INTRODUCTION

During the past 15 years, the Cohesion policy of the European Union (EU) has become one of the most intensively evaluated policies in Europe. Since the reform of the Structural Funds in 1988, successive phases of regulation have created an increasingly rigorous system of appraisal, monitoring and evaluation covering all EU-funded regional development interventions. Evaluation of Structural and Cohesion Funds programmes has to be conducted at defined points in the programming cycle: ex ante to verify targets; at the mid-point to establish the need for corrective action; and ex post to assess outcomes. These national and regional evaluations are complemented by meso- and meta-evaluation studies and thematic evaluations by the Commission services, and by extensive Commission-sponsored research and debate on evaluation concepts, methods and practices.

The growing importance accorded to the evaluation of EU Cohesion policy is part of a broader international ‘boom’ in policy and programme evaluation, reflecting a perceived need to legitimise and justify government intervention - albeit with major differences between countries (POLLITT, 1998; FURUBO J-E, RIST R.C. and SANDAHOL R, 2002). In the context of the EU, two factors help to explain the trend. First, the Cohesion policy of the European Union has become, in budgetary terms, the most significant area of Community policy. In the mid-1980s, the European Regional Development Fund accounted for only 7.5 percent of the Community budget, (2.3 million ECU at 1985 prices) (MICHIE and FITZGERALD, 1997) and was often dismissed as a ‘side payment’ (MARKS, 1992; POLLOCK, 1995; ALLEN 1996). In the 2007-2013 period, by contrast, the Structural and Cohesion Funds will comprise 36% of planned Community spending (€308 billion in 2004 prices) (EUROPEAN COUNCIL, 2005). Second, the 1988 reform of the Structural Funds gave the European Commission much greater influence in the distribution of regional development funding, in particular with respect to the designation of eligible areas, the approval of Member State development plans, the management and delivery of programmes and the control of expenditure. This influence has often been a source of tension between the Commission and Member State governments, with protracted arguments over the spatial and thematic allocation of funding (BACHTLER and TUROK, 1997; BACHTLER and WISHLADE, 2004; 2005).

The combination of large amounts of expenditure and contested decisions on its usage has led to pressure for more accountability in spending, reflected in the creation of a steadily more
extensive EU evaluation regime, as well as more stringent obligations for the monitoring, financial management and auditing of expenditure. However, Cohesion policy evaluation has had to serve the objectives of many different organisations, ranging from programme managers and partners through regional and national government authorities to the various European institutions - the European Council, the Commission, the European Parliament and the European Court of Auditors - each with its own motives and interests in the results and implementation of EU spending.

Further complexity arises from the nature of the policy being evaluated. Structural and Cohesion programmes are implemented under a common regulatory framework but in widely differing national and regional circumstances with varied institutional arrangements for managing and delivering regional development policy. Programmes comprise a range of interventions - targeting physical infrastructure, economic infrastructure, business development, human resources, research, technological development and innovation, environmental improvement, tourism and community development - through a mix of financial instruments and many different types of beneficiary. Also, EU support has to be co-financed with national public or private funding that may originate in several different organisations or schemes.

With the greater importance attached to evaluation, the methodology for evaluating EU Cohesion policy itself has come under scrutiny and been subject to considerable debate. This is not surprising given the much increased scale and role of Cohesion policy, but it is clear that there is no uniformity of approach. Current evaluation methodologies range from the ‘bottom-up’, survey-based assessments of project and beneficiary outcomes to the ‘top-down’, input-output models of aggregate programme impact, as well as process studies of Structural Fund implementation. In addition, there are differences in the kinds of data collected, the sophistication of application and the kinds of question addressed. On the one hand, this is to be expected, given the diversity of policy instruments and purpose, and the differing institutional arrangements for management and implementation. It could even be viewed as a strength of EU Cohesion policy evaluation. On the other hand, it makes the evaluation of the policy appear uncoordinated, and raises issues about the comparability and consistency of results.

In fact, the results reported for the effects of EU Cohesion policy in the periodic reports of the Commission claim substantial levels of job creation, investment and other outputs attributable to the policy (EUROPEAN COMMISSION, 1996; 2001; 2004). However, the poor quality of monitoring data on which many evaluations have had to rely, the difficulty of isolating effects attributable to EU funding, and the other methodological problems and limitations highlighted above, mean that some of the reported results have been treated with scepticism, in particular the degree to which national or regional convergence is attributable to EU Cohesion policy (EDERVEEN S, de GROOT and NAHIUS, 2002; TARSCHYS, 2003; SAPIR, 2004).

Apart from disputes over the outcomes and impacts of Structural and Cohesion Funds programmes, there has also been debate about the less-tangible effects of EU Cohesion policy, referred to as Community ‘added value’. Under this, it has been argued that the regulatory obligations, combined with the role of the Commission in promoting ‘good practice’ in evaluation, have encouraged a greater commitment to, and use of, evaluation for the effective
management of programmes (EUROPEAN COMMISSION, 2002; 2004), although, again, the influence of the Structural Funds in this area has been questioned (ECOTEC, 2003). More broadly, evidence has been presented on the growing emphasis of Cohesion policy evaluation being carried over into national regional policymaking, with a more widespread evaluation of domestic interventions (BACHTLER and TAYLOR, 2003).

As the third programme period since 1988 draws to a close, this is a timely opportunity to take stock of the evaluation of EU Cohesion policy, to consider what has been learned and to identify research questions and policy issues for the future. This Special Issue of Regional Studies aims to provide an overview of the current state of EU Cohesion policy evaluation. The contributors are researchers from different academic disciplines and policymakers drawn from across the EU. To cover the variety of evaluative approaches, the instruments and circumstances in which Cohesion policy is applied, as well as different perspectives, the Special Issue is divided into four parts, some of which are cross-cutting in nature.

The first part (LEONARDI, MAIRATE) provides overviews from different perspectives on the rationale for EU Cohesion policy and its effects, including ‘added value’. The second part (BATTERBURY, BRADLEY, and MARTIN and TYLER) provides a critical assessment of Cohesion policy evaluation, focusing on the current regulatory framework. It examines contrasting macro-based approaches for determining the effect of policy, comprising a model-based approach to ex-ante evaluate the Structural Funds and a shift-share, residual-based method to evaluate interventions in Objective 1 regions. This part offers a new estimate of the employment effect of Cohesion policy. The third part of the Special Issue (FLORIO, BASLÉ, BLAZEK and VOZAB, ESER and NUSSMUELLER, and ARMSTRONG and WELLS) examines the evaluation experience in a variety of circumstances, including different policy measures (Structural Funds, Community Initiatives and Cohesion Fund) and different kinds of evaluation (ex ante, mid-term and ex post). It explores other issues, such as the use of cost-benefit analysis, the evaluation of community economic development initiatives and the preparation of programming documents in a new Member State. Finally, the fourth part has shorter contributions (BARCA, HUBER, JAKOBY and RAINES) that take a policymaker perspective on evaluation culture in a variety of EU regions and countries.

As a starting point for the Special Issue, this paper provides an introductory discussion of the main issues, setting the context for the subsequent papers and highlighting some of the main issues being addressed by contributors to this volume. It begins by tracing the evolution of Cohesion policy evaluation from 1988 to the present. The paper then discusses the concepts and methods of evaluation, the credibility of the results obtained, and the organisational and cultural differences in evaluation practice across the EU. It concludes with some questions on the way forward for evaluation.

**THE EVALUATION OF EU COHESION POLICY**

The current approach to EU Cohesion policy evaluation originated in the reform of the Structural Funds in 1988. Prior to the reform, the monitoring and evaluation of regional policy spending had been accorded a low priority and was subject to disagreement between the Commission and Member States over the necessity for evaluation, and how it should be organised and conducted (BACHTLER and MICHIE, 1995).
varied greatly between Member States, reflecting different policy and public management traditions and practices (MEANS, 1993).

The 1988 Regulations introduced a structured approach to the monitoring and evaluation of the Funds. They required partnership-based Monitoring Committees to be established for each Community Support Framework (CSF) and programme, and structural operations to be evaluated ex ante and ex post to establish their effectiveness with respect to economic and social cohesion, as well as the impact of CSFs and individual operations (EUROPEAN COUNCIL, 1988). In the subsequent programme period (1989-1993), the basic structures for monitoring and evaluation were established, but compliance with the requirements for prior appraisal was patchy and often ineffective (EUROPEAN COMMISSION, 1990; 1993). Many Member States were either unwilling or unable to provide projections on the outcomes of Community assistance, leading to the Commission undertaking its own evaluations to establish the likely effects of policy (EUROPEAN COMMISSION, 1992). Criticisms of the inadequacies of monitoring and evaluation were made by the Council, Parliament and Court of Auditors (BACHTLER and MICHIE, 1995).

In 1993, the Regulations were revised for the 1994-1999 programme period. They strengthened the monitoring and ex-ante and ex-post evaluation obligations of the Member States and stipulated specific requirements for the quantification of programme objectives, projections of impact and identifying the outcomes of previous EU-funded operations (EUROPEAN COUNCIL, 1993). An amendment to the Regulations also introduced the requirement for an ‘interim evaluation’ to be undertaken after three years to provide a critical analysis of progress and to allow adjustments to be made.

The Commission used these regulatory changes to exert a powerful influence on evaluation practice in the Community, in four main ways. First, the Commission created specialist evaluation units within its Directorates-General to coordinate evaluation and ensure that evaluation requirements are met by the Member States. Second, as part of the programming process for the 1994-99 programmes, the Commission obtained an independent ex-ante appraisal of each development plan submitted by the Member States in order to strengthen its influence in the negotiation and approval of plans (involving, inter alia, pressuring Member States to provide better and quantified indicators to permit monitoring). Third, the Commission significantly increased its own evaluation activity. Apart from ex-ante appraisals, it undertook ex-post evaluations of the 1989-93 period, as well as a series of thematic evaluations on subjects such as partnership, RTDI, SMEs, transport and financial engineering. The COURT OF AUDITORS (1998) estimated that expenditure on evaluation by the Directorate-General for Regional Policy (DGXVI) increased from €9.96 million in the 1989-93 period to €25.29 million in the 1994-99 period. Lastly, the Commission invested heavily in evaluation research through the MEANS (Evaluation Methods for Actions of Structural Nature) programme, which undertook wide-ranging analyses of evaluation practices across the Community, drafted guidance and evaluation manuals, and provided support for evaluators and evaluation commissioners, such as through ‘exchange of experience’ fora and evaluation conferences (BACHTLER and MICHIE, 1995; MEANS, 1999).

A further reform of the Structural Funds in 1999 introduced new regulations for the 2000-2006 period. With the aim of simplifying the implementation of the Funds, the regulations
decentralised several aspects of programme management to the Member States, including responsibility for monitoring and evaluation. The requirement to evaluate programmes ex ante, at the mid-term point and ex post was maintained, and in some respects was strengthened (EUROPEAN COUNCIL, 1999). The content of the ex-ante evaluation was more clearly specified with special attention given to environmental, labour market and equality issues. In an effort to improve the quality of monitoring information, the Commission produced extensive guidance on ex-ante evaluation and monitoring indicators. Further, the results of the mid-term evaluation were to be used for a mid-term review of programmes and a ‘performance reserve’ (constituting 4% of programme allocations) held back for allocation to the ‘best-performing’ programmes mid-way during the programme period. The mid-term evaluation was to be updated one year later to check that the evaluation recommendations had been implemented and to provide strategic input to the preparation of the 2007-2013 programmes. The Commission also continued to invest in evaluation studies and research, with further conferences and an updated guide on evaluation design, implementation, methods and techniques (TAVISTOCK INSTITUTE, 2003).

The current situation is that the regulatory obligations for evaluating EU Cohesion policy comprise three main stages of evaluation: ex ante, mid-term and ex post.

- **Ex-ante evaluation** - or, perhaps more appropriately, appraisal - has to be carried out as part of the programme planning phase. It involves a structured assessment of the social and economic situation in the programme area, with particular attention being given to the environmental situation and equal opportunities, the expected impact of the proposed measures, and an analysis of the relevance of the proposed implementation and monitoring arrangements. In the 1999-2000 programming phase, this was generally carried out by independent evaluators, either as a separate exercise or integrated within the programme development process. In previous programme periods, an independent ex-ante evaluation was undertaken by the Commission, but in the current period this responsibility was left to the Member States.

- **Mid-term evaluation**, as the name suggests, is carried out at the mid-point of the programme. In the case of the 2000-2006 period, the mid-term evaluations were mostly undertaken during 2003. Conducted by independent evaluators, the main purpose is to establish whether the programme strategy is still relevant in the light of economic and social trends, and to assess the initial results of assistance and the extent to which targets have been attained. In some respects, this can be regarded as the most valuable part of the evaluation system, since it enables the programme to be adjusted in the light of experience.

- **Ex-post evaluation** is the final stage in the evaluation process. Whereas the other two stages are the responsibility of the Member States, ex-post evaluation is the duty of the Commission, in collaboration with Member States. Carried out by independent evaluators, its remit is to examine the utilisation of resources and the effectiveness and efficiency of the assistance and its impact, with a view to drawing conclusions regarding policy in economic and social cohesion. It is these studies that form the basis for the periodic Cohesion Reports published by the European Commission. Although comprising the most comprehensive and detailed assessment of the outcome of EU Cohesion policy, ex-post
evaluations are clearly limited by problems by data availability and reliability, and methodological difficulties in isolating the effects of EU support from other influences, especially in areas where Structural Funds play a relatively small role.

Reflecting on the evolution of Cohesion policy evaluation since 1988, it is possible to identify several important trends and changes (BACHTLER, 2001; TAYLOR, BACHTLER and POLVERARI, 2001; TAVISTOCK INSTITUTE, 2003). First, evaluation has been fully integrated into the programme cycle from the start. Driven by regulatory changes and the influence of the Commission, programme management authorities have increased the range and quality of monitoring indicators, benchmarks and targets, as well as monitoring systems and organisational arrangements to deliver more reliable data. More evaluation information is becoming available and is being used for the planning of EU-funded interventions, through the implementation phase of programmes, to the ex post evaluation of outcomes.

Second, evaluation has evolved “from externally imposed evaluation obligations to internally driven demand for evaluation coming from programme managers and policy makers themselves” (TAVISTOCK INSTITUTE, 2003, p.12). Instead of being regarded solely as an administrative requirement for accountability purposes, evaluation is becoming more widely regarded as a management tool that can support the delivery of programmes and contribute to knowledge management.

Third, over consecutive programme periods, the gradual acceptance of the partnership principle and its application to all stages of programme management have extended the range of partner organisations involved in the design and implementation of evaluations. Evaluation studies have also served as a communication tool to engage local authorities, economic and social partners, voluntary and community groups and environmental bodies in the design and management of programmes.

Lastly, notwithstanding these trends, there are still important research questions and policy challenges concerning the future development of EU Cohesion policy evaluation. Some of these are measurement issues related to the deficiencies of evaluation data that could be resolved through changes in monitoring practices. There are also important debates about the philosophical approaches, rationales and methods of Cohesion policy evaluation. Other issues are intractable, deriving from the design of EU Cohesion policy, the different interests of European institutions and Member State governments, national evaluation cultures and resource constraints on evaluation. These issues are addressed in the following sections of this Introduction.

**EVALUATION CONCEPTS AND METHODS: POSITIVISM VS REALISM**

At the heart of many of the debates and conflicts about the evaluation process are the differing philosophical traditions underpinning evaluation. The Commission-sponsored guide distinguishes between three main traditions (TAVISTOCK INSTITUTE, 2003): the classical, positivist approach based on the premise that objective knowledge can be obtained through observation; the more recent, realist approach to evaluation which explores the mechanisms that account for changes in policies and programmes through social enquiry among practitioners; and constructivism, which rejects objective knowledge and involves evaluation through orchestrated interaction
with stakeholders to understand the different views, values and interdependencies. These traditions, discussed in several papers in this volume (BATTERBURY; ESER and NUSSMUELLER; ARMSTRONG and WELLS), underlie the different purposes and methods of Cohesion policy evaluation.

The longest standing purpose of evaluation of Cohesion policy is accountability, with a view to making judgements on the economy, efficiency and effectiveness of Structural Funds. Adopting a positivist approach, so-called ‘summative’ or ‘allocative’ evaluations have been conducted with a view to justifying the use of Cohesion policy expenditure (to the Council, Parliament and Court of Auditors), measuring what has been achieved in relation to targets, and determining the appropriateness of the policy and specific types of intervention (in particular, for the Commission). For the larger programmes (in the Cohesion countries and major Objective programmes), macro-economic models have been used for ex-post evaluation - notably HERMIN (see BRADLEY in this issue, and ESRI, 2002), PARADISE (BLAAS, 1991), QUEST (RÖGER, 1996) and dynamic input-output analysis (BEUTEL, 1997) - but also for ex-ante purposes in making projections of impact for the EU15 for the 2000-2006 period (BEUTEL, 2002) or in the new Member States in 2007-2013 (BRADLEY and MORGENROTH, 2004) to establish impacts on GDP growth and employment. Techniques involving the use of control groups and other statistical methods have also been used, but much more rarely (MEANS, 1999). In smaller programmes, a range of bottom-up techniques has been used to trace the micro-economic consequences of programme measures and activities through a combination of programme monitoring data, surveys of beneficiaries and regional or sub-regional statistics, principally to capture employment effects (EPRC and FRASER OF ALLANDER, 2000).

These methods involve a range of methodological and measurement difficulties. In the case of Objective 1, BRADLEY concludes that macro models are potentially capable of extracting the pure Structural Funds policy impacts from the background of all the other domestic and external shocks that affect the economy at the same time. “However, such results only provide imprecise, ball-park impact estimates, and the challenge remains that of reconciling bottom-up, micro analysis with top-down macro-analysis” (p. ??). A key constraint is the quality of available data, which according to MARTIN and TYLER “varies considerably from being very poor to (at best) barely adequate” (p. ??).

These problems increase in the evaluation of interventions outside Objective 1. Based on experience of ex post of evaluation of Objective 2, BASLÉ argues that it is near-impossible to obtain “a reliable and credible assessment of impact and true legitimisation for the EU interventions” (p. ??), particularly because of the multiplicity of actors, the difficulty of isolating cause and effect and the locality of actions and areas. Establishing effects in such contexts would require a better ‘logical diagram of impact’ (i.e. a theory of cause-and-effect), greater knowledge of the key factors determining success and failure, and better regional quantitative and qualitative data from better management and monitoring.

The emphasis on measuring impact and performance, using such methods, is criticised by BATTERBURY as missing the main question of ‘why things work (or not) in a specific context’. Following a more realist evaluation approach, this type of evaluation question has grown in importance since the mid-1990s. ‘Formative’ evaluation techniques have been applied to Structural Funds programmes to examine the effectiveness and relevance of implementation
procedures with a view to improving programme design and delivery. The use of this approach, typified in mid-term evaluations and, to a certain extent, ex-ante evaluation, has been driven by the growing recognition of the usefulness of evaluation as a management tool and has been primarily of interest to programme managers and partners. The methods used are mainly stakeholder interviews and case-study research to assess programme management, partnership, project appraisal and selection and monitoring arrangements. Over time, such methods have contributed to an important ‘learning effect’ both within and across programmes (see MAIRATE).

The ‘learning’ function implicit in formative techniques has not been without problems. ESER and NUSSMUELLER use the example of the mid-term evaluation (MTE) to illustrate the conflict between accountability and learning. When the MTE was included as part of the EU evaluation structure in the mid-1990s, insufficient consideration was given to the question of whether the main function and purpose of the evaluation was to justify expenditure or to learn, and whether the role of the evaluator was to be a judge or moderator. Such tensions were heightened in the current programme period (2000-2006), with the decision to make the allocation of a performance reserve conditional on programme achievements at the mid-point of the period, to be determined on the basis of the MTE results. These problems are at least partly inherent to the ‘hybrid governance’ structure of Structural Funds implementation which involves a mix of management by results and ‘bureaucratic rule steering’ (LANG, 2001).

Further, some regard evaluation practice hitherto as being largely ‘managerialist’ with the absence of a ‘democratic’ dimension that might open up the process to the wider community (TOULEMONDE, 2000). This is the main theme of the paper by BATTERBURY, which considers the evaluation framework for Cohesion policy to be limited to three core purposes – accountability, improved quality and performance, and improved planning – with other potentially useful functions neglected (i.e. capacity-building and learning). While the learning function would open up the possibility of carrying Structural Funds lessons into wider organisational and institutional systems, the Commission has not yet engaged sufficiently with this broader approach. Other, ‘pluralistic’ or ‘participatory’ evaluation techniques to develop a sense of ownership among programme participants or to empower specific groups of individuals have also been largely ignored.

Lastly, it is important to note that Cohesion policy evaluation is conducted primarily at the programme, regional or Member States level. The paper by FLORIO draws attention to the critical role of project-level evaluation in the context of the Cohesion Fund. This finances large infrastructure investments in transport and the environment, which potentially have long time horizons, and for which cost-benefit analysis is a formal requirement. However, market prices will generally be inadequate to examine the economic benefits and costs of these projects, necessitating the calculation of special ‘shadow’ prices that reflect the real costs of capital and labour, and of time itself, the social discount rate. This paper raises intriguing questions about the appropriate prices, which are by no means trivial.

While project-level evaluation is an obligation for the Cohesion Fund, ARMSTRONG and WELLS consider that there is a need, within the Structural Funds also, to “apply the full power of economic analysis” (p. ??) at the project level and within project design and monitoring processes. This is particularly relevant for community economic development (CED) initiatives
which have expanded within Structural Funds programmes over the past decade, but which according to these authors have not been adequately evaluated. ARMSTRONG and WELLS see potential for a third, constructivist evaluation approach, within the mix of philosophical traditions, as having potential merit in order to overcome the general hostility of CED organisations to orthodox economic evaluation.

**THE CREDIBILITY OF EVALUATION RESULTS**

As noted above, the origins of Cohesion policy evaluation lie in the results-oriented approach to Structural Funds management introduced by the 1988 reforms. Despite shifts in the conception, methods and organisation of evaluation, the need to justify the results of policy and account to the Council, Parliament and Court of Auditors remain important objectives, especially through the ex-post evaluations (BOUGAS, 2001).

As LEONARDI points out, EU Cohesion policy has provided significant amounts of investment (5-5% of national GDP) in the Cohesion Countries and has “favoured the convergence of less-developed regions towards the EU mean in terms of GDP per capita rates of annual economic growth, employment levels and unemployment during the 1990s at both national and regional levels” (p. ?). According to the evaluation research reported in the Third Cohesion Report (EUROPEAN COMMISSION, 2004), GDP in real terms at 1999 was between 2.2% and 4.7% higher than it would otherwise have been in the four EU15 Cohesion countries – Greece, Spain, Ireland and Portugal. The analysis of employment change in assisted regions conducted by MARTIN and TYLER also shows good evidence of a relative turnaround in employment growth in Objective 1 regions in the 1990s compared to the 1980s, i.e. the period when Structural Funds support was increased. They estimate employment creation in the order of one million jobs at 2002.

Nevertheless, qualification of such figures for impact is important. The paper by BRADLEY highlights the importance of taking account of other factors; research suggests that the direct impacts of the Structural Funds in isolation are modest, and that the real, long-term benefits of EU Cohesion policy are associated with the responsiveness of lagging economies to external opportunities for trade and investment. In similar vein, research by EDERVEEN et al (2002) found that European support as such did not improve national growth performance but that it ‘enhanced’ growth in countries with the ‘right’ institutions. With respect to other effects, BOLDRIN and CANOVA (2000) found that “there is no statistical evidence that Structural Funds have had a positive impact upon the growth rates of either labour or total factor productivity in the poorer regions” (p.242). Regarding to the location of economic activity, MIDEFART-KNARVIK and OVERMAN (2002) concluded that, while the Structural Funds expenditure had attracted R&D-intensive industries to the less-favoured regions, this effect had mainly acted counter to regional and national comparative advantage (except in the case of Ireland).

In the context of Objective 2 region, Commission-sponsored research has attempted to calculate employment effects associated with Structural Funds during the 1990s (ERNST & YOUNG, 1997; CSES, 2003). After making adjustments for ‘additionality’, displacement and indirect effects, ERNST & YOUNG (1997) estimated some 450,000-500,000 net additional jobs to the Objective 2 programmes in the 1989-93 period, while CSES (2003) estimated net job creation at 560,000 for the 1994-999 period. Such estimates are, however, hedged with major reservations about the quality of the data, as noted above. In the UK, for example, research on
the quality of monitoring data for the 1994-1999 period demonstrated significant deficiencies in data collection and interpretation practices and the realism of impact estimates (EPRC and FRASER ASSOCIATES, 2002). Problems of disentangling Structural Funds effects from other interventions in countries with small amounts of EU funding are also significant; research in Sweden, for example, was unable to identify any discernible effects of EU programmes on regional development indicators such as income per head, population and employment (OXFORD RESEARCH et al., 2004). In addition, BASLÉ points to the possibility of an ‘announcement’ effect, rather than an ‘implementation’ effect, and casts some doubt on the latter without suitable calibration of the results, i.e. cross-checking of micro-based results against relevant and objectively-measured aggregate-level indicators.

The effects of the Structural Funds are not just identifiable in terms of cohesion impacts. MAIRATE cites influences related to policy, implementation rules, learning and visibility, although varying according to the programme scale and type, administrative systems and maturity. Both MAIRATE and LEONARDI cite the importance of the multi-level governance dimension of Structural Funds in enabling previously excluded actors (i.e. sub-national authorities, socio-economic partners, community groups) into the policy process and for building institutional and administrative capacity for economic development in the regions. Cohesion policy is also regarded as a political expression of ‘solidarity’ between Europe’s wealthier and poorer regions.

**DIFFERENCES IN EVALUATION PRACTICE**

*Evaluation culture and capacity*

There are major differences in the way that evaluation has evolved and is used in Member States, and in particular the way in which evaluation is integrated within the politico-administrative culture of individual countries (TOULEMONDE, 2000). This is reflected in the different approaches to the evaluation of regional policy and responses to the evaluation obligations relating to Structural and Cohesion Funds (BACHTLER, 2001).

Historically, evaluation experience in the EU15 has been presented in terms of a ‘north-south’ divide (WADLEY, 1986; TAVISTOCK INSTITUTE, 2003). In parts of northern Europe – the Netherlands, UK, Germany and the Nordic countries – the evaluation of regional policy has a stronger tradition, with regular evaluations of national regional policy interventions (TOEPEL, 2000; TAYLOR et al., 2001). By contrast, in Italy and Spain, policy evaluation has not been well-established within the public administrative culture (CASAVOLA and TAGLE, 2003; OUSNA and MARQUEZ, 2000). In France, the focus of research has been on the political process of producing policy rather than the results, and while the evaluation of regional policy has grown since the mid 1990s, the ‘managerialist’ accountability approach has been contested (FONTAINE and MONNIER, 2002; BARBIER, 2003).

These differences in evaluation culture are reflected in varied evaluation capacity across the EU, such as the availability of qualified and experienced evaluators, and the presence of appropriate knowledge and skills among evaluators and evaluation commissioners, and the institutional frameworks to manage evaluation and promote evaluation practices; a point that is made by a number of papers in this Special Issue. Such factors have influenced the way in
which Member States have responded to the obligations to evaluate EU Cohesion policy. During the 1990s, countries such as Greece, Italy, Portugal and Spain restricted their evaluation activity to complying with the accountability requirements of the regulations, while programme management authorities in other countries (such as Ireland, Denmark and the UK) used evaluation more actively for improving the management of the Funds (BACHTLER, 2001).

The development of evaluation in the new Member States is a case study of these issues. The first experience of Structural Funds evaluation has been the ex-ante evaluation of the 2004-2006 programmes. In the context of the Czech Republic, as well as other new Member States, BLAZEK and VOZAB (this volume), conclude that the ex-ante evaluation, while effective in some respects, was constrained by a shortage of experience and knowledge of practices and techniques, the lack of clarity in the roles and responsibilities of evaluators and programmers and the absence of an institutional structure to facilitate evaluation. The timing of the evaluation findings was also critical: unless results were generated relatively early in the programming process, they would not be considered.

**Influence of EU evaluation practice**

The last question is: what effect EU evaluation has had on regional development policies and programmes? Previous research argues that the evaluation obligations of EU Cohesion policy have acted as a ‘driver’ of policy and evaluation in the Member States (TAYLOR et al., 2001; TAVISTOCK INSTITUTE, 2003). The final papers in this volume consider this question from the perspective of policymakers in four EU countries. In general, they suggest that the EU evaluation requirements and practice have influenced policy choices, enhanced the role of evaluation as part of the policy process and stimulated policy learning.

In Italy, BARCA states that EU evaluation obligations have been used as part of a conscious strategy to improve policy verifiability and to modify policy objectives and instruments. It has been part of a more general strategy of capacity building. According to JAKOBY, in the North-Rhine Westphalia region of Germany, evaluation has been a catalyst for changing programmes and management systems at different stages of the policy life-cycle. It has also initiated a learning process among stakeholders in the region and encouraged an exchange of experience at the international level.

From an Austrian viewpoint, HUBER identifies an added value for national regional policy through the EU evaluation requirements, especially with regard to evaluation methods. However, he also sees drawbacks in the overly bureaucratic and prescriptive way that the obligations are interpreted and applied by the Commission, and is of the opinion that ‘tacit knowledge’ is undervalued. RAINES echoes some of these points, from the experience of EU evaluation in Scotland. While the evaluation of the Structural Funds has played an important role in policy development and encouraged policy learning, as currently practiced, it does not have sufficient flexibility to adapt to national and regional local circumstances, policy needs and evaluation practice.

Finally, and perhaps contrary to the above, BLAZEK and VOZAB find that evaluation is unable to influence the conceptual basis, rationale or objectives of regional development policies, nor can it initiate major system changes. Likewise, SCHEFFEL (2001) notes the objection of the
German government authorities to evaluation that challenges the underlying policy objectives and expenditure choices. It illustrates the need for evaluation to operate within the prevailing policy and institutional paradigm, and its inability to question the theoretical propositions and assumptions on which the development and implementation of policy are based (SANDERSON, 2000). This is a point that is reinforced in the paper by HUBER and draws attention to the potential incompatibility of the different values, and wider social systems, within which politicians, administrators and evaluators operate and the different way that they perceive the logic of evaluation (HUMMELBRUNNER et al, 2005).

WHICH WAY FORWARD?

Looking forward, in many respects the evaluation of EU Cohesion policy is at a watershed. Over the past 18 years, an extensive evaluation regime has been constructed to account for spending on Structural and Cohesion Funds, to assess the economy, effectiveness and efficiency of the policy and to support the planning, management and delivery of programmes. This has been supported by extensive investment in evaluation methods, the promotion of ‘good practice’ and the progressive creation of an ‘evaluation culture’ in most Member States.

Notwithstanding this activity, the results of evaluation are not conclusive on the impact of Cohesion policy. With respect to the accountability objective of evaluation, on the one hand there is clear evidence of economic catching-up by poorer parts of the EU15 over a time period when major EU support was provided; research suggests that, at the very least, Structural and Cohesion Funds have helped to accelerate growth where the national policy environment was receptive. On the other hand, even in the Cohesion countries, where EU transfers account for a sizeable proportion of GDP and investment, impact estimates are considered imprecise, and the effects attributable to EU funding alone appear (according to some studies) to be limited. As the paper by BRADLEY indicates, there are major methodological challenges to be addressed if this situation is to improve - for example, through the quantification of programme objectives, better physical monitoring data, and the reconciliation of macro- and micro-economic methods for impact assessment. These are challenges for the Commission, but also for the research community. At present, independent (i.e. not officially commissioned) regional economic research on the effects of Structural Funds is limited by comparison with research on national policies or research in the political science literature on issues such as governance and partnership. The papers by LEONARDI, BRADLEY and MARTIN and TYLER identify a series of gaps where independent work is urgently needed.

This is not just an academic question. The recent debate on the EU financial perspective for 2007-2013 showed that spending on EU Cohesion policy is heavily contested by the ‘net contributors’, part of the criticism being justified on the grounds of inadequate policy performance (EDERVEEN et al, 2002; TARSCHYS, 2003; HMT et al, 2003; SAPIR, 2004). The December 2005 budget agreement includes a provision for fundamental review of the structure of the EU budget in 2008-2009 in preparation for negotiations on the post-2014 budget. Major questions on spending and policy direction will always be determined mainly by political factors, but there is a need for better evaluation and analysis to inform decisions.

A second set of research and policy challenges are how to make better use of evaluation at programme and project level. In this area, it is arguable that much of the evaluation effort is
currently wasted. As BASLÉ points out, considerable efforts are made to quantify programme performance in regions where the Structural Funds comprise a relatively small part of economic development expenditure, and where it is impossible to disaggregate the effects of different funding streams. This problem will be exacerbated in the 2007-2013 programme period, where many Regional and Competitiveness programmes will have less money than at present, potentially spread over a larger geographical area. Further, as ESER and NUESSMULLER note, there is an inherent conflict between the accountability and ‘learning’ function of evaluation.

A different approach would be to focus programme-level evaluation efforts more strongly on capacity-building and learning objectives (rather than accountability), as BATTERBURY recommends, with a view to assisting policy choices and the implementation of programmes. Decisions such as the approach to evaluation, the choice of objectives and methods, and the administration and timing of evaluation should be made within the programme partnership. This would allow for a more diverse range of methods and, as advocated by ARMSTRONG and WELLS, should encompass more evaluation at the project level and in areas such as community economic development which are not well-served by standard evaluation methods. The aim would be to gain more insight into the effectiveness of interventions and delivery mechanisms and to understand ‘success’ factors and pre-conditions. The role of the European Commission, under such an approach, would be to ensure that an evaluation system is in place and, as at present, to promote exchange of experience.

Lastly, the institutionalisation of an evaluation culture still varies greatly among Member States. As the papers by BARCA, JAKOBY, HUBER and RAINES illustrate, EU evaluation practices have brought tangible improvements to the design and delivery of EU programmes in different parts of the EU, but there remains an asymmetry in evaluation capacity and culture (BOUGAS, 2001). The major institutional challenges are in the new Member States, as outlined by BLAZEK and VOZAB, (and, to a certain extent, in the EU15 Cohesion countries) which need to invest in evaluation management units, and the evaluation skills and competences of future evaluators. In other parts of the EU, the situation is different. As Structural Funds focus more on the Lisbon/Gothenburg agendas in the next programme period (BACHTLER, GROSS and McMASCETER, 2005), EU funding may be used much more for specialized areas of support such as regional innovation networking, specialist advisory support, environmental sustainability, financial engineering, innovative applications in the sue of IT and specialist training. These interventions present particular implementation difficulties for programme managers and challenges for evaluators, requiring new thinking and more interchange in evaluation approaches and methods across policy boundaries.
ACKNOWLEDGEMENTS

The editors would like to thank all of the referees of the papers, who gave freely of their time, and whose efforts are greatly appreciated. They are also grateful for the secretarial assistance of Lynn Ogilvie and Rhona Walker of the European Policies Research Centre (University of Strathclyde). In addition, thanks to the *Regional Studies* Editorial Board at the University of Newcastle for their support and to the current and former Editorial Assistants at Newcastle, Gail Welsh and Sarah Wray, who processed the material efficiently and were able to offer excellent assistance.
REFERENCES


MEANS (1993) *Methods to give meaning to the evaluation obligation: the conclusions of the MEANS programme*, MEANS/93/12/EN, CEPOS, Lyon.


---

1 Cohesion policy refers to the Structural Funds and Cohesion Fund, which are the European Union’s main instruments for supporting social and economic restructuring across the EU at both regional and national levels. There are four Structural Funds: the European Regional Development Fund, the European Social Fund, the European Agricultural Guidance and Guarantee Fund: Guidance Section and the Financial Instrument for Fisheries Guidance, while the Cohesion Fund finances major projects in environmental and transport infrastructure. The Structural Funds are available in priority ‘Objective’ areas (i.e. Objective 1, 2 and or 3) and in support of Community Initiatives (i.e. EQUAL, INTERREG III, URBAN II and LEADER +). The Cohesion Fund is available in less prosperous Member States, which are currently Greece, Portugal, Spain and the ten new Member States.

2 A detailed discussion on the effectiveness and efficiency of EU Cohesion policy is contained in the ‘mini-theme special issue’ in *Regional Studies* 39(7), October 2005, as part of a debate on the report of the Sapir Group.
To see how these policy measures relate to Cohesion policy, see endnote i. Cohesion policy is wide-ranging, and, inevitably, for the purpose of this Special Issue it was necessary to be selective, so that some elements of Cohesion policy are not well represented, including the European Social Fund, agriculture, fisheries and rural development measures.

The latest being the Third Cohesion Report (EUROPEAN COMMISSION, 2004).

However, it points to a general problem of the comparability of results when different evaluation methodologies are used. MARTIN and TYLER measure the jobs in net cumulative terms at a point in time, but which is not directly comparable to the estimates in the Third Cohesion Report, which measure the net jobs over a period of time, i.e. a programme period (EUROPEAN COMMISSION, 2004).

Indeed, for the current period, the Italian government supplemented the EU performance reserve with a national reserve to increase the pressure on programmes to perform, and to raise the importance accorded to evaluation.

This is through a benchmarking project with Scotland as part of the mid-term evaluation.