

EPRC

EUROPEAN POLICIES RESEARCH CENTRE



Evaluating the Effectiveness of Regional Policy



EoRPA Paper 14/6

Laura Polverari, John Bachtler and Arno van der Zwet

European Policy Research Paper
Number 91

Published by:

European Policies Research Centre

University of Strathclyde

40 George Street

Glasgow G1 1QE

ISBN: 1-871130-97-2



March 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 3061
Fax: +44 (0) 141 548 4898

e-mail: john.bachtler@strath.ac.uk
fiona.wishlade@strath.ac.uk

The place of useful learning

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

Preface

This report was prepared by the European Policies Research Centre (EPRC) under the aegis of EoRPA (European Regional Policy Research Consortium), which is a grouping of national government authorities from countries across Europe. The Consortium provides sponsorship for EPRC to undertake regular monitoring and comparative analysis of the regional policies of European countries and the inter-relationships with EU Cohesion and Competition policies. Over the past year, EoRPA members have comprised the following partners:

Austria

- Bundeskanzleramt (Federal Chancellery), Vienna

Finland

- Työ- ja elinkeinoministeriö (Ministry of Employment and the Economy), Helsinki

France

- Commissariat Général à l'Egalité des territoires (General Commissariat for Territorial Equality, CGET, previously DATAR), Paris

Germany

- Bundesministerium für Wirtschaft und Energie (Federal Ministry for the Economy and Energy), Berlin
- Ministerium für Wirtschaft, Bau und Tourismus, Mecklenburg-Vorpommern (Ministry for the Economy, Construction and Tourism, Mecklenburg-Western Pomerania), Schwerin

Italy

- Ministero dello Sviluppo Economico (Ministry of Economic Development), Dipartimento per lo sviluppo e la coesione economica (Department for Development and Economic Cohesion), Rome

Netherlands

- Ministerie van Economische Zaken (Ministry of Economic Affairs), The Hague

Norway

- Kommunal- og moderniseringsdepartementet (Ministry of Local Government and Modernisation), Oslo

Poland

- Ministerstwo Infrastruktury i Rozwoju (Ministry of Infrastructure and Development), Warsaw

Sweden

- Näringsdepartementet (Ministry of Enterprise, Energy and Communications), Stockholm

Switzerland

- Staatssekretariat für Wirtschaft (SECO, State Secretariat for Economic Affairs), Bern

United Kingdom

- Department for Business, Innovation and Skills, London
- The Scottish Government, Enterprise, Transport and Lifelong Learning Department, Glasgow

The research for the country reviews was undertaken by EPRC in consultation with EoRPA partners. It involved a programme of desk research and fieldwork visits among national and regional authorities

in sponsoring countries during the first half of 2014. The EoRPA research programme is coordinated by Professor John Bachtler, Fiona Wishlade, Dr Sara Davies and Heidi Vironen.

This paper should be referred to as: L. Polverari, J. Bachtler and A. van der Zwet (2014) *Evaluating Effectiveness of Regional Policy*, EoRPA Paper 14/6, Paper prepared for the 35th meeting of the EoRPA Regional Policy Research Consortium at Ross Priory, Loch Lomondside, 5-7 October 2014

Country-specific research was contributed by the following research team.

Austria: Stefan Kah, EPRC

Norway: Fiona Wishlade, EPRC

Finland: Heidi Vironen, EPRC

Poland: Dr Martin Ferry, EPRC

France: Frederike Gross, EPRC

Sweden: Heidi Vironen, EPRC

Germany: Dr Sara Davies, EPRC

Switzerland: Stefan Kah, EPRC

Italy: Dr Laura Polverari, EPRC

United Kingdom: Rona Michie and Dr Martin Ferry, EPRC

Netherlands: Dr Arno Van der Zwet, EPRC

Many thanks are due to everyone who participated in the research. Thanks also to Dr Keith Clement, Lynn Ogilvie and Alyson Ross for editorial, coordination and secretarial support respectively, and to Dr Andrew Judge for research assistance. In addition, the European Policies Research Centre gratefully acknowledges the financial support provided by the members of the EoRPA Consortium.

Disclaimer: It should be noted that the content and conclusions of this paper do not necessarily represent the views of individual members of the EoRPA Consortium.

TABLE OF CONTENTS

1. INTRODUCTION	3
2. THE EVALUATION OF REGIONAL POLICY.....	5
2.1 Overall importance of evaluation.....	5
2.2 Who demands evaluation and why?	6
2.3 Narrow versus integrative approaches to evaluation	9
3. IMPLEMENTATION AND METHODS	15
3.1 Who commissions and undertakes evaluation?.....	15
3.2 Scope	16
3.2.1 Project-level evaluation – Measuring the impact of the New Regional Policy projects in Switzerland	17
3.2.2 Instrument-level evaluation: Evaluating the Contratti di Programma in Italy	18
3.2.3 Instrument-level evaluation: Evaluating the Selective Financial Assistance in Northern Ireland	19
3.2.4 Programme-level evaluation: Evaluation of the Regional Strategic Programmes of West Finland and South-West Finland	21
3.2.5 Theme-level evaluation: Evaluating financial engineering instruments in Poland	22
3.3 Focus.....	23
3.4 Methods.....	25
4. CONCLUSIONS	33
4.1 The limits of existing approaches.....	33
4.2 The impact of evaluation on policy.....	33
4.3 Measures to improve evaluation and future plans	34
4.4 Questions for discussion	35

Executive Summary

This paper provides a review and assessment of the current the state of play in regional policy evaluation in the context EoRPA countries, against the backdrop of persisting budget cuts and policy reform. The paper discusses:

- The importance attached to the evaluation of regional policy effectiveness in EoRPA countries;
- How the evaluation of effectiveness is interpreted and operationalised;
- The limits of existing approaches and how these are being addressed;
- The evaluation of effectiveness as a tool for improving the policy-making process and the effectiveness of regional policy;
- Emerging trends in regional policy evaluation.

There is a strong emphasis on regional policy evaluation in most EoRPA countries and a general consensus that regional policy evaluation will be high on the policy-maker agenda in 2014-20. Reasons include: an increase in the resources devoted to evaluation; a spillover effect from Cohesion policy onto domestic policy; improved skills and capabilities related to evaluation; concerns about the financial scale and redistributive effects of regional policies, and a stronger role for evaluation across all policy fields. However, in 2013-14, evaluation has sometimes been seen as less important than the need to finalise Cohesion policy programmes and regional aid maps for 2014-20; to address spending delays; and to introduce anti-crisis measures and structural reforms.

EoRPA countries vary in terms of their views of evaluation and the degree to which evaluation is undertaken systematically. External evaluation in the strict sense is carried out virtually everywhere. In some cases, however, this occurs in a more disjointed fashion than would be needed to provide a comprehensive view of policy effectiveness and impact. This appears to be particularly the case where policy is devolved to sub-national authorities and lacks strong coordination or oversight from the centre (e.g. Italy, England). In some countries, evaluation is part of a wider and composite framework of knowledge acquisition that includes not just evaluations by external independent evaluators, but also a range of further sources, such as auditing, reporting, internal reviews, external peer reviews and 'What works centres' (e.g. Germany, the United Kingdom or the Netherlands).

Responsibility for evaluations is in line with the broader allocation of responsibility for policy design and implementation. Evaluations are commissioned predominantly at national level in the Netherlands, Norway, Sweden and the United Kingdom (England and Scotland), and at sub-national level in Austria. In Finland, France, Germany, Italy and Poland, evaluations are commissioned by both national/federal and *Land*/regional levels.

Most evaluation work is entrusted to actors who are external to the public administration, primarily consultants and academics, largely on grounds of objectivity/independence and specialist skills. However, a fairly recent trend is the internalisation of evaluation, for instance within dedicated evaluation units (in Italy and Scotland) or government departments (in France and Austria). Policy-

makers may feel that they have a good understanding of a policy or scheme, and this can limit perceptions of the added value of external evaluators (e.g. in Finland and Germany).

Most evaluations in EoRPA countries focus on the effectiveness, impact and implementation of programmes. Value-for-money considerations are particularly relevant in the United Kingdom (England), Germany, the Netherlands and Norway, whilst utility is also appraised in Italy and Sweden. A number of countries also evaluate implementation arrangements. There is variation across evaluations in terms of the definition of effectiveness and e.g. in relation to treatment of unintended effects, timeframe (short-term or long-term), and causality. This has implications for the type of lessons that can be drawn from evaluations.

A variety of methodologies is used, notably: macro-economic modelling, econometric statistical analysis, micro-econometric counterfactual analysis, qualitative case studies, and theory-based evaluation. Each has advantages and disadvantages. Meta-evaluations (increasingly used in Poland) are useful in presenting evidence in a format that addresses the knowledge needs of policy-makers.

Constraints on the evaluation of effectiveness and on the influence of evaluation on policy include: institutional issues (e.g. lack of understanding of 'evidence-based' policy, short-term mind-sets, political reluctance, or scepticism), systemic shortcomings (e.g. lack of data or evaluation capacity), and scarce resources.

Although evaluation generally has only a limited influence on regional policy, there are examples of impact, e.g. with regard to the closure of policy instruments, the adjustment of project selection procedures, and institutional changes. A key challenge with regard to effectiveness evaluation is that it is difficult to ensure that evaluations provide a full picture of policy effects.

Evaluation is likely to be high on the regional policy agenda in 2014-20 because of the expectations of EU Cohesion policy and the EU Regional Aid Guidelines, and a domestic political drive in a number of countries (notably France and Poland). EoRPA countries are investing in data and methods, introducing organisational arrangements and using more experiential knowledge as part of their evaluation efforts. Against this background, the paper proposes the following questions for discussion:

- What kind of evaluations (conceptual approaches, methods, outputs) are most useful for regional policy-makers?
- What would favour a better integration of evaluation activities into the policy cycle (e.g. in terms of organisational arrangements or capacity-building activities)?
- Two trends identified in this report are a shift towards integrated approaches and internalising evaluation activities. What are the risks and benefits of these trends?
- Regional policy-makers must make judgements based on imperfect and at times contradictory evidence: Is the use of experiential knowledge valuable for assessing the effectiveness of regional policy?

1. INTRODUCTION

The past two decades have seen increasing attention being given to the evaluation of regional policy at both European and national levels.¹ As a previous EoRPA noted, 'evaluation is a well-established component of the policy process in a number of countries and continues to be used for regional policy management and accountability, feeding into adjustments of on-going interventions and informing future plans'.² The growing importance accorded to evaluation has been driven by the need to justify public spending as well as an interest in understanding 'what works' in regional policy.

This paper provides a review and assessment of the current the state of play in regional policy evaluation in the context EoRPA countries, against the backdrop of persisting budget cuts and policy reform. The paper addresses the following questions:

- How much importance is attached to the evaluation of the effectiveness of regional policy in the EoRPA countries?
- How is the evaluation of the effectiveness interpreted and operationalised?
- What are the limitations of the approaches adopted and how are they being addressed?
- How effective is evaluation in terms of its ability to improve the policy-making process and, ultimately, the effectiveness of regional policy programmes?
- What are the emerging trends in regional policy evaluation?

In investigating these questions, the report provides illustrative examples of evaluations recently conducted in the EoRPA countries. The paper concludes with some questions for discussion at the EoRPA annual meeting 2014.

¹ Bachtler J and Wren C (2006) The Evaluation of EU Cohesion Policy: Research Questions and Policy Challenges, *Regional Studies*, 40(2), Special Issue: The Evaluation of EU Cohesion Policy, 143-153; Bachtler J, Polverari L, Oraže H, Clement K and Tödtling-Schönhofer H, with Gross F and Naylor I (2009) Ex post evaluation of Cohesion policy programmes 2000-2006 co-financed by the ERDF (Objective 1 and 2), Work Package 11: Management and implementation systems for Cohesion policy, Final Synthesis Report to the European Commission (DG Regio), Brussels; Olejniczak J, Kozak M and Bienias S (2011) *Evaluating the effects of regional interventions. A look beyond current Structural Funds' practice*. Ministry of Regional Development, Warsaw 2011 Polverari L and Bachtler J (2004) Assessing the Evidence: The Evaluation of Regional Policy in Europe, EoRPA Paper 04/05, European Polices Research Centre, University of Strathclyde, Glasgow; Hoerner J and Stephenson P (2012) Theoretical perspectives on approaches to policy evaluation in the EU: The case of cohesion policy, *Public Administration*, 90(3), 699-715.

² Bachtler *et al.* (2013) *op. cit.* 41.

2. THE EVALUATION OF REGIONAL POLICY

2.1 Overall importance of evaluation

European countries have different traditions in terms of what constitutes regional development policy, and similarly different traditions in relation to the evaluation of public policy and of regional policy specifically. For some, evaluating public policy is a long-established practice in which evaluation is embedded in the policy cycle (e.g. Germany, the Netherlands, the UK); for others, policy evaluation is more recent practice, in some cases stimulated by EU accession (e.g. Austria, Italy and Poland, but also in Switzerland).

In recent years, evaluation in the regional policy context has been strengthened in virtually all EU Member States. The reasons for this trend include the following:

- an increase in the **resources** devoted to evaluation, derived from the progressive financial strengthening of EU Cohesion policy and, in parallel, the progressive increase in Technical Assistance budgets;
- a **spillover effect** of increasing EU evaluation requirements on domestic systems, whereby the EU Cohesion policy regulations have been deemed to have influenced national policies, for example by raising the profile of evaluation³ or by generating a 'virtuous circle' of evaluation in the Member States;⁴
- a widening of the **evaluation communities** and improvement of the **skills and capabilities** related to evaluation, as a result of dedicated efforts in this direction by the European Commission and domestic authorities, which have recently found renewed impetus as a result of the enlargements of the EU;⁵
- concerns among policy-makers to justify the **financial scale and redistributive effect** of regional policies, leading to demands for evidence of the effectiveness and value for money of spending;⁶
- a more pronounced **role assigned to public policy evaluation** *per se* under the New Public Management agendas pursued within many European countries.⁷

³ Bachtler J (2001) Quod erat demonstrandum? The Evaluation of Regional Policy, in H-F Eckey, D Hecht, M Junkernheinrich, H Karl, N Werbeck and R Wink (eds) *Ordnungspolitik als konstruktive Antwort auf wirtschaftspolitische Herausforderungen – Festschrift zum 65. Geburtstag von Paul Klemmer*, Lucius & Lucius Verlag, Stuttgart, 3-20; Bachtler and Wren (2006) *op.cit.*; Aalbu H (1998) The Structural Funds as an agent for system changes: a mid-term evaluation Sweden's Objective 6 programme, Nordregio Working Paper No. 2, Stockholm; Barca F (2006a) European Union evaluation between myth and reality: Reflections on the Italian experience, *Regional Studies*, 40(2), 273–276, Barca F (2006b) *Italia Frenata. Paradossi e lezioni della politica per lo sviluppo* (Donzelli, Roma); Huber W (2006) Evaluation of EU Cohesion policy: window-dressing, formal exercise or coordinated learning process? *Regional Studies*, 40(2), 277-280.

⁴ Geddes (2006); Bachtler *et al.* (2009) *op. cit.*

⁵ Polverari and Bachtler (2004) *op.cit.*

⁶ Bachtler and Wren (2006) *op.cit.*; Hoerner and Stephenson (2012) *op.cit.*

⁷ In parallel, there has also been a general rise in independent academic interest in the appraisal of policy outcomes and effectiveness, linked to the disappointing socio-economic performance of the European regions despite the considerable resources spent: with such significant investment, why have regional disparities not been re-absorbed? See as examples the work of Andres Rodriguez-Pose (LSE), Guido Serravalli (Parma), Sasha Becker (Warwick) and many others (for a review, see also Polverari and Bachtler, 2014). However, the present paper focuses on evaluation intended as a government-led activity. Independent academic studies, which are prominent in some countries (e.g. Germany, Italy and the UK), are left out of the investigation. There is no clear-cut evidence regarding the actual impact that such studies have on the policy process.

These trends can be observed ex post over a longer-term perspective. However, from a more recent viewpoint, the evolution of evaluation in the field of regional policy is less clear-cut. On the one hand, the recent austerity and budget reductions have led to an increased demand for accountability about the results of policy and instruments, and about value for money of investments (for example in France and the UK). On the other hand, however, a reduction of evaluation activities has also been noted in practice in Austria and, to an extent, Italy. In these countries, attention has shifted to more pressing needs, such as: focusing on *ad hoc* measures to counter the impact of the economic crisis; concentrating efforts on the structural reforms required by fiscal compact and 2014-2020 Cohesion policy conditionalities; and tackling spending delays to ensure the absorption of 2007-13 programmes and manage the stringent timetable for the finalisation of the new Cohesion policy programmes. Against these more pressing needs, evaluation has become a secondary goal.

Overall, EU Cohesion policy has acted not just as a catalyst for evaluation, but also as a vehicle for disseminating standards and methods across the EU Member States.⁸ This has had a marked impact on the evaluation cultures of Member States, especially those that had less extensive evaluation traditions prior to their accession to the EU. The evaluation requirements of Cohesion policy and the accompanying capacity-building activities have had a considerable harmonisation effect in terms of creating a common understanding of evaluation as a useful and necessary tool for improving the policy, so much so that today the actual weight assigned to evaluation in different countries is generally high. All EoRPA countries consider the evaluation of regional policy as a priority (see Table 1). The reality of how this priority is translated into practice varies, but there is unanimity that evaluation is an important and necessary aspect of the policy cycle.

2.2 Who demands evaluation and why?

The demand for evaluation comes from the policy-making and political levels; only rarely is there also demand from stakeholders and the public at large. The policy demand is driven by various factors, both endogenous and exogenous to the policy. The need to deal with actual or expected spending cuts has been an important impetus for evaluation in Finland, Poland and France. In Finland, in the wake of reduced funding, attention has been placed increasingly on understanding what works in the context of regional policy. In Poland, where the attention is already shifting to the post-2020 period, awareness that regional policy will be funded largely by domestic resources is leading the Polish authorities to want to understand the effectiveness of different types of interventions, so as to inform policy directions after 2020. From a more systemic perspective, in France the change of government in 2012 saw the new administration engaging in a large-scale public policy modernisation plan (*Modernisation de l'action publique*) that has evaluation at its heart. Part of the impetus was due to a desire to respond to public concerns over the previous government's General Public Policy Review (*Revue générale des politiques publiques*), accused of having introduced cuts without adequate consideration of evidence or indeed stakeholder involvement.

In other cases, the demand for evaluation is driven by statutory requirements. In Norway, for example, the Ministry of Local Government and Modernisation has to report annually to Parliament, and evaluations are used to support the process.⁹ In the United Kingdom (England), the 2010 evaluation strategy of the Department for Business, Innovation and Skills (BIS) commits the department to

⁸ Bachtler and Wren (2006) *op.cit.*; Polverari and Bachtler (2004) *op.cit.*

⁹ A similar requirement exists in Italy, but the role of evaluation in the process has been declining in the past few years, with the report focusing especially on providing an overview of the financial and implementation aspects.

evaluating major policies every five years (although there have been implementation delays with planned evaluations).¹⁰ In Finland, Germany and the Netherlands, the commitment to monitor and evaluate programmes is embedded in the Regional Development Law, GRW coordination framework and national legislation on Government Accounts respectively. Another driver, in Switzerland, has been the desire to test the appropriateness of a new policy approach (the New Regional Policy introduced in 2008), providing evidence about whether it is working and on how it could be further improved. Lastly, in countries such as Germany and the Netherlands, evaluation is also performed because it is a routine procedure of the policy process, and it is generally taken for granted that it should be undertaken.

¹⁰ According to a 2013 National Audit Office (NAO) report, as of October 2013, no evaluation programme had yet been published, and so far BIS has published only three of the five evaluations that it committed to in the 2010 strategy. NAO (2013) Evaluation in government, http://www.nao.org.uk/wp-content/uploads/2013/12/10331-001-Evaluation-in-government_NEW.pdf (last downloaded 20 September 2014).

Table 1: Weight assigned to regional policy evaluation

Country	Weight currently assigned to evaluation	Driver	Current demand	Recent trend
Austria	Medium – currently, focus on spending 2007-13 Cohesion policy allocations rather than evaluating policy.	There is stakeholder interest, but the main impetus is represented by the EU Regulations .	Low	Less prominent in 2007-13, but set to increase again in 2014-2020.
Finland	High – evaluation institutionalised in domestic regional development law and high political and stakeholder interest to understanding effects of policy instruments in a context of reduced funding.	Reduced resources focused attention on ‘what works’.	High – political and stakeholder	On the increase.
France	High – evaluation institutionalised as part of wider reform for the Modernisation of Public Policy, whose main ambition is to enhance public policy effectiveness.	Controversy over spending cuts introduced by past government (General Public Policy Review) and budget cuts heightened political and public interest on public policy evaluation.	High – political and public	On the increase. Limited until 1990s, increased since 2000s, and markedly increased since 2012.
Germany	High – long-established evaluation culture, institutionalised in the GRW Coordination Framework	Legislation, political and public awareness, past evaluation tradition.	High – political (e.g. Land and Federal Parliaments), administrative and stakeholder	Will continue to be high on the agenda.
Italy	Medium – tradition of evaluating business aids and steady increase of programme evaluation thanks to EU CP. However, current focus is on absorption, crisis and preparation of new programmes.	EU Regulations. Political and policy-maker interest (in the regions).	Variable – low at the national level, high (political and/or administrative) in some regions	Less prominent nationally in 2007-13, but activism in some regions (e.g. Campania, Sicily, Apulia, Lombardy).
Netherlands	High – established evaluation tradition where evaluation is embedded in policy-making process.	Domestic legislation on government accounts, past evaluation tradition.	High – political and administrative	On the increase. Over past decade, more emphasis on evaluation and more formalisation of evaluation system.
Norway	High – strong evaluation culture, continuous evaluation embedded in policy-making.	Political and legislative – Ministry has to report annually to Parliament and evaluations are used to support the process.	High – political and administrative	Will continue to be high on the agenda.
Poland	High – virtually non-existent prior to EU accession, but rapidly and continuously growing.	Size of funding thus far and, prospectively, awareness of post-2020 EU funding cuts (effectiveness of different interventions).	High – political and administrative	On the increase.
Sweden	High and embedded in regional growth policy (increased in 2007-13).	Policy-makers: how to improve policy delivery (learning evaluation).	Medium – policy-maker, not so much politicians, stakeholder or public	On the increase.
Switzerland	High and increased since the introduction of the New Regional Policy (NRP) in 2008.	Understanding NRP impacts and related projects.	High – politicians and administration	On the increase.
UK – England	Medium – high importance in principle, but with implementation delays.	Political and policy-maker interest.	High – political, administrative and public	Less prominent since 2010, but supplemented by other sources.
UK – Scotland	Medium – evaluation mostly on implementation/process issues and confined to Cohesion policy.	EU Regulations and policy-maker demand for knowledge on how to improve the programmes.	Medium – political agenda focussed elsewhere	Continuity with 2007-13 as a process largely internal to the SG administration.

Source: EoRPA research.

2.3 Narrow versus integrative approaches to evaluation

Notwithstanding the common view that evaluation is high on the agenda, EoRPA countries vary in relation to how they view evaluation and the degree to which evaluation is undertaken systematically and consistently.

External evaluation in the strict sense is carried out virtually everywhere. In some cases, however, this occurs in a more disjointed fashion than would be necessary to allow a comprehensive view of the effectiveness and impact of the policy in a given country. This appears to be particularly the case where policy is devolved to sub-national authorities and lacks strong coordination or oversight from the centre. In Italy, for example, the central Evaluation Unit has been employed over the past couple of years in supporting the new programming work related to the 2014-20 Partnership Agreement (PA) and Operational Programmes, for instance coordinating the PA ex-ante evaluation, defining the operational framework for the results-orientation of programmes, and drafting parts of the PA. By necessity, this has entailed a reduced emphasis on evaluation activities (though a lot of the programming work, for example in the area of urban development, could rely on previous evaluation activities). At the same time, some of the regional authorities have been active in pursuing their own independent evaluation agendas, including undertaking ex-post evaluations of programmes, fields or schemes (examples include the ex-post evaluation of the 2000-06 ROP in Campania, the ex-post evaluation of the aid scheme 'programming contracts' in Apulia, the ex-post evaluation of the 2000-06 ROP impact on transport infrastructure in Sicily, various evaluations in Lombardy and many others). There is no national overview of these evaluations, however.

Similarly in England, BIS has noted that it has been impossible to develop a coherent narrative across the UK of the effects and impacts of the 2007-13 Cohesion policy programmes, because evaluation was carried for each individual programme/fund/region separately during that period, with no common national evaluation framework. This has prompted the UK Programme Board that was put together for the development of the Partnership Agreement to ask the Performance Management Group to look at evaluation and what can be put in place to achieve increased coordination between nations and funds. This work has just started.

In some countries, evaluation is part of a **wider and composite framework** of knowledge acquisition that includes not just evaluation work from external independent evaluators, but also a range of further sources, such as **auditing, reporting, internal reviews and external peer reviews** (e.g. Germany, the United Kingdom or the Netherlands).

Germany is a prime example. The Joint Task for the Improvement of Regional Economic Structure (*Gemeinschaftsaufgabe Verbesserung der Regionalen Wirtschaftsstruktur*, GRW) coordination framework includes formal evaluations undertaken regularly by external experts (e.g. on business aids, as illustrated in case study Box 5, in Section 3.4), focused on whether and to what extent the GRW has contributed to the most important regional policy goals and on the actual impact of funding compared to what would have been the case in the absence of GRW funding (counterfactual situation). In addition, these evaluations are supported by a comprehensive set of reporting obligations by the *Länder* to the federal authorities that ensure that the statistical base on which the evaluations rely is solid (e.g. central database with data on GRW recipients, reports on the *Länder* legal frameworks and regional policy goals, reporting on project-level commitments, expenditure and clawed-back funds etc.). These data are complemented by the information provided by the audits

undertaken by federal and *Land* authorities, which also provide information that is useful for appraising the effectiveness and impact of GRW funding.

The UK (England) provides an interesting case of an integrative approach to evaluation. Although evaluation activity within central government departments in England has been less visible since the 2010 election, there have been various parallel, external sources of scrutiny. One of the main bodies with an active role in examining government policy in England is the National Audit Office (NAO), which scrutinises public spending on behalf of the UK Parliament to hold government departments to account and support public service managers in improving the performance of public policies (Parliament Committees also publish relevant reports). The NAO audits the financial statements of all government departments and a wide range of other public bodies, and it also produces around 60 value-for-money reports per year. In recent years, the NAO and the UK Parliament have had a strong interest in local growth policy and have published the following reports:

- National Audit Office (2014) Progress Report on the Regional Growth Fund;¹¹
- National Audit Office (2013) Funding and Structures for Local Economic Growth;¹²
- House of Commons Business, Innovation and Skills Committee (2013) Local Enterprise Partnerships;¹³
- House of Commons Committee of Public Accounts (2012) The Regional Growth Fund;¹⁴
- National Audit Office (2012) The Regional Growth Fund;¹⁵
- National Audit Office (2010) Regenerating the English Regions: Regional Development Agencies' Support to Physical Regeneration Projects;¹⁶
- House of Commons Business, Innovation and Skills Committee (2010) The New Local Enterprise Partnerships: An Initial Assessment.¹⁷

NAO reports take an audit approach with an emphasis on value for money, using a range of quantitative and qualitative techniques, including financial analysis, interviews and case studies, document review, literature review, economic analysis (efficiency analysis and benchmarking against comparators), peer review and expert panels, as summarised in the diagram in Figure 1 below.

¹¹ <http://www.nao.org.uk/wp-content/uploads/2015/02/Progress-report-on-the-regional-growth-fund.pdf> (last downloaded 20 September 2014).

¹² <http://www.nao.org.uk/wp-content/uploads/2013/12/10285-001-Local-economic-growth.pdf> (last downloaded 20 September 2014).

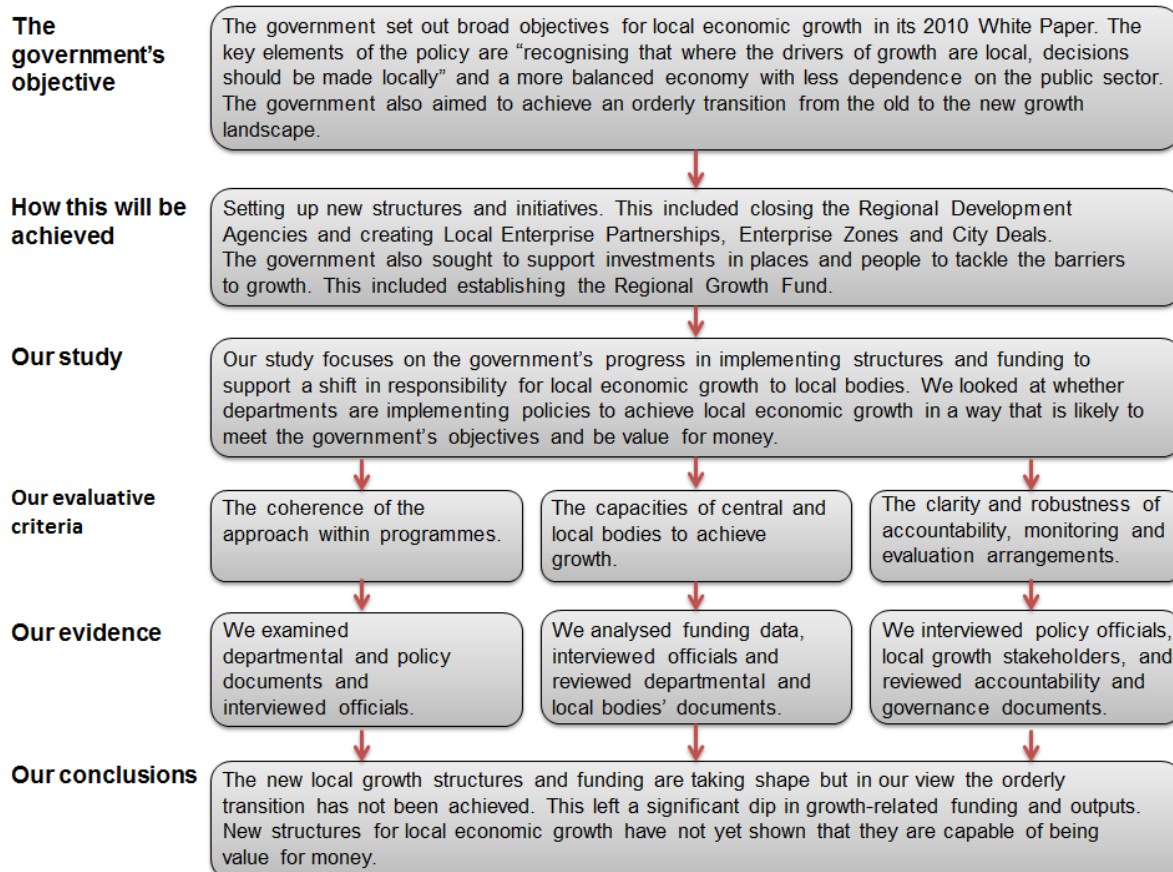
¹³ <http://www.publications.parliament.uk/pa/cm201213/cmselect/cmbis/598/598.pdf> (last downloaded 20 September 2014).

¹⁴ <http://www.publications.parliament.uk/pa/cm201213/cmselect/cmpubacc/104/10402.htm> (last accessed 20 September 2014).

¹⁵ <http://www.nao.org.uk/wp-content/uploads/2012/05/121317.pdf> (last downloaded 20 September 2014).

¹⁶ <http://www.nao.org.uk/wp-content/uploads/2010/03/0910214.pdf> (last downloaded 20 September 2014).

¹⁷ <http://www.publications.parliament.uk/pa/cm201011/cmselect/cmbis/434/434i.pdf> (last downloaded 20 September 2014).

Figure 1: NAO approach to auditing the ‘Funding and structures for local economic growth’

Source: Adapted from <http://www.nao.org.uk/wp-content/uploads/2013/12/10285-001-Local-economic-growth.pdf> (last downloaded 23 September 2014), p. 41.

Further, taking inspiration from the United States,¹⁸ in the UK, the past four years have seen the launch of six 'What Works Centres', responsible for synthesising evaluation evidence on the effectiveness of policy in a range of fields, with five thematic centres in England, including one on local growth ('What Works Centre for Local Economic Growth', described in Box 1 below)¹⁹ and one in Scotland on public service delivery and reform across all public services ('What Works Scotland Centre', Box 2). The objective of the Centres is to improve the evidence that government can use in making decisions about public services. They do so by producing, synthesising and disseminating research to local decision-makers to support them in investing in those services and those delivery mechanisms that determine the best outcomes for citizens and value for money for taxpayers.

The What Works Centres realise a fundamental conceptual shift from traditional external evaluation to a more systemic approach to evidence-based policy-making. This shift recognises that the evidence upon which public policy choices rest is diverse, contestable and hampered by practical limitations that affect its quality (e.g. data, methods and timetable of availability) and, at the same time, that it is essential to involve in the production and use of evidence not just politicians and policy-makers, but also frontline practitioners and end-users. Through their work, the Centres are intended to favour a reconciliation between multiple sets of evidence – expert, experiential (from practitioners and end-users) and political – and thus enable a design and delivery of public services that is both efficient and responsive.

¹⁸ Following the lead of evidence-assessment bodies in the medical field, the US Department of Education was one of the first to create an institutional base for public policy, with a WW Clearing House in 2002 as a resource for informed education decision-making.

¹⁹ The other themes are health and social care, education attainment, ageing, crime reduction and effective early intervention. See <https://www.gov.uk/what-works-network> and <http://www.esrc.ac.uk/collaboration/public-sector/what-works-evidence-centres.aspx>.

The impact of the What Works centres is as yet unclear, but the approach is being closely monitored by some other countries. Furthermore, at a time of budget cuts and austerity measures, it has been noted that What Works Centres represent good value for money, in particular when compared to carrying out primary research, as much of their work is based on making use of existing evidence.²⁰

Box 1: What Works Centre for Local Economic Growth in England

The What Works Centre for Local Economic Growth is a collaboration between the London School of Economics and Political Science (LSE), Centre for Cities and Arup, and it is funded by the Economic and Social Research Council (ESRC), Department for Communities and Local Government and BIS. The Centre has so far produced three reports, examining business advice (looking at evaluations of business information, advice and mentoring programmes), employment training (looking at evaluations of programmes aiming to improve adult skills and labour market outcomes), and the economic impact of major sports and cultural events.

Each evidence review assesses available evidence to find evaluations that are robust and clearly identify policy impact, using a five-stage process:²¹

- Scope of review: A User Panel and Academic Panel agree review question, key terms and inclusion criteria using existing literature reviews and meta-analyses.
- Searching for evaluation: Evaluation evidence is collected using a wide range of sources, from peer-reviewed academic research to government evaluations and think-tank reports. This review includes academic databases, specialist research institutes, UK central and local government departments, and think tanks. The process also includes a call for evidence.
- Sifting evaluations: The full set of evidence is refined based on relevance and the robustness of the research method using the Scientific Maryland Scale:
 - o Level 1 - Correlation of outcomes with presence or intensity of treatment, cross-sectional comparisons of treated groups with untreated groups, or other cross-sectional methods in which there is no attempt to establish a counterfactual.
 - o Level 2 - Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention ('before and after' study).
 - o Level 3 - Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention, and a comparison group used to provide a counterfactual (e.g. difference in difference).
 - o Level 4 - Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention, and a comparison group used to provide a counterfactual (i.e. difference in difference).
 - o Level 5 - Reserved for research designs that involve randomisation in treatment and control groups.
- Scoring evaluations: Provide a full appraisal of each evaluation on the shortlist, collecting key results and using the SMS to give a final score that reflects the quality of method and quality of implementation.
- Synthesising evaluations: Conclusions drawn are based on a combination of these findings and existing literature.

Source: <http://whatworksgrowth.org/>.

²⁰ Rutter T (2013) What Works policy labs: are they working? – round up, *the Guardian*, available on: <http://www.theguardian.com/public-leaders-network/2013/dec/03/what-works-policy-labs-working-roundup> (last accessed 20 September 2014).

²¹ Graphic showing the process available here: <http://whatworksgrowth.org/our-work/about-these-reviews/#.U6rS4EC4M88> (last accessed 20 September 2014).

Box 2: What Works Scotland (WWS)

The What Works Scotland Centre is a new centre, hosted by the Universities of Glasgow and Edinburgh, whose aim is to support policy-makers and practitioners in reforming public service delivery in Scotland. It is supported by the Economic and Social Research Council (ESRC) and the Scottish Government (c. £3 million) with match-funding from the Universities of Glasgow and Edinburgh (c. £700,000).

The Centre was launched in June 2014. It will collaborate with the Community Planning Partnerships and engage with stakeholders, to:

- elaborate data, evidence and the application of improvement methodologies in order to support service performance across a range of fields including: health, economics, inequality, employment, social care, criminal justice, asylum, migration, housing, and urban renewal;
- engage with stakeholders, to develop services that are shaped and co-produced by both service providers and the citizens and communities who utilise them;
- build on the strengths and assets of individuals and communities.

Source: http://www.esrc.ac.uk/news-and-events/press-releases/30919/What_Works_Scotland_to_drive_scale_and_pace_of_public_service_delivery_and_reform.aspx.

Another example of an integrative approach is represented by the Netherlands, where, in addition to the evaluation work in the strict sense, including a recent territorial review assessing sub-national policies and governance challenges commissioned by the OECD, the government relies on annual audit reports produced by the Dutch Court of Auditors, so-called 'EU Trend Reports'. Published in February each year, the report provides insights into the financial management of EU Funds in the EU Member States with a particular focus on the Netherlands. It also provides an analysis of the effectiveness and efficiency of ERDF funds in the Netherlands. The main objectives are to: (i) inform the Minister of Finance and assist him in forming an opinion on the EU budgetary issues; (ii) inform the debate in Parliament concerning EU finances; and (iii) inform the public on EU finances. Based on monitoring data from EU programmes and on qualitative assessments, it provides recommendations that can directly feed into the policy-making cycle (see Box 3 for more detail on this year's report) and is generally followed up by government, including through a letter where the government outlines its response to the recommendations contained in the report.

Box 3: 2014 EU Trend Report – The Netherlands

The 2014 EU Trend Report provides the following recommendations:

- More transparency, so that programmes and projects are forced to demonstrate to the public why certain projects are carried out and what the effects of these projects are (as for example Europa om de Hoek).²²
- Develop general standards for recurring cost components in order to compare the effectiveness of projects.
- When considering value for money component, not only consider the ERDF part of the project funding but also the other components of the co-financing.
- Target values should be more related to the desired effect and less to intermediate results of the projects. Indicators should also be closely connected to the final intended result.
- More objective target values that have a clear rationale and generalise the definition of the indicator.
- Promote competition between project applications in order to increase chances that the most effective projects are selected (rather than first come, first served).
- Use implementation contracts.

The report also considered the impact of some measures taken for the 2014-20 programme period that aim to increase effectiveness. For example, it notes concern over the potential perverse effect of the performance reserve, which, based on the experience of the 2000-06 period, is viewed as potentially encouraging programmes and projects to set low(er) targets. However, it provides a positive assessment of the new programme indicators, which are considered better able to track programme achievements and effectiveness, as opposed to simply reviewing who obtains funding.

The 2014 EU Trend Report also outlined a number of challenges that hinder the degree to which the effectiveness of EU programmes can be appraised, notably:

- **Lack of clarity in terms of indicators:** First, the indicators and targets that measure if projects contribute to the objectives are determined at the regional level and are linked to EU goals. In most programmes, it was unclear whether these indicators were set based on achievements in previous programme periods. Second, project beneficiaries are asked to provide project targets that contribute to the programme's targets. The way in which these targets are set is mostly unknown. Thirdly, Management Authorities have developed 'calculation tools' to increase the reliability of targets set by beneficiaries. However, the development of calculation tools has relied on targets already provided by beneficiaries.
- **Excessive caution in defining targets:** The target values for indicators that the ERDF Management Authorities set at the programme level are often very largely exceeded. This raises questions on whether the target values for the indicators are set at realistic levels. Accordingly, the target values are not a good indicator of the effectiveness of the intervention.
- **Adjustment of indicators in itinere:** In some instances, indicators are changed mid-programme, which limits the usefulness and comparability of the data.
- **Too broad indicators:** The indicators are not always useful to measure the effectiveness of interventions at the regional or programme levels, as they are very broadly defined. This is particularly problematic when appraising the cost-effectiveness (efficiency) of projects.
- **Vague definition and variable interpretations of value for money:** Although efficiency and effectiveness are on paper a key concern, in practice there does not always seem to be a great focus on these issues. When calculating value for money, only the ERDF funds are considered, and the target values used are too broad and unsubstantiated. There is a lack of consideration of the various cost components for projects that are structurally similar, and none of the Managing Authorities set clear 'internal' norms or standards to assess value for money.
- **Limited ex-post analysis:** Effectiveness of projects is considered by reporting the achieved targets. However, most of the time this is done according to the statement of the beneficiary, without structural ex-post appraisal by the Management Authority.

Source: Algemene Rekenkamer (2014) *EU-Trendrapport 2014*, http://www.rekenkamer.nl/Publicaties/Onderzoeksrapporten/Introducties/2014/02/EU_trendrapport_2014 (last accessed 20 September 2014).

²² <http://www.europaomdehoek.nl/> (last accessed 20 September 2014).

3. IMPLEMENTATION AND METHODS

3.1 Who commissions and undertakes evaluation?

Largely reflecting the allocation of responsibility over policy design and implementation, evaluations are commissioned predominantly at national level in Netherlands, Norway, Sweden and the United Kingdom (England and Scotland), and at sub-national level in Austria. In Finland, France, Germany, Italy and Poland, evaluations are commissioned by both national/federal and *Land*/regional levels. In Germany, the federal level is the dominant actor in relation to the domestic GRW, whilst Cohesion policy evaluations are largely the responsibility of individual **Länder**, whereas in Italy and Poland regional-level commissioning of evaluations has increased in parallel with the progressive strengthening of evaluation capabilities within the regions (a longer-term trend in Italy, more recent in Poland). In France, policy effectiveness is evaluated predominantly at national level, but the regions are responsible for the evaluation of Cohesion policy programmes.

The majority of evaluation work is entrusted to actors who are external to the administration, primarily consultants and, in some cases, individual academics or academic institutions. This is largely done on grounds of objectivity/independence and linked to the specialist skills that evaluation teams can bring into the process. However, a **fairly recent trend is the increasing internalisation of evaluation**, for instance within dedicated evaluation units (in Italy and Scotland) or government departments (in France and also Austria, where internal evaluations have been undertaken for a number of years). The entrusting of evaluation to actors within the administrations is due in part to cost-effectiveness considerations and in part to a desire to improve the effectiveness of the ‘policy→evaluation→policy’ loop. Even where evaluation is predominantly outsourced, some concerns have been raised about the added value entailed by external evaluations and the extent to which they provide new insights (e.g. in Finland, Germany and the UK).

There is sometimes a sense within the administrations that policy-makers have a sufficient understanding of the working of a policy or scheme, which limits the added value of involving external evaluators (e.g. in Finland and Germany). Interviews undertaken in the German *Land* of Mecklenburg-West Pomerania, for example, noted that the staff involved in implementation already have very good knowledge about ‘what works’, even though they do not regard the various in-house sources of knowledge or activities as evaluation. Staff are in regular contact with the users of funding (e.g. firms looking to invest) and thus have a good overview of the types of instrument/funding that such users need and of the kinds of funding where there is strong take-up. They also undertake analyses of regional economic problems and in-house estimates of effects – for example, of the leverage effect of GRW funding, in terms of jobs created relative to the amount of funding – and have good data sources both from the federal and *Land* statistical offices and the monitoring systems (e.g. the GRW’s). Nevertheless, formal evaluations of the GRW have been undertaken at federal level for Germany as a whole, and within Mecklenburg-West Pomerania in the context of Cohesion policy evaluations of business aid.

Table 2: Actors of the evaluation process (commissioning bodies and evaluators)

Country	Type of evaluator	Territorial level commissioning
Austria	Consultants and academics Occasionally internalised	(Predominantly) <i>Land</i>
Finland	Mainly external (objectivity, but some concerns that added value/new insights not always there)	National and regional
France	Varied: politicians, policy-makers, civil servants, external experts, auditors – not so much academics or consultants – with Ministries in charge of putting together evaluation teams	National and regional
Germany	Consultants and academics	Domestic GRW – mainly federal EU CP – mainly <i>Land</i>
Italy	Consultants and academics Regional and national evaluation units	National and regional (in recent years especially regional)
Netherlands	Consultants, public research institutes, National Audit Office	National
Norway	Consultants and academics	National
Poland	Consultants (mostly) and academics (increasing)	National and, increasingly, regional
Sweden	National agency (<i>Tillväxtanalys</i>) is responsible for organising evaluations, which can be carried out internally or externally.	National
Switzerland	Consultants	Federal and cantonal
UK	Consultants (England – might change due to budget cuts) Internal SG Unit (Scotland, increasingly) and consultants	National

Source: EoRPA research.

3.2 Scope

The scope of evaluation varies across countries. As illustrated in Table 3 below, evaluations can relate to:

- individual projects**, as in Switzerland in relation to the New Regional Policy introduced in 2008;
- schemes or policy instruments**, as the evaluation of the ‘programming contracts’ instrument in Italy or of the Regional Selective Assistance Scheme in Northern Ireland;
- whole programmes**, as in the example from Finland in the case study box to follow;
- a theme**, for example, the evaluation of the implementation of financial instruments in Poland.

Examples of each approach are provided in the paragraphs to follow.

Table 3: Scope of evaluation

Country	Scope
Austria	Not individual schemes (e.g. they need often to be assessed jointly, not in isolation).
Finland	Instruments or policy programmes (e.g. evaluations covering effectiveness of centre of expertise programme, regional strategic programmes, individual projects).
France	Themes, not individual schemes but entire themes (e.g. business aids). Themes are selected by government, according to governmental priorities and perceived need. Inter-ministerial Committee for the Modernisation of Public Policy decides the evaluation programme.
Germany	Instrument/specific aspects of GRW funding (e.g. business-oriented infrastructure; business aids).
Italy	Varied: from themes to programmes to specific schemes/instruments (e.g. programming contracts).
Netherlands	Varied: from single instrument, to overall programme, to theme.
Norway	Varied: from single instrument to as theme, to the overall policy or governance. But all programmes are evaluated at some point in their lives at programme level.
Poland	Varied: Instruments, support to specific types of recipients, programmes.
Sweden	Whole programme, large projects (€1mill+), implementing organisation.
Switzerland	Projects (to gauge the NRP's impact as a whole).
UK	Varied: from single instrument, to overall programme, to theme.

Source: EoRPA research.

3.2.1 Project-level evaluation – Measuring the impact of the New Regional Policy projects in Switzerland

In 2012, SECO launched a pilot study²³ on measuring impacts at the project level. For this, 21 NRP projects from five different thematic priorities (innovation, tourism, education/health, energy, natural resources) were selected, with NRP funding ranging from €8,200 to €1.4 million per project. The study looked at the effects that NRP funding had at project level (outputs and outcomes) and at regional level (impact), as well as for the participating firms. Indicators used to measure the impact on the regional economy were employment and added value. The report distinguished between three forms of effects that policy interventions can have:

- Direct effects: additional turnover for firms participating in NRP projects, additional employment created by firms participating in projects. Comparatively easy to measure.
- Indirect effects: additional turnover for (regional) firms providing goods/services to firms participating in NRP projects, and related potentially additional employment. Relatively challenging to measure; reliance on estimates based on multipliers.
- Induced effects: additional turnover for (regional) firms due to increased consumption of employees benefiting directly or indirectly from NRP support. Very challenging to measure; only rough estimates are possible.

²³ Regiosuisse (2013) Wirkungsmessung NRP-Projekte 2012. Synthese der Wirkungsmessung, http://regiosuisse.sswm.info/sites/default/files/wirkungsmessung_nrp-projekte_2012_synthese_schlussbericht_0.pdf

Estimates for the impact indicators 'employment' and 'added value' were calculated on the basis of

- Jobs created directly, (additional) turnover, services/good from suppliers, import quota;
- An existing regional employment multiplier²⁴ that is based on import quota and regional interdependencies;
- Sector-specific information, e.g. average turnover and added value per full-time equivalent.

The estimated impacts on regional added value ranged from an additional €16,000 to €26 million per year. In addition to quantitative impacts, the report also highlighted a series of qualitative ones, e.g. increased cooperation. However, the pilot study also showed that measuring impacts is complex and requires resources. It is difficult to impose such approaches on project managers and policy-makers who tend not to have the necessary evaluation expertise. The result of the study fed into a conference in 2013 looking at the impact of NRP and INTERREG,²⁵ which resulted in wide agreement that effort in terms of impact evaluation must be increased.

3.2.2 Instrument-level evaluation: Evaluating the Contratti di Programma in Italy

In 2006, the Italian national Evaluation Unit undertook an evaluation of the effectiveness of one of the main business aid schemes, the *Contratti di Programma* (programme contracts).²⁶ The evaluation covered five themes: additionality; territorial impact (economic effect and longer-term sustainability); R&D; role of the negotiation stage in project selection; and role of the ex-ante appraisal. With a budget of €260,000, the work entailed a mixed micro-and-macro perspective with:

- 10 qualitative case studies, all in the Southern regions (especially for the assessment of additionality and effects on the territory);
- analysis of administrative and statistical data relating to the case studies analysed;
- fieldwork through semi-structured interviews with entrepreneurs, institutional actors, elite witnesses;
- econometric modelling for the appraisal of the national effects; and
- for a selected Programming Contract (that relating to the FIAT Melfi Plant), a meta-evaluation comprising systematic documental and bibliographic analysis.

Following on from this study, a similar evaluation was undertaken in 2013 by the region of Apulia.²⁷ The evaluation focused on the effect of the programme contracts on the firms supported and on the local productive system, and on the effectiveness and efficiency of implementation (actors and processes of negotiation, launch and implementation). Evaluation questions were:

- To what extent and in which way has the public incentive influenced the choices of the entrepreneur in relation to the size, place and timing of the investment?

²⁴ Giuliani G and Berger S (2010) *Leitfaden für die regionalwirtschaftliche Beurteilung von Entwicklungsstrategien und -projekten*, Zürich, <http://www.regiosuisse.ch/regionalpolitik/projekte/projekte/download/publikationen/handbuecher-leitfaden-methodenbuecher/leitfaden-regionalwirtschaft-bericht.pdf> (last accessed 20 September 2014).

²⁵ <http://www.regiosuisse.ch/weiterbildung/formation-regiosuisse/formation-regiosuisse-konferenz-2013> (last accessed 20 September 2014).

²⁶ http://www.regione.puglia.it/web/packages/progetti/nvip/Rapporto_Finale_CdP.pdf (last downloaded 20 September 2014).

²⁷ <http://www.dps.mef.gov.it/documentazione/uval/Relazione%20CDP.pdf> (last downloaded 20 September 2014).

- What has been the impact on the territory of the investment realised from a social, economic and environmental perspective? (firms' performance, relationship between firms and firm-other actors, territorial effects)
- What factors have influenced the decisions taken and the results obtained?

The evaluation relied on a similar methodology as the preceding national study, comprising:

- 12 qualitative case studies, representative of the projects financed (i.e. different territorial areas and different sectors);
- desk research (primary and secondary);
- interviews with those responsible for the implementation of the contracts within the supported firms or consortia (though there were difficulties with tracking down those who had been in charge of the negotiations or the external actors involved, e.g. research centres and universities involved in the activities of R&D due to the time elapsed since the launch of the projects);
- a specific methodology for the training component of the scheme (CATI method);
- a social accounting matrix (SAM) for the overall assessment of the effects and economic impact of the regional economic system.

The Apulian evaluation reported significant impacts on firms, of different and varied nature, and on the regional economic system with the exception, in relation to the latter, of R&D.

3.2.3 Instrument-level evaluation: Evaluating the Selective Financial Assistance in Northern Ireland

The aim of the evaluation was to provide a robust evaluation of the operation, outcomes and impact of Selective Financial Assistance (SFA), the main business aid scheme in Northern Ireland, in order to provide a solid evidence base on which to assess the extent to which SFA met its economic objectives and to inform on-going discussions with the BIS and the European Commission over Northern Ireland's Assisted Area status. The evaluation was commissioned in September 2012 by the Department for Enterprise, Trade, and Investment (DETI) in Northern Ireland to SQW Ltd as lead contractor, working alongside Aston Business School and Ipsos MORI as sub-contractors. The Final Report was published in July 2013.²⁸ It covered the projects approved over the financial years 2004/05 to 2010/11. Key elements of this assessment included examination of the following elements:

- Rationale for intervention: testing the extent to which SFA met a genuine need over the evaluation period in terms of market or other failures, and its alignment with the economic and strategic contexts.
- Additionality: the effects of SFA over and above what would have occurred in any case, taking into account deadweight, displacement, and substitution effects.
- Net economic impact: the employment and Gross Value Added (GVA) effects delivered and the overall contribution of SFA to the Northern Ireland economy over the evaluation period.
- Value for Money: in terms of its economy, efficiency and effectiveness and overall return on publicly-funded investment.

²⁸ http://www.detini.gov.uk/final_report_-_sfa_evaluation_2004-11.pdf (last downloaded 20 September 2014).

The study adopted a logic-chain approach, designed to enable a robust assessment of (and the linkages between) the rationale for intervention, the emerging objectives, inputs, activities and outputs delivered, and the resulting outcomes and impacts. At the level of benefiting firms, the evaluation applied both a 'self-reported' approach to assessing the additionality and impact of SFA in line with Green Book thinking, and econometric modelling techniques that sought to ascertain the net effects of SFA assistance, after controlling for the effects of 'selection bias'. The methodological approaches comprised a mix of primary research and desk-based analyses:

- Detailed analysis of SFA monitoring data, covering approved and recorded expenditures, and contracted gross outputs at a project level.
- A review of key SFA and strategic documents including relevant SFA guidance, relevant institutional corporate plans, and wider economic policy documents and statements.
- Telephone surveys of beneficiaries and non-beneficiaries of SFA.
- Case studies with 11 firms who had received SFA support, involving a review of project documents and a detailed discussion with the lead contact at the firm.
- An online survey of Invest NI Client Executives, providing evidence on the SFA process from the perspective of Client Executives who had worked with firms in progressing SFA projects.
- Telephone consultations with senior representatives from across Northern Ireland's business, finance, and policy communities.

The evaluation found that the on-going case for intervention remained sound in 2011, and significant firm-level benefits and economy-wide impacts were delivered. A range of GVA impacts were estimated, up to £1.5 billion based on net jobs created and assuming four years persistence. Employment effects were evidenced, with an estimated c.10,000 new jobs created and 7,500 net jobs safeguarded over the evaluation period. Metrics from the econometric analysis were different, but positive, with a small impact on beneficiary firm growth over a one-year period, although not over five years. Value for money was found to be acceptable and the cost per net job created/safeguarded was generally within benchmarks. SFA's additionality was characterised as medium and acceptable, in line with wider evidence on business interventions. Less positively, SFA supported a wide range of firms, including many small projects delivering limited value and low additionality, and there was underspend and underperformance against agreed targets.

The report made three policy recommendations:

- In general, SFA should be continued, as its rationale remains valid given the on-going economic challenges faced by Northern Ireland, it performed adequately over the evaluation period, and it remains an integral part of the business support landscape in Northern Ireland, given its flexibility and scale. However, the existing intervention needs refreshing, tightening and re-purposing.
- SFA's delivery should remain with Invest NI, given Invest NI's business focus, track record, delivery capacity and existing relationships with client firms and the business representative community.
- DETI and Invest NI should undertake a formal and thorough options assessment to consider the most appropriate function and form of SFA as a policy intervention for the future.
- In addition, there were a further seven recommendation themes and actions, which included a formal theory of change to deliver greater clarity and consistency on SFA purpose and focus.

The work was complex and technically challenging. Amongst the challenges, the most significant related to: monitoring data; separating out SFA projects (generally provided as part of a bundle); memory decay; lack of formal documents on the scheme; and establishing the effect of the recession on the scheme.

3.2.4 Programme-level evaluation: Evaluation of the Regional Strategic Programmes of West Finland and South-West Finland

The evaluation focused on the implementation and impact of the regional strategic programmes in the regions of West Finland including South Ostrobothnia, Central Finland, Pirkanmaa, Ostrobothnia, Satakunta and South-West Finland (Varsinais-Suomi). The obligation for the evaluation was set out in the former regional development law (1651/2009), according to which an external evaluation needed to be carried out for the strategic regional programme at least once during the programme period. The evaluation was undertaken by Ramboll Management Consulting and commissioned by the concerned regions, and it was carried out between October 2012 and January 2013.²⁹

The evaluation aimed to provide the following:

- An assessment of the performance of the regional strategic programme as an instrument of regional development in the individual regions and in the area as a whole. In addition, it provided an impact evaluation, especially concerning the common development perspective of the regional actors, the cooperation, the commitment of the actors, and the achievement of the regional development objectives.
- An evaluation of how the recommendations of a previous evaluation (2009) had been implemented.
- An evaluation on the performance of the regional strategic programme, its results and impact, as well as a related analysis that provided tools for the new programme period (in terms of how to increase the impact of the different themes and of the different operational models embedded in the programme).
- The evaluation report aimed to provide information for the purposes of the development of the new regional development law and for the preparation of the regional development guidelines for the next phase.

The evaluation methods consisted of: documentary research, monitoring results of each region, interviews, and an online survey. In addition, information was gained through workshops organised in each region focused on specific themes.

The evaluation results cover a number of different themes, including the common development perspective of the regional actors and cooperation, the commitment of the actors, and the achievement of the regional development objectives. With regard to the achievement of the programme objectives, the evaluation noted that the period had been particularly challenging. The regional programme work at its best achieved a slowdown in the negative implications of the sudden structural changes that the regions were faced with, particularly in the field of ICT. At the same time, the various initiatives started in the regions had not been launched sufficiently rapidly to respond to

²⁹ Ramboll Management Consulting (2014) [Maakuntaohjelma aluekehityksen välineenä: näkökulmina vaikuttavuus ja alueen kokonaiskehittäminen – arvio maakuntaohjelmakaudesta 2011-2014 Länsi-Suomen maakunnissa ja Varsinais-Suomessa](#) (last downloaded 20 September 2014).

the problems. The key recommendation concerning the effectiveness of the programme was to introduce a model that entails forecasting, monitoring and evaluation, as well as a more flexible and future-oriented approach to regional planning.

The evaluation raised concern over the difficulty of measuring impact. It is very difficult to say what result or achievement is due to the regional strategic programme, and what is not. In addition, the regional strategic programmes often use terms such as 'increase', 'improve' or 'influence', without concretely specifying how this is monitored and what indicators are used for this purpose.

3.2.5 Theme-level evaluation: Evaluating financial engineering instruments in Poland

The evaluation was commissioned by the Polish Ministry of Regional Development (now Ministry of Infrastructure and Development) from a consortium, Policy & Action Group Uniconsult and Taylor Economics. The evaluation had a budget equivalent to c. €87,565 and was realised in the period August 2012 to March 2013. Poland has little experience of using financial engineering instruments in regional policy, and the evaluation was commissioned to draw together initial experience of their effectiveness and to inform on their use in the 2014-2020 period.

The primary objective of the study was to evaluate the implementation and effectiveness of financial engineering instruments within the framework of the Operational Programme Innovative Economy, 2007-2013 Regional Operational Programmes and the Operational Programme Development of Eastern Poland, and to build a map of financial engineering instruments and other repayable financial instruments supported from public funds in Poland. The study also aimed to develop insights to inform preparations for the use of these instruments in the 2014-2020 programme period.

The methodology of the study comprised a wide range of qualitative and quantitative methods, in particular broad desk research (domestic and foreign), c. 90 individual interviews with representatives of public administration bodies, entrepreneurs, financial intermediaries, beneficiaries implementing repayable financial instruments and experts. Six focus-group interviews (FGI) were held with entrepreneurs, who had been granted repayable and non-repayable assistance, and 1,098 computer-assisted personal interviews (CAPIs) with entrepreneurs benefiting from various forms of repayable instruments supported from the public funds and 500 CAPIs with entrepreneurs obtaining no support under financial engineering instruments (as counterfactual control group) were also undertaken.

The main conclusions were as follows:

- In general, repayable financing products offered with public assistance are effective in matching the needs of entrepreneurs and reducing the scale of the gap in capital financing in Poland. Nevertheless, the products involved often do not address the highest-risk projects.
- The impact of public funding on the availability of repayable financing is positive in expanding the opportunities for capital financing, (particularly in seed capital and access to debt financing for start-ups) though full and reliable data are not always available.
- Effectiveness is limited by competition between instruments financed from public sources and purely commercial instruments.
- A key problem relates to the durability of financial instruments and the monitoring of their results. It is critical to ensure that on one side the financial intermediary knows in advance the

status of funds upon project completion and on the other side that mechanisms are in place, enabling the most active and effective intermediaries to continue using the transferred funds.

- A serious problem in implementing certain financial engineering instruments (in particular within the scope of ROP) was double-financing.
- The cost effectiveness of respective instruments was viewed in most cases as positive, although the situation is somewhat differentiated among various FEIs.
- Loans allow enterprises to increase the level of debt and to increase investment expenditure. However, a loan's impact on the level of net income generated by a company was not confirmed. Also, the net effect in the form of an increase in the number of jobs could not be confirmed.
- Guarantees contributed to job creation but did not help to increase investment expenditure in businesses that were granted such guarantees – they only accelerated certain investments.

The report provides a number of recommendations, including the absolute necessity to introduce clear mechanisms to coordinate the structuring and implementation of repayable instruments financed from both national and European resources, and that a considerable number of the instruments examined should be continued, albeit with some modifications. With particular relevance to the evaluation of the effectiveness of the schemes, the report recommended the establishment of a system for the periodic gathering of data about the results of lending / guarantees supported: joining such a system should become a condition for applying for further public assistance.

3.3 Focus

The focus of evaluation also varies (Table 4). Effects, impacts, effectiveness, value for money and implementation are all different foci and tend to dominate differently in different national contexts.

Table 4: Predominant evaluation foci

Country	Impact	Effectiveness	Value for money/efficiency	Implementation	Utility
Austria		X		X	
Finland		X		X	
France	X	X			
Germany	X	X	X		
Italy		X		X	X
Netherlands	X	X	X		
Norway		X	X	X	
Poland	X			X	
Sweden				X	X
Switzerland	X				
UK (England)	X		X		
UK (Scotland)				X	

Source: EoRPA research.

Appraising the effectiveness, impact and implementation of programmes are the most common aims amongst the EoRPA countries. Value-for-money considerations are particularly relevant in the United

Kingdom (England), Germany, the Netherlands and Norway, whilst utility is also being appraised in Italy and Sweden.

As illustrated in Table 4, the evaluation of effectiveness is prominent in a number of countries. A straightforward understanding of effectiveness is that it is about whether intended objectives have been achieved. The Evalsed evaluation guide,³⁰ for instance, links the concept of effectiveness to the following questions:

- To what extent have the objectives been achieved?
- Have the interventions and instruments used produced the expected effects?
- Could more effects be obtained by using different instruments?

However, this concept can be interpreted in different ways, as effectiveness can be appraised considering a plurality of different levels:

- a) In relation to whether the policy has contributed to the realisation of its overarching goals (e.g. for Cohesion policy, the goals assigned by the Treaty); and/or
- b) In relation to whether programmes have achieved
 - a. the macro-objectives declared in their strategy sections
 - b. the specific targets assigned to priorities and measures.
- c) In relation to projects, i.e. whether the projects funded have realised their declared goals and targets, which may be only a subset of the targets or goals of a measure or scheme (e.g. increased productivity, product innovation or personnel employed in the case of a business aid scheme).

Terms of reference or evaluation studies may not necessarily specify how effectiveness is understood in the context of the specific study. There is therefore variation across evaluations both within and across countries in terms of the interpretation of effectiveness. For instance:

- In Germany, effectiveness is generally defined in relation to the specific **goals of a particular regional policy instrument** (e.g. creation of jobs, attraction of new firms, increase in business investment, creation of new cooperation networks etc.).
- In France, where assessing the effectiveness of public policies is currently at the heart of the government's plans, a recent guidance document by the General Administration Inspectorate states that appraising effectiveness entails answering the following questions: What is the specific impact of a public intervention? Are the observed results and effects in line with initially defined objectives? To what extent do different instruments contribute to these objectives?³¹ The focus is thus on **whether the effects of instruments contribute to the objectives of the policy**.
- In a guidance note produced by the Polish National Evaluation Unit in 2012, the definition of effectiveness is: 'The extent to which goals are realised (i.e. whether the planned goals

³⁰ Evalsed (2013) *EVALSED: The resource for the evaluation of Socio-Economic Development*, September 2013, http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/guide_evalsed.pdf (last downloaded 20 September 2014).

³¹ Inspection Générale de l'Administration, Inspection Générale des Finances et Inspection Générale des Affaires Sociales (2012) *Cadrage méthodologique de l'évaluation des politiques publiques partenariales*, Guide, http://www.igf.finances.gouv.fr/webdav/site/igf/shared/Nos_Rapports/documents/2012/2012-M-097-01%20Guide%20m%C3%A9thodologique.pdf, p.2.

were achieved), the effectiveness of the methods and institutions used and the impact of external factors on ultimate effects'. The focus is thus on **effectiveness and the reasons underlying it**.³²

- In a methodological document by the Italian Evaluation Unit, the evaluation of effectiveness is defined in quite a broad sense as 'if and to what extent the intervention has been able to reach the expected results or, in any case, to generate other useful results'.³³ The focus is thus on both **effectiveness in the strict sense and utility more widely**.

The way effectiveness is defined, even if only implicitly, has implications for the operationalisation of evaluation, as it determines whether unintended effects are considered, the timeframe of the enquiry (e.g. whether short-term or long-term) and whether the evaluation seeks to disentangle causality. It also has implications for the type of lessons that can or cannot be drawn from evaluation.

For instance, at EU level, the evaluation of the effectiveness of Cohesion policy in past programme periods has entailed essentially a matching exercise between 'the system of output, result and impact indicators[and] the trend of ex-ante-defined target variables ... a sort of performance evaluation' that might indicate 'the direction in which some desired variables travel, without however ... revealing much on the causes of the change or the contribution (positive or negative) of the different components of the policy to such change'. This type of analysis cannot lead to the formulation of judgements on the effects of the interventions.³⁴

3.4 Methods

Overall, the evaluation of regional development policy is undertaken according to a variety of methodological approaches, ranging from heavily quantitative, to mainly qualitative, to a mix of the two. These vary according to a number of dimensions, including (Hoerner and Stephenson, 2012):

- **Epistemological assumptions:** Is the evaluation considered to deliver generalisable results or context-specific results?
- **Top-down or bottom-up:** Are studies large/meta or in-depth? Are evaluation questions defined top-down/at the outset, or kept fluid/emerging during the evaluation process?
- **Data elaboration techniques:** Are quantitative or qualitative methods used? How is data elaborated upon?
- **Data sources:** What are the data sources? e.g. interviews, survey, statistical databases, quantitative monitoring data, etc.
- **Stakeholder involvement:** To what extent are stakeholders involved? Are a lot of different stakeholder organisations involved, or only a few larger ones/none at all?
- **Role of the evaluator:** Is the evaluator considered as a neutral, *super partes* actor or as part of the process?

³² Ministerstwo Rozwoju Regionalnego (2012) *Ewaluacja Poradnik dla pracowników administracji publicznej*, https://www.mir.gov.pl/rozwoj_regionalny/Polityka_rozwoju/Projekt_Zarzadzanie_Strategiczne_Rozwojem/Przydatne_dokumenty/Documents/EWALUACJA_z%20okladkami.pdf (last downloaded 20 September 2014).

³³ Marchesi G, Tagle L and Befani B (2011) Approcci alla valutazione degli effetti delle politiche di sviluppo regionale, *Materiali Uval*, 22/2011, p. 16, http://www.dps.mef.gov.it/documentazione/uval/materiali_uval/MUVAL_22_Valutazione_impatto.pdf (last downloaded 20 September 2014).

³⁴ Ibid. Own translation from Italian.

These dimensions are significant, in that they indicate often implicit choices as to the theoretical anchoring of evaluation – along the positivist-constructivist continuum (and the intermediate realist approach) – and, related, about the actual reach of evaluation, i.e. the extent to which it is expected to disentangle cause-effect relations or rather to provide at best speculative assessments of the degree of influence extorted by the policy.³⁵

The main methods employed in the evaluation of regional policy include: macro-economic modelling, econometric statistical analysis, micro-economic / counterfactual analysis, qualitative case studies, and theory-based evaluation.

Macro-economic modelling provides an overview and/or forecast of a policy's effectiveness at the aggregate level. It is based on a systematic analysis of variables that are considered key target values of the policy (e.g. GDP, employment, etc). Sophistication levels in models differ considerably and are often highly dependent on data availability. Although the detail and highly specialist approach used in macro-economic models makes them difficult to interpret, as the results are quantifiable and can often be translated into easy messages, they are popular with policy makers and politicians. An example of the use of a macro-economic model to assess the impact of a regional development programme is represented by the evaluation of the socio-economic impact of the Śląskie ROP 2007-2013 in the Śląskie region, which was undertaken through an adaptation of the HERMIN model (see Box 4 below).

³⁵ Hoerner and Stephenson (2012) *op. cit.*

Box 4: Macro-economic modelling to evaluate the impact of the Śląskie ROP 2007-2013

This is an evaluation commissioned by the Marshal's Office, Regional Self-government Board of the Śląskie region, to establish the impact of the Regional Operational Programme of Śląskie 2007-2013 on the socio-economic development of Śląskie, using the 5-sectoral macro-economic HERMIN model. The evaluation was undertaken by the Wrocław Regional Development Agency Joint Stock Company in a short time-span (30 September to 17 December 2013). It had a budget equivalent to c. €30,000 (0.002 percent of the ROP). The rationale was to inform Cohesion policy interventions in the 2014-2020 period and inform debate on development directions in the region post-2020.

Evaluation questions:

What has been the impact of the ROP according to macro-economic indicators, including:

- The two main indicators related to the key objectives of the programme (change in GDP and net job creation).
- A set of other indicators covering other socio-economic issues: supply side of the economy; convergence processes at national and EU levels; social development reflected by the situation on the labour market; the economic structure; standard of living of the population; and the public finance sector and foreign trade.

Evaluation methods:

The evaluation used the most current version of the national HERMIN model for Poland, which is part of the Cohesion System of HERMIN Models (CSHM). It also employed the 'moving' or 'mean' average method to analyse data points by creating a series of averages of different subsets of the full data set. This was applied to data from Śląskie's four sub-regions. The evaluation involved four stages:

- Updating the statistical base
- Updating and adapting the HERMIN model
- Creating scenarios for the region and for the sub-regions 'with ROP' and 'no ROP'
- Analysis and interpretation of simulation results.

Main results:

- For the 2008-2016 programme funding period, the average ROP contribution to the growth of GDP is between 0.8 percent (in the central sub-region) and 1.7 percent (in the northern sub-region). For the entire region, the average value of the programme's impact on GDP is expected to be around 1 percent.
- The number of additional jobs created thanks to the support of the programme in the period 2008-2016 is estimated to range between 2,000 (in the western sub-region) and 7,100 (in the central sub-region). The total additional jobs in the region is expected to be 13,800.
- The effect of GDP growth is expected to reduce the GDP gap with the EU average: with the financial support from the programme, regional GDP per head is expected to reach 72.3 percent, 1.1 percent higher compared to the 'no ROP' scenario.
- The positive impact of the funds is also visible in other indicators, especially labour productivity and gross fixed capital formation.

Recommendations:

Applying the same intervention logic and economic structure of intervention used for the 2007-2013 ROP to the new 2014-2020 ROP – with infrastructural projects dominant – will generate a relatively weaker effect than in 2007-2013. This is due to the fact that the resources of material capital and human capital are now higher than at the beginning of the implementation of the 2007-2013 NSRF. Their continued growth – as a result of Cohesion policy – will generate relatively modest results in terms of stimulating socio-economic development. Thus, funds under the 2014-2020 ROP should gravitate towards a greater share of expenditure on R&D. As part of the smart specialisation agenda, this should support cooperation between science and the business sector and facilitate the exploitation of the potentials of the region.

Source: Wrocław Regional Development Agency (2013) *Wpływ realizacji Regionalnego Programu Operacyjnego Województwa Śląskiego na lata 2007-2013 (RPO WSL) na rozwój społeczno-gospodarczy województwa śląskiego z wykorzystaniem makroekonomicznego 5-sektorowego modelu HERMIN*, Katowice, 2013; https://www.ewaluacja.gov.pl/Wyniki/Documents/1_200.pdf.

A simpler alternative to macro-economic modelling comprises approaches based on **econometric regression analysis**, which allow for the evaluation of specific hypotheses in single equations. A further, even more streamlined, approach was utilised in the already-discussed evaluation of projects funded by the New Regional Policy in Switzerland.

Micro-econometric studies are based on the use of control groups that compare outcomes for those that have received assistance through the interventions with those that have not (counterfactual). An example is provided in the case study box below (Box 5), relating to an evaluation appraising the impact of aid for business investment granted via the German GRW. Other examples of counterfactual evaluation can be found in Italy (for example, an evaluation of the impact of innovation measures contained in the 2007-13 Marche ERDF ROP) and the Czech Republic (an evaluation of the employment effects of the Human Resources NOP 2007-13).³⁶

Box 5: Micro-econometric impact analysis for the appraisal of the impact of Germany's Joint Task for the 'Improvement of the Regional Economic Structure'

The evaluation was commissioned by the Federal Ministry for the Economy and Energy (previously the Federal Ministry for the Economy and Technology) from academics at the University of Dortmund. It was commission in the context of domestic discussions over the future of the Regional Joint Task, its effectiveness and orientation, and whether any changes were needed.

The core aim of the evaluation was to compare the development of employment and wages in aided and non-aided firms. The call for the evaluation was published in 2008; the work was undertaken in 2009-10; and the final version of the evaluation was published in September 2010.

The study entailed a micro-econometric impact analysis, based on a 'matching' approach, which compared 4,622 firms aided by the GRW in 2001-06 with non-aided firms with similar characteristics (e.g. business location, sector, size, skills, type of activity, age of firm and previous record of growth).

The evaluation found that employment in aided firms grew by an annual average of 1.9 percent between the year of aid and 2008, while employment in non-aided firms with similar characteristics fell by 6.8 percent annually.

It made recommendations for the creation of an on-going approach to evaluating the effectiveness of GRW aid for business investment.

A follow-up study was published in January 2013,³⁷ using the same matching methodology but focusing on the question on whether the effectiveness of GRW investment aid varied across firms of different sizes and, specifically, on the effectiveness of aid to large firms. This showed that effectiveness did not differ significantly between firms of different sizes, in terms of employment created in aided versus non-aided firms.

Source: <http://bmwi.de/DE/Mediathek/Publikationen/publikationen-archiv,did=376256.html>.

More qualitative approaches are often based on **case study** designs. Case studies are 'an intensive study of a single unit or a small number of units (the cases), for the purpose of understanding a larger class of similar units (a population of cases)',³⁸ and they can vary both in relation to the type of knowledge pursued – from illustrative, to exploratory, to explanatory³⁹ – and in relation to the types of methods utilised, from solely qualitative to mixed. Case studies generally rely on a mix of primary and secondary sources, comprising literature and document reviews, statistical and administrative data,

³⁶ T33 (2014) POR CRO FESR Regione Marche 2007-2013 – Servizio di valutazione degli interventi attuati a favore della ricerca e dell'innovazione, 21.12.2011,

http://ec.europa.eu/regional_policy/sources/docgener/evaluation/library/italy/1112_italy_marche_eval_it.pdf;

Potluka O and Brůha J (2014) *EU Cohesion Policy attribution to employment: a case of the Czech Republic*, Paper presented at the 2014 ECPR General Conference, Glasgow, 7 September 2014,

<http://www.ecpr.eu/Filestore/PaperProposal/e13a3696-4251-4f98-8800-d72e79c343ba.pdf>.

³⁷ Bade F-J (2013) *Bedeutung und Wirksamkeit der Förderung größerer Unternehmen durch den gewerblichen Investitionszuschuss im Rahmen der Gemeinschaftsaufgabe 'Verbesserung der regionalen Wirtschaftsstruktur' (GRW)*, Dortmund, <http://www.bmwi.de/DE/Mediathek/publikationen,did=574466.html> (accessed 11 July 2014).

³⁸ Gerring J (2007) *Case Study Research. Principles and Practices*, Cambridge University Press, p. 37.

³⁹ Illustrative case studies describe the main characteristics of a real-world example to clarify an idea or reinforce an argument; exploratory case studies aim to understand what happened within a case by looking beyond descriptive features and studying the surrounding context; explanatory case studies aim to explain why certain behaviours occurred by determining causes and effects. Commonwealth Association for Public Administration and Management (2010) *Overview of Case Study Models and Methodology*, <http://www.capam.org/documents/reportoncasestudymethodologies.pdf>.

and interviews, surveys, direct observation, focus groups and other techniques utilised to collect new data. Triangulation of these different sources is at the heart of the approach. The case study method aims to construct a 'narrative' and has been a preferred method by the Commission (DG Regio) in the context of the ex-post evaluations of Cohesion policy programmes in the 2000-06 and 2007-13 periods.⁴⁰

More recently, **theory-based impact evaluations** have received increased attention. Such an approach is concerned with the logic and utility of interventions and focuses heavily on the mechanisms that explain the changes that a policy or programme aims to achieve. 'The perceived value of theory-driven evaluation is, in part, generating knowledge such as not only knowing whether a program is effective or efficacious (i.e., causal description; *that* a causal relationship exists between A and B) but also explaining a program's underlying causal mechanism (i.e., causal explanation; *how* A causes B)'.⁴¹

Table 5 attempts to classify these different methods according to the various dimensions outlined above.

⁴⁰ Gaffey V (2009) 'Case studies in the framework of ex post evaluation 2000-2006: expectations and experiences of the DG for Regional Policy', http://ec.europa.eu/regional_policy/archive/conferences/evaluation2009/abstracts/gaffey.doc

⁴¹ Coryn C L S, Noakes L A, Westine C D and Schroeter D C (2011) A Systematic Review of Theory-Driven Evaluation Practice from 1990 to 2009, *American Journal of Evaluation*, 32(2), 199-226; Leeuw F L (no date) *Theory-based Evaluation*, http://ec.europa.eu/regional_policy/information/evaluations/pdf/impact/theory_impact_guidance.pdf; Stame N (2004) Theory-Based Evaluation and Types of Complexity, *Evaluation*, 10(1), 58-76.

Table 5: Methodological approaches

Methodol. approach	Epistemol. assumption	Bottom-up/ top-down	Data elaboration techniques	Data sources	Stakeholder involvement	Role of evaluator	Key advantages	Key disadvantages
Macro-economic	Generalisable	Top-down	Statistical – multiple equations	Statistical databases with macro-economic data	No	Neutral	Simulates policy intervention Incorporates a large number externalities and spillovers High level of consistency and comparability Results communicable and politically desirable	Contested theoretical assumptions Only suitable when policy funding accounts for large percentage of GDP Highly complex and difficult to interpret Can incorporate limited number of control variables Limited use for regional interventions
Econometric	Generalisable	Top-down	Statistical – limited number of equations	Surveys Statistical databases with macro-economic data	No	Neutral	Less complex than macro-economic models Allows hypothesis-testing Includes a number of explanatory variables Allows for falsification of results	Theoretical assumptions not universally held Causality often difficult to prove 'Black box approach' – ignores complex processes and feedback mechanisms Does not account for independence of cases, particularly between regions Does not account for spatial heterogeneity
Micro-economic	Generalisable	Top-down	Statistical / Counterfactual	Monitoring data Surveys	Possible (limited)	Neutral	In-depth detailed evaluation taking context into account Counterfactual assessment Provides cross-instrument comparison	Establishing control groups is challenging Does not always take account of exogenous factors Lack of generalisability of results Endogeneity issues
Single or multiple case study/ies	Context-specific	Bottom-up	Qualitative/ interpretive	Literature reviews Monitoring data Interviews Focus groups Surveys	Yes (variable degree)	Neutral	In-depth, rich exploration Placed in institutional context Suited for complex situations where various factors/processes interact with each other. Information on mechanisms Internalises stakeholder views	Context specificity Limited generalisability Risk of bias
Theory-based	Context-specific	Bottom-up	Theory-mapping Process-tracing	Literature reviews Monitoring data Interviews Focus groups Surveys	Yes	Part of the process	Focuses on the mechanisms that explain policy changes Produces a narrative that can explain hidden goals of policy	Conceptual clarity can be weak Does not always take account of exogenous factors Limited experience & results in practice

Source: Adapted from Hoerner and Stephenson (2012) *op. cit.* and Polverari and Bachtler (2014) *op. cit.*

Methods are of course dictated by and large by the evaluation questions, and each presents specific strengths and weaknesses, as summarised in the table above. However, different countries appear to have different preferences when it comes to the evaluation methods utilised to appraise the effectiveness of regional policy, with Germany and the Netherlands edging predominantly towards quantitative methods (though there are also qualitative evaluations and these are considered useful too), and France utilising especially qualitative methods (such as online consultations and focus groups - often not evaluations in the strict sense, but critical reviews of policy instruments based on monitoring and audit data). However, by and large countries utilise both approaches, and methods are generally adapted to the scope and focus of evaluation. Theory-based evaluation is being introduced in Norway and Poland.

Meta-evaluations are not evaluations in the strict sense but are useful in terms of elaborating evaluative evidence in a format that addresses the knowledge needs of politicians or policy-makers. Two recent examples from Poland and the UK relate to the effectiveness of EU Cohesion policy (see Box 6 and Box 7 below). This type of evaluation is set to become increasingly utilised in Poland.

Box 6: Meta-evaluations on the effectiveness of Cohesion policy in Poland

Evaluation of effectiveness was a crucial theme in the Polish presidency of the EU Council in 2011. Poland organised a series of high-level conferences and seminars to highlight and discuss the key issues. This included a conference on 'Evidence-based Cohesion policy' held in Gdansk in July 2011. The conference aimed to provide a forum for an informed discussion on how to use evaluation evidence to improve policy performance. It was used to launch the results of a meta-analysis of evaluation and other empirical studies (from the European Commission and Member States) organised by the Polish Government and conducted by a team of external experts. Almost all EU countries (24 countries) contributed to the project by sharing the most important evidence concerning the effectiveness and efficiency of Cohesion policy interventions.

Around 200 evaluations were analysed. The main aim of this document was to initiate evidence-based debate on how Cohesion policy could contribute to the successful implementation of the Europe 2020 Strategy. The second aim was to further explore the evaluation evidence available for the process of decision-making in order to focus Cohesion policy on results. The results of the evaluation process were presented in the report in line with the three main priorities of the Europe 2020 Strategy: smart, sustainable and inclusive growth.

Conclusions - Under the heading Smart Growth, the research found that non-repayable support is more effective for innovative, R&D intensive projects far from the market, while financial engineering is more appropriate when financial market failure is causing a lack of private financing. Under Sustainable Growth, it was found that the most effective and desirable forms of stimulating green technologies were: financial support for demonstration projects (the effect of promotion and incentives); wide dissemination of knowledge about the most effective solutions; and proactive advisory support. In terms of Inclusive Growth, it was found that interventions concentrated on early childhood are the most cost-effective and the most needed, as they translate into higher employment, lower school drop-out rates and lower poverty.

Source: Ferry M (2011) Building a place-based model: regional policy developments in Poland, 2009-2010, EoRPA country paper, European Policies Research Centre.

Box 7: Meta-evaluations on the effectiveness of Cohesion policy in the United Kingdom

In the United Kingdom, a recent report on the impact of Cohesion policy for the Balance of Competences Review initiated by the UK Government assessed the effect of EU policies and legislation on the UK. The Cohesion review assessed the evidence concerning the extent to which the objectives of the relevant policies have been met, and whether these policies and funds have delivered value for money. The findings conclude that the effectiveness of Cohesion policy has been appraised using four main methodological approaches, which lead to varying results in terms of impact: (i) the two main macro-economic models applied to SCF funding – HERMIN and QUEST – find clear positive effects in the net recipient Member States; (ii) results from econometric regression analysis tend to vary widely depending on the technical specifications applied, and the time series and country/regional datasets used; (iii) micro-economic studies using control groups have tended to report positive, but differing, results; and (iv) case study evaluations also generally show positive effects, although these are often stated in gross terms.

Some of the conclusions of the report are that: value for money of EU Cohesion policy is not fundamentally different from domestically funded policy initiatives; EU Member States derive benefits from Cohesion policy investments undertaken in other countries; the multi-level governance model of policy implementation pioneered in EU Cohesion policy is one of the policy's main areas of added value and is credited with having a significant impact on regional policy practice in the Member States and regions; there has been an excessive dispersion of Cohesion policy funding across too many goals and fields of interventions in many countries.

Source: Polverari L and Bachtler J (2014) Balance of Competences Cohesion Review, BIS Research Papers, February 2014, Department for Business Innovation & Skills.

4. CONCLUSIONS

4.1 The limits of existing approaches

Regional policy-makers in EoRPA countries have highlighted several limitations that constrain the evaluation of effectiveness and the influence of evaluation on policy. These are partly **institutional problems** linked to the context in which policy is conceived and implemented, such as the following issues.

- *There is a lack of experience and understanding of 'evidence-based' policy.* This is reflected in weak connections between those responsible for policy design and development and those responsible for evaluation and analysis. Thus, programmes and initiatives may not be designed with evaluation of effectiveness in mind (e.g. vague or contradictory programme objectives).
- *Evaluation is often regarded as a response to short-term obligations or information needs rather than taking a long-term perspective.* This limits the scope to evaluate effectiveness, and significant elements of the objectives of the initiatives may be omitted from evaluations.
- *Political reluctance or scepticism.* A common problem is a reluctance to make use of results that are critical, or there may be political pressure concerning the form and timing of publication of evaluation results.

Second, there are **systemic problems** related to the nature of the policy or its implementation. These include lack of evaluation capacity, either within the public administration or the evaluator market. Also, there is often a lack of relevant data, indicators or systems for generating the basic information needed for evaluation. And the growing breadth of regional policy interventions can make it difficult to produce evaluation findings on the aggregate outcomes of policy interventions.

Third, a **contingent problem**, linked to financial austerity, is the lack of appropriate financial and human resources to carry out evaluation to the scale and breadth that would be necessary to support policy-making.

4.2 The impact of evaluation on policy

Previous EoRPA research indicated a limited influence of evaluation on policy. It was noted that 'evaluation generally does not play a major role in determining the overall direction of policy. Rather, evaluation is used to justify political objectives and policy choices....Although the influence of evaluation on the overall direction of regional policy may be limited ... evaluation can play a major role in decisions on policy processes or instruments – ranging from the rationale accorded to a policy, programme or scheme, the policy mix chosen to achieve certain policy goals, to the operationalisation of specific instruments'.⁴²

This conclusion remains valid. There are examples of evaluation evidence having an impact, notably with regard to:

⁴² Polverari and Bachtler (2004) *op. cit.*

- the **termination or continuation of policy instruments**, as with the scheme for new firms 'Nykvest' in Norway, which was found to be redundant by an evaluation study and thus abolished, or the continuation of GRW aid to business investment, which recent evaluations (2010 and 2013) found effective;
- **adjustment of project-selection procedures** based on the effectiveness of the interventions, as in Poland (see Box 8 below);
- **structural institutional change**, as in England, where NAO criticism across a number of reports related to local growth policy, according to which economic growth initiatives were developed in isolation, resulted in the merger into a Single Team, with a single policy director, of the implementation structures responsible for local economic growth and the City Deals, involving three different government departments (BIS, DCLG and the Cabinet Office City Policy Unit).⁴³

However, these examples are only a handful, and even though impact on policy is intended to be a cornerstone of evaluation in a number of countries, whether this ambition will be realised is an open question. Specifically with regard to effectiveness, being able to get a full picture of the effects of the policy remains a key difficulty. Evaluating specific instruments can be relatively straightforward (e.g. business support), but grasping the full effects and impact of policy remains methodologically challenging. This limits the degree to which policy lessons can be drawn from evaluation.

Box 8: Evaluation-led change in Śląskie (Poland)

At the programme level, a 2012 evaluation assessing the impact of the Śląskie ROP 2007-13 looked at which approaches to project generation and selection (i.e. pre-selected sub-regional project platforms, key projects or competitive calls) were most effective. Effectiveness was measured in terms of contribution to the associated indicators and targets in the ROP priorities. For each of the different approaches, the evaluation looked at the percentage of indicators achieved by a specific approach and the percentage of the allocation used by a specific approach. On the basis of this analysis, the evaluation concluded that projects in sub-regional project platforms were the most successful. This was particularly the case for certain 'network-type' projects developed in response to complex challenges in a given territory. These required integrated approaches implemented based on cooperation and mutual complementarity in fields such as urban revitalisation and integrated local transport systems. Looking ahead to the 2014-20 programme period, the evaluation recommended that the selection of part of the projects under these platforms should be continued, and as a result they will be implemented in the ROP 2014-2020.

Source: EoRPA research.

4.3 Measures to improve evaluation and future plans

There is a general consensus that regional policy evaluation will be high on the policy-maker agenda during the 2014-20 period because of the evaluation expectations from EU Cohesion policy, the EU Regional Aid Guidelines, and a domestic political drive in a number of countries (most notably in France and Poland).

Some EoRPA countries are improving the **organisational arrangements** to deliver increased evaluation activity at different levels, such as the creation of a performance management group in the UK Department of Business, Innovation & Skills to improve the coordination of evaluation across

⁴³ The Cabinet Office's Cities Policy Unit was created in 2011. It is made up of civil servants and staff seconded from local government, think-tanks and the private sector. It works jointly with BIS and DCLG and across Whitehall.

nations and funds, and the requirement for Local Enterprise Partnerships to develop evaluation frameworks.

Further, efforts are being made to improve the **data and methods** for better evaluation, such as: the intention to make more use of mixed methods that combine quantitative and qualitative approaches in Finland; more use of impact evaluation in Sweden; better reporting of data (to allow more econometric research) and methodological improvements (e.g. control-group use, self-selection bias) in Norway, particularly to allow comparisons between programmes. Other examples can be found: in the Netherlands, where a working group within the DG Enterprise and Innovation produced a report on the different methodologies to be used to measure the effectiveness of instruments ('Dare to measure'); in Switzerland, where SECO has been working on an impact assessment model (*Wirkungsmodelle*) that will be used from 2016 to evaluate ex-post the 2008-2015 NRP; and in Poland, where the national Evaluation Unit plans to support more advanced evaluation methodologies, particularly counterfactual analysis and theory-based evaluation, through training and methodological guidelines by the national Evaluation Unit.

There are, however, concerns in some countries in relation to the financial and human resources available for evaluation, as well as the limitations entailed by pressing and competing demands upon administrations.

Lastly, there is an interesting development in the use of more **experiential knowledge** as part of evaluation efforts. This involves creating organisational arrangements for capturing and sharing knowledge from practice. Such arrangements are particularly characteristic of multi-level governance systems such as those of federal countries, evident in the examples of the ÖROK (Austrian Spatial Planning Conference) in Austria, the *regiosuisse* in Switzerland, and (perhaps) the GRW Planning Sub-Committee in Germany. The use of What Works Centres in the US and, more recently, in the United Kingdom indicates a trend towards combining both external evaluation (and academic research, to an extent) with experiential knowledge in ways that provide quicker and potentially more relevant insights into policy performance that may have a greater influence on policy formation and delivery.

4.4 Questions for discussion

In the above context, questions for discussion at the EoRPA meeting are:

- What kind of evaluations (conceptual approaches, methods, outputs) are most useful for regional policy-makers?
- What would favour a better integration of evaluation activities into the policy cycle? For instance, what organisational arrangements or capacity-building activities would enhance this integration?
- Two trends identified in this report are a shift towards integrative approaches and towards internalising evaluation activities. What are the risks and benefits of these trends?

- Regional policy-makers have to make judgements based on imperfect and at times contradictory evidence: Is the use of experiential knowledge seen as valuable for assessing the effectiveness of regional policy?