7)

The development of the 2012 Survey built upon the experiences of the three previous surveys (2009, 2010 and 2011, also undertaken by the research team responsible for the 2012 report). Particular emphasis was placed upon maintaining continuity in the design of the surveys over time, allowing data from the surveys to be merged together to facilitate more detailed levels of analysis than could be achieved with the data from a single year. In the 2012 survey, only a limited number of relatively minor changes were made to the questionnaire where it was felt the data collected in previous years did not contain sufficient detail.

7.6.2 Delivery

The evaluation was commissioned by the MA (WEFO). The evaluators were a team of several units at Cardiff University with several consultancies. The evaluation examined two Priorities under each of the Welsh ESF OPs for 2007-13.

Telephone interviews were conducted with over 4,000 people who had left an ESF project delivered under Priorities 2 and 3 of the Convergence OP and Priorities 1 and 2 of the Competitiveness OP during 2012. Not all approved projects were included in the survey due to the availability of participant data at the time the sample was drawn. The telephone survey collected information on: pre-entry characteristics of ESF participants; motivations for participating in an ESF project; skills acquired as a result of the intervention and career details of respondents since completing the project. Also

⁴² This case study has been developed by Rona Michie. The reference of the full evaluation report is: Davies R, Jones S, Munday M, Williams G, Winterbotham M (2013) *The 2012 European Social Fund Leavers Survey*, Report to Welsh European Funding Office (WEFO) <http://gov.wales/docs/wefo/report/140127esfleaverssurvey2012reporten.pdf> (last downloaded 21 May 2015).

identified were employment outcomes, including 'softer' benefits from learning (such as increased confidence) and entry into further learning. The interviews included questions to explore participants' perceptions regarding their level of satisfaction with their course, their awareness of ESF and perceptions of additionality, i.e. did participants feel that they would have gained the same employment impact without intervention. The methodology is described in detail in Chapter 6 (pp. 48-56) and Annex I of the report (pp. 69-73).

7.6.3 Challenges

The main challenges concerned limitations on what was possible due to the data. For example, for unemployed people, the control group was very good, but for economically inactive people, there is less confidence about the control group (if an individual takes part in an ESF programme, they can be considered to be more motivated than other economically inactive individuals so it is difficult to do a fair comparison. For unemployed people, they have to be seeking employment in any case, so there is more consistency across groups).

7.6.4 Assessment

As stated above, this impact evaluation approach was trialled in 2010 to see if it would work; it did, and so every Leavers Survey since has included it. The MA plans to continue using this approach for ESF and extending it where possible to ERDF interventions, where it has been less successful to date. This intention can clearly be seen in the draft Evaluation Plan for 2014-20.

8. ANNEX II – EVALUATIONS FORESEEN IN THE CZECH PARTNERSHIP AGREEMENT EVALUATION PLAN⁴³

Table 9: Planned evaluations

Evaluation (title/topic)	Focus/scope	Timetable	Resources	Actors
Assessment of implementation – set-up of processes in OPs and in uniform methodical environment	Pilot assessment of crucial processes at the level of the MA, comparison of periods 2007-13 and 2014-20, identification of failure demand, barriers and identification of barriers of the system at the level of the MA, uniform methodical environment	October 2014 – June 2016	CZK 200,000 (c. €7,200)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> internal, with the possibility of external consultations
Assessment of effectiveness of communication activities of the NCA	Ongoing assessment of effectiveness of realised communication activities	2014 (already finished) / 2017, 2020	CZK 400,000 (c. €14,400)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Assessment of meeting of objectives linked to uniform methodical environment	Verification of meeting of objectives linked to the uniform methodical environment, e.g. administrative burden, monitoring system	3Q 2015 – December 2016	CZK 2,000,000 (c. €72,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external and internal capacities
Evaluation of the System of education in 2007-13	Ex post evaluation of the system of education of the period 2007-13	2015 / 2017, 2021 (for the system of education 2014-20)	CZK 1,000,000 (c. €36,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Meta-evaluation of evaluative activities	Evaluation of quality of evaluative outputs	2015 / further every year	CZK 0	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> internal
Ex post evaluation of 2007-13 programme period	Ex post evaluation of 2007-13 programme period, assessment of benefits at the level of selected interventions	3Q 2015 – 3Q 2017	CZK 2,500,000 (c. €90,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure

⁴³ An English version of the Czech Partnership Agreement Evaluation Plan can be accessed from this link: http://dotaceeu.cz/getmedia/aad363e8-c0a8-4702-9450-bb1d2823b172/Revised_The-Evaluation-Plan-of-the-Partnership-Agreement.pdf (last accessed 5 June 2015).

Evaluation (title/topic)	Focus/scope	Timetable	Resources	Actors
Verification of developing needs	Assessment of modifications of developing needs at the level of PA since its approval	4Q 2015 – 1Q 2017/ further 2017-19	CZK 1,000,000 (c. €36,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Assessment of meeting of objectives of the PA	Assessment of meeting of priorities of financing and expected results of the PA	4Q 2015 – August 2017/ further 2017-19	CZK 2,000,000 (c. €72,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Assessment of set-up and monitoring of synergic and complementary connections	Assessment of working of mechanisms of coordination of synergic and complementary connections and assessment of processes setting in respective guidelines and monitoring system	3Q 2015 – 1Q 2017	CZK 2,000,000 (c. €72,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external and internal capacities
Relevance of “capability approach” for assessment of relevance of ESI Funds interventions	Pilot assessment of relevance of selected interventions of the ESI Funds of “wellbeing” of relevant target groups, on the basis of the capability approach	2Q 2015 – 4Q 2016	Out of the Technical Assessment budget	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> Technology Agency of the Czech Republic
Survey on extent of fulfilment or extent of satisfaction of employees / relevant stakeholders	Survey based on analysis of three results indicators of job satisfaction, defined in the OP Technical Assistance 2014-20	3-4Q 2015	CZK 200,000 (c. €7,200)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Assessment of set-up of implementation of instruments of territorial approaches	Assessment of setting of processes of implementation of instruments of territorial approaches (integrated instruments); verification of monitoring settings and information on realisation	3Q 2016 – April 2017	CZK 500 000 (c. €18,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external and internal capacities
Evaluation of administrative capacity of the PA	Assessment of fulfilment of the methodical guideline human resources and proposals for modification of the guideline	2Q 2016 – 4Q 2016	CZK 1,000,000 (c. €36,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure

Evaluation (title/topic)	Focus/scope	Timetable	Resources	Actors
Regular territorial analysis of interventions with focus on specific topic	Analysis of attained results and effects in selected thematic areas	2016 – and further every Q one topic	CZK 0	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> internal capacities
Survey on fulfilment of results-indicators for publicity	Survey on fulfilment of three results-indicators connected to publicity that were defined in the Common communication strategy 2014-20	2016, 2018, 2019, 2013	CZK 200,000 (c. €7,200)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Survey on fulfilment of indicators of satisfaction of employees / relevant stakeholders	Survey on fulfilment of three results-indicators connected to job satisfaction, defined in the OP Technical Assistance 2014-20	3-4Q 2016	CZK 150,000 (c. €5,400)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Assessment of fulfilment and realisation of synergic chains	Assessment of set-up and working of defined synergic chains in the PA and assessment of fulfilment of synergic chains and other connections defined in the PA	3Q 2017 – 1Q 2019	CZK 3,500,000 (c. €126,900)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Verification of modification of developing needs (relevance PA)	Assessment of modifications of developing needs since the PA approval	3Q 2017 – March 2019	CZK 1,000,000 (c. €36,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Assessment of fulfilment of objectives of the PA	Omnibus assessment of fulfilment of financing priorities and expected results of the PA	4Q 2017 – August 2019	CZK 2,000,000 (c. €72,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Assessment of results of CLLD implementation	Mid-term evaluation of integrated instruments realisation	2Q 2018 – 1Q 2019	N/A	<i>Commissioning body:</i> MRD <i>Evaluator:</i> external and internal capacities
Assessment of results of implementation of ITI and Integrated territorial development plans	Mid-term evaluation of integrated instruments realisation	2Q 2018 – 1Q 2019	N/A	<i>Commissioning body:</i> MRD <i>Evaluator:</i> external and internal capacities

Evaluation (title/topic)	Focus/scope	Timetable	Resources	Actors
Assessment of results of implementation of territorial instruments	Assessment of results of implementation of territorial instruments	2Q 2018 – 1Q 2019	N/A	<i>Commissioning body:</i> MRD <i>Evaluator:</i> external and internal capacities
Horizontal impacts of selected crucial areas – base for programming 2021+	Appraisal of impacts of the PA	2019-20	N/A	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external and internal capacities
Survey on fulfilment of results-indicators for publicity	Survey on fulfilment of three results-indicators connected to publicity that were defined in the Common communication strategy 2014-20	2018, 2019, 2013	CZK 600,000 (c. €21,600)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Questionnaire survey at the level of the Czech Republic: awareness of EU funds	Survey on awareness of the Czech public of EU funds and survey on fulfilment of results-indicators for publicity	2017	CZK 2,000,000 (c. €72,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Assessment of communication activities of the NCA	Assessment of affectivity of realised communication activities of the NCA	2017	CZK 2,000,000 (c. €72,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Evaluation of the system of education 2014-20	Evaluation of the system of education.	2017	CZK 2,000,000 (c. €72,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Evaluation of administrative capacity of the PA	Assessment of fulfilment of the methodical guideline human resources and proposals for modification of the guideline	2019	CZK 1,000,000 (c. €36,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Ex post evaluation of the PA 2014-20	Ex post evaluation of the programme period 2014-20	2023/2024	CZK 3,000,000 (c. €108,000)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external and internal capacities
Assessment of effectiveness of communication activities of the MA	On-going assessment of effectiveness of realised communication activities	2017, 2020	Recommended evaluation for MAs	<i>Commissioning body:</i> MAs <i>Evaluator:</i> external, to be selected through tendering procedure

Evaluation (title/topic)	Focus/scope	Timetable	Resources	Actors
Meta-evaluation of evaluative activities	Assessment of quality of evaluative outputs, survey on questionnaires for commissioning bodies and evaluators	2015 and further every year	CZK 0	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> internal capacities
Regular territorial analysis of interventions with focus on specific topic	Analysis of attained results and effects in selected thematic areas	2017-23 - every Q one topic	CZK 0	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> internal capacities
Survey on extent of fulfilment of extent of satisfaction of employees / relevant stakeholders	Survey based on analysis of three results-indicators of job satisfaction, defined in the OP Technical Assistance 2014-20	3-4Q of every year	CZK 150,000 (c. €5,400)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> external, to be selected through tendering procedure
Report summarising conclusions of evaluations at the level of the PA	Summary of outputs from evaluations realised at the level of the PA	2022	CZK 200,000 (c. €7,200)	<i>Commissioning body:</i> MRD – NCA <i>Evaluator:</i> Internal capacities

Note: The data described in the table relate to the level of the PA, whose Evaluation Plan has already been approved. Additional evaluations will be foreseen in the EPs of individual OPs.

Source: MRD (2015) *Evaluation Plan of the Partnership Agreement for programming period 2014-2020*, approved in February 2015, available at <http://www.strukturalni-fondy.cz/cs/Fondy-EU/Narodni-organ-pro-koordinaci/Evaluace/Evaluacni-plan> (last downloaded 30 March 2015).