This paper examines lessons to be learnt from the European Commission’s ex-post evaluation of the 1989-93 Objective 2 programmes.

In particular, the paper considers the extent to which it proved possible to translate the Commission’s approach to ex-post evaluations into practice, highlighting experience from the Objective 2 study that might be more generally applicable. The focus is on methodological issues, but the substantive results of the evaluation are also reviewed.

1. Ex Post Evaluation Framework

The European Commission defines the purpose of ex-post evaluation as being threefold. To quote from a working paper¹, the aim is to:

- determine the real and lasting results which can be actually attributed to intervention;
- compare the outcomes with the expected results defined at the ex-ante stage and possibly amended during implementation;
- assess the extent to which the assistance has actually achieved its objectives.

Guidance on ex-post evaluation methods has been provided by the European Commission but the various components have to be pieced together from quite a large number of different sources such as the various MEANS publications, the Commission’s own guidelines and previous evaluation studies. A further complication is that ESF and ERDF evaluation methods have been developed separately by DG V and DG XVI with little integration or cross-referencing of key methodological aspects.

A model of sorts nevertheless exists for the ex-post evaluation of Structural Fund programmes. This is underpinned by a ‘theory of action’ in which support for the development of productive structures, human resources and physical infrastructure has various short-term job and wealth creation effects, together with longer term structuring impacts that are manifest in the diversification of assisted regions away from economic structures based on traditional industrial activities². Ultimately, the impact of Structural Fund interventions is to be measured in terms of the contribution programmes make to regional convergence and the European Union’s economic and social cohesion.

The Commission’s model combines ‘bottom-up’ and ‘top-down’ methods to evaluate Structural Fund impacts. Figure 1 provides an overview of the various aspects of the methodology - as developed by Ernst & Young for the Objective 2 evaluation - highlighting the key features of the two dimensions. (It should be noted that the Commission’s various econometric models are not considered here).

**Figure 1: Structural Fund Ex-Post Evaluation Model**

### Top Down

- Measures difference between baseline situation and position at end of Structural Fund programme period using key indicators such as GDP, unemployment and industrial structures.
- Captures global impacts but does not identify contribution of specific measures.
- Based on published statistics relating to regional trends.
- Compares performance of assisted regions with non-assisted areas to help assess the added value of intervention and extent of convergence.
- Calculates gross outputs from a sample provides a check on top-down estimates of impacts.
- Identifies contribution of specific measures but difficult to capture wider programme impacts.

### Bottom-Up

- Supplements 'top-down' approach by providing insights to causality (the cause-effect 'black box').
- Based on empirical research to examine programmes. Grossing up programme outputs from a sample provides a check on top-down estimates of impacts.
- Assesses net added value using feedback from beneficiaries and comparison with non-assisted activities to assess additionality, displacement and indirect effects.
- Identifies contribution of specific measures but difficult to capture wider programme impacts.

The top-down dimension focuses on an analysis of the ‘before-and-after’ situation, thereby capturing the overall impact and net effects of policy intervention. However, the weakness of the top-down method is that as a result of both data limitations and the varying impact of other influences and trends, it is difficult to isolate particular Structural Fund impacts or to identify the relationship between cause and effect (the so-called ‘black box’ problem). The bottom-up technique overcomes some of these disadvantages since it enables specific Structural Fund outputs to be identified, provides empirical perspectives on the counter-factual and, hence, sheds light on causality. Ideally, the two dimensions will be brought together by aggregating outputs attributable to Structural Fund intervention at the programme level (bottom-up analysis) and then assessing what contribution these outputs make to changes in regional trends (top-down analysis).

A number of other features of the Commission’s model are worth highlighting. First, the Commission’s ex-post evaluation model is to be seen as part of three-staged process involving prior appraisal, monitoring/interim evaluation and ex-post evaluation. There are forward and backward linkages at each stage. Thus, ex-post evaluation relies on prior appraisal to define the performance indicators and targets that are needed to subsequently judge the success of programmes in achieving their objectives; likewise, interim evaluation should prepare the way for ex-post evaluation by ensuring that data on gross outputs is in place; whilst ex-post evaluation
takes the analysis forward by examining final outputs and net impacts, feeding best practice lessons back into the appraisal stage of new programmes.

A second feature of the Commission’s approach is that considerable emphasis has been placed on developing a standardised approach to Structural Fund evaluations. This is seen as being necessary to allow the relative effectiveness of different interventions to be assessed (the learning function of evaluation) and to make it possible to consolidate Structural Fund impacts at a European level (accountability function). Against this, some argue in favour of maximum discretion at a local and regional level to design studies so that they can be tailored to essentially local needs and priorities. In this sense the challenge is one of reconciling the principle of subsidiarity on the one hand, with the prescriptions of European legislation with regard to demonstrating accountability on the other, the latter presupposing a high degree of harmonisation in evaluation methods.

Third, the Commission’s model depends heavily on synthesis techniques. Here, it is helpful to distinguish between the concept of ‘external synthesis’ (or cluster evaluation) where the task is to integrate research from a series of standalone studies, on the one hand, and ‘internal synthesis’ where a common methodological framework exists from the outset, on the other. In both cases, the challenge is one of achieving a cumulation of evaluation findings across space and time, thereby achieving a transition from the description of particulars to generalisations that have universal validity. The concept of synthesis evaluation is in many respects far less developed that other more conventional types of evaluation. However, in the EU context it is in a key to producing results from evaluation studies that can feed into improvements in policy and practice.

Having briefly reviewed the Commission’s model, the question is: how does this approach to Structural Fund ex-post evaluation work stand up to experience in the field, what can be learnt from experience and to what extent are any shortcomings due to weaknesses in the model itself as opposed to failures of practical implementation? The Objective 2 evaluation sheds some light on theses issues.

2. Objective 2 Study

The ex-post evaluation of the 1989-93 Objective 2 programmes covered 60 regions from nine EU Member States. It was carried out in 1996-97 by a team of 25 evaluators with co-ordination at a European level by Ernst & Young.

The purpose of the study was defined by the European Commission as being to provide ‘an ex-post evaluation report of the CSFs and OPs relating to Objective 2 areas in order to assess the main results obtained during the two programming periods and to contribute to improving the functioning of current and future interventions’. In the subsequent elaboration of the study’s terms of reference, a number of key evaluation issues were highlighted as being particularly significant:

- What impacts were achieved by Objective 2 intervention?
- How did these impacts contribute to the industrial reconversion process in Objective 2 regions?
- To what extent did the 1988 Structural Fund reforms improve the way programmes were implemented and the results achieved?
- What can be learnt from the evaluation of the 1989-93 Objective 2 programmes that can be used as a guide to best practices in the future?
The Commission’s tender specifications grouped the sixty Objective 2 regions to be evaluated together into ten lots. Three of the lots consisted of regions from single Member States (France, Italy and the United Kingdom). But the others in each case brought together two or more Objective 2 regions from different countries in multinational groupings with broadly similar characteristics, namely the Atlantic Regions, Industrial Core North, Industrial Core South, Metropoles and North Sea regions.

Nine contractors were selected to co-ordinate the regional grouping with a second tier of researchers being sub-contacted by them to carry out the research in each of the 60 regions. Altogether, some 25 organisations were chosen to work on the ex-post evaluation - ten professional consultancy firms, nine universities and six independent research institutes. These organisations were selected by the Commission and contracted with DGXVI rather than Ernst & Young. The way in which the study was organised is shown below in Figure 2.

**Figure 2: How the Objective 2 Evaluation Was Organised**

There were three main components to the research:

- a general assessment of the before and after situation in each region;
- in-depth studies in 22 of the 60 regions covered by Objective 2 programmes;
- thematic studies to identify best practices in respect of particular types of interventions.

Below, we review each of these components, commenting first on the study findings and then on methodological aspects.

**General Assessments (Top Down Dimension)**

The ‘top-down’ dimension - general assessments covering all 60 Objective 2 regions - mainly focused on an analysis of trends in their performance using key economic indicators such as changes in GDP and employment levels. The rationale for Structural Fund intervention, type of strategies adopted, partnership structures and programme management practices were also examined.
The timing and intensity of the crisis affecting the coal, shipbuilding and textiles industries varied from one Objective 2 region to another. The general assessments highlighted the fact that by the end of the 1980s, industrial restructuring was well underway in most regions and their performance against key indicators was improving. The convergence in economic performance with the rest of the EU proved to be short-lived, however, with the situation worsening again in the early 1990s. From 1991 onwards, there was a marked slow-down in the EU economy with a net loss of three million jobs between then and the end of the period. Indeed, by the end of 1993, unemployment rates in the EU were back up to the peak levels of the mid-1980s.

However, the research suggested that compared with other EU regions, those covered by Objective 2 of the Structural Funds performed relatively well. Having fallen more rapidly than elsewhere in the late 1980s, unemployment rates in Objective 2 regions then rose more slowly during the recession of the early 1990s. As a result, the difference between the average unemployment rate in Objective 2 regions and for the EU as a whole narrowed from 4.0% (1989) to 1.7% (1993) whereas the differentials might have been expected to widen. Part of the explanation for this relatively favourable performance lay in demographic trends with the highly urbanised Objective 2 areas experiencing an ageing population and drop in new entrants to the labour market. Also, a significant proportion of those affected by job losses in traditional industries withdrew completely from the labour market. However, the impact of factors such as these does not detract from the fact that during the 1989-93 period, Objective 2 regions were more successful in generating new jobs than other EU regions.

The general assessments also included a review of the Objective 2 strategies and programmes, highlighting similarities and differences between regions with regard to factors such as appraisal methods, strategic objectives, types of measures implemented, the level of financial inputs and private sector leverage, programme management and partnership structures. This aspect of the evaluation was mainly limited to desk-based research, i.e. a critical analysis of programme documentation, existing studies and statistical information, but supported by some interviews with regional and national authorities.

By and large, the ‘top-down’ general assessments posed few significant difficulties. There were some complications in smaller Objective 2 areas where data on the situation at the beginning and end of the 1989-93 period could not be disaggregated down to a NUTS 3 level, at least on a harmonised basis. The fact that administrative boundaries, and more particularly the geographical units used for statistical purposes, did not coincide with the definition of some Objective 2 areas caused some difficulties. Also, even where we obtained a good analysis of regional trends for key indicators such as unemployment rates and industrial structures, there were gaps in the analysis with regard to some of the other performance measures that we had suggested should be used for the general assessments.

In Depth Studies (Bottom Up Dimension)

The ‘bottom up’ aspect of the methodology - the in-depth and thematic studies - was designed to provide insights to causality, i.e. to explain why the situation before and after Objective 2 intervention was different and to shed light on best practices.

The first step involved fieldwork in the Objective 2 areas to collect and critically analyse monitoring information held by regional authorities on Structural Fund outputs. It was envisaged
that the in-depth studies would also include survey work in 22 of the 60 Objective 2 regions to obtain feedback directly from beneficiaries. We placed considerable emphasis on this aspect of the research. Apart from investigating key evaluation issues in more depth, an important purpose of the survey work was to provide information on ERDF and ESF projects needed for an empirically-based estimate of the net outputs attributable to Objective 2 intervention. The survey work was also viewed as a bottom-up method of ‘checking’ the accuracy of Objective 2 programme monitoring data. Here, the task was to scale up the output estimates obtained from the sample of projects to provide an estimate for Objective 2 programmes as a whole, comparing the results with monitoring data. It was accepted that this scaling up exercise would have to be limited to a number of key indicators such as the number of jobs created or saved, the number of firms assisted and positive outcomes (jobs and qualifications) attributable to ESF-supported training programmes.

Turning to the results, it was estimated that between 1989-93 approaching 850,000 gross additional jobs were created as a result of Objective 2 intervention. After making adjustments for additionality, displacement and indirect effects this estimate was reduced to a figure of between 450,000-500,000 net jobs created, saved or redistributed. The average cost per net job was estimated at some ECU 13,500. Separate estimates were calculated for each Objective 2 region and the global estimate was very much a bottom-up calculation. Table 1 below provides a summary of the estimated employment outputs for each of the Member States containing regions that received Objective 2 assistance during the 1989-93 period.

<table>
<thead>
<tr>
<th>Region</th>
<th>ERDF and ESF Gross Jobs</th>
<th>ERDF Net Jobs</th>
<th>ESF Net Jobs</th>
<th>Total Net Jobs</th>
<th>ECU Cost per Net Job</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>68,901</td>
<td>25,816</td>
<td>10,572</td>
<td>36,388</td>
<td>5,870</td>
</tr>
<tr>
<td>Denmark</td>
<td>4,117</td>
<td>1,468</td>
<td>693</td>
<td>2,161</td>
<td>11,477</td>
</tr>
<tr>
<td>Germany</td>
<td>82,364</td>
<td>41,893</td>
<td>3,552</td>
<td>45,445</td>
<td>12,778</td>
</tr>
<tr>
<td>Spain</td>
<td>103,951</td>
<td>27,655</td>
<td>25,084</td>
<td>52,739</td>
<td>25,402</td>
</tr>
<tr>
<td>France</td>
<td>162,024</td>
<td>47,491</td>
<td>35,745</td>
<td>83,236</td>
<td>14,721</td>
</tr>
<tr>
<td>Italy</td>
<td>48,388</td>
<td>18,434</td>
<td>7,174</td>
<td>25,608</td>
<td>15,108</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>1,460</td>
<td>494</td>
<td>268</td>
<td>761</td>
<td>15,102</td>
</tr>
<tr>
<td>Netherlands</td>
<td>40,618</td>
<td>18,113</td>
<td>3,849</td>
<td>21,962</td>
<td>7,499</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>334,527</td>
<td>138,827</td>
<td>40,224</td>
<td>179,050</td>
<td>10,781</td>
</tr>
<tr>
<td>Total</td>
<td>845,988</td>
<td>320,190</td>
<td>127,161</td>
<td>447,351</td>
<td>13,696</td>
</tr>
</tbody>
</table>

Other measurable outputs included an estimated 470,000 SMEs that received Structural Fund assistance during the 1989-93 period and some 920,000 beneficiaries of ESF-supported training programmes.

Although a rather crude indicator, we used the cost per net job estimates to compare how well the different Objective 2 regions performed in using Structural Fund assistance. The analysis
suggested that there was a considerable variation, ranging from a cost per net job of around ECU 1,500 in Clwyd (UK) to almost ECU 60,000 in Lorraine (France). Further analysis suggested that this generally reflected the type of priorities pursued in the different Objective 2 regions. In particular, where infrastructure investment accounted for a high proportion of expenditure, induced job creation effects took longer to materialise than in regions where the emphasis was on measures such as support for SMEs which cost less and tended to have more immediate effects. However, there appeared to be other explanatory factors. In particular, smaller Objective 2 programmes generally produced better results than larger ones, possibly because partnerships tended to be more cohesive and programmes were easier to manage. Taking both EU and national expenditure on Objective 2 programmes together, the global or real cost per net job was around EUC 40,000 which is within the expected range for regional development interventions.

It was the ‘bottom-up’ research that posed the most serious complications. The first problem was that it proved difficult in quite a high proportion of the Objective 2 areas to obtain good-quality monitoring data on physical outputs from regional authorities. In only 42 of the 60 areas covered by the study was it possible to collect this information for both ERDF and ESF programmes. In some regions the data had simply not been collected whilst in others there seemed to be a reluctance to make it available or it was still in the process of being collated. In general, data on ESF outputs was more complete than that on ERDF outputs. Apart from poor quality data and evidence of double and triple counting, there was also the complication that few regional authorities adopted standardised definitions for employment outputs, the category ‘jobs safeguarded’ being subject to particularly broad interpretation.

Next, despite the emphasis placed on the importance of carrying out survey work, only 11 of the 20 evaluators who were involved in the in-depth research actually did so. In some regions quite extensive survey work was undertaken. Thus, in Piedemont (Italy), for example, the evaluators interviewed 46 project sponsors; in South Wales (UK) a sample of 42 projects accounting for a fifth of Objective 2 expenditure was examined; and in Twente (Netherlands) the research covered 30 ERDF and ESF projects. Where survey work such as this was undertaken, it provided very valuable first-hand insights on how Objective 2 funding had been used as well as shedding light on the question of additionality and value added. In several regions, the survey work was on a sufficiently large scale to enable bottom-up estimates of the physical outputs to be scaled up and used to check the accuracy of information obtained from the analysis of regional authorities’ monitoring data.

The most problematical aspect of the methodology, however, was the process of adjusting gross Objective 2 outputs to take into account additionality, displacement and indirect effects. Whilst we hoped that the survey work would shed some light on the extent of additionality, it was accepted that the parameters for displacement and indirect effects would probably have to be derived from secondary sources rather than new research in the regions. Nevertheless, very few evaluators were able to produce an estimate of net effects. In some cases, there appeared to be a complete lack of familiarity with the key concepts. For example, financial additionality was widely assumed to be the same thing as project additionality. Nevertheless, we did obtain sufficient information from the in-depth studies for an estimate net Objective 2 outputs to be made in the synthesis report. But this was based on assumptions that did not fully reflect the particular circumstances of the various Objective 2 regions.
As part of the research, the evaluators were also asked to examine the extent of synergy effects. This too proved difficult. In this context, synergy was defined as being a situation where projects and programmes combine to produce impacts that are greater than would be achieved if measures were implemented in isolation. For the Commission, there was particular interest in investigating the extent of synergy between the ERDF and ESF components of the Objective 2 programmes. It was agreed that the MEANS methodology for measuring synergy effects would be used in the evaluations. Because this methodology was still untested, the Objective 2 evaluations were seen as an opportunity to obtain feedback on the feasibility of applying the methodology in practice. However, only three of the 25 evaluators actually attempted to apply the MEANS methodology, the general view being that it was too complex and time-consuming. In another case, the co-ordinators adopted a modified version based on a scoring system to assess synergy (and other) effects but the basis of the scoring was not at all transparent. On the whole, better results were obtained where the evaluators adopted a more straight-forward approach.

Overall, there were some very good examples of the methodology for the in-depth studies being put into practice but as the assessment in the next section suggests, the performance of the evaluators in this and other respects was far from uniform across the different regions.

**Synthesis, Impact Assessment and Policy Recommendations**

The final stage in the ex-post evaluation involved synthesising the results of the research in the various Objective 2 regions, first at the level of the nine trans-national groupings and then at an overall EU level. To do this it was necessary to bring together the ‘bottom-up’ and ‘top-down’ aspects of the research by estimating what effect the sum total of net outputs attributable to Objective 2 intervention had on trends in key before-and after indicators for the various assisted regions over the 1989-93 period.

The overall conclusion of the study was that although macro-economic trends were unfavourable for much of the 1989-93 period, Objective 2 programmes nevertheless had a significant impact and helped to sustain the process of structural adjustment in regions affected by the decline of traditional industries. At a global level, the impact assessment suggested that as a result of Objective 2 assistance, unemployment rates in assisted regions were 0.8% to 1.4% lower by the end of 1993 than they would have been in the absence of intervention. Whilst the impact of Objective 2 intervention on unemployment is certainly an important yardstick for measuring short-term impacts, the strategies were essentially designed to address deep-seated structural weaknesses over a longer period. From this point of view, the contribution of Objective 2 programmes to the diversification of assisted regions and to reducing their dependence on traditional industrial activities as a source of jobs and income is a critical measure of success. Here, we estimated that Structural Fund intervention helped to promote economic diversification in assisted regions with the jobs created or maintained being equivalent to an estimated 0.7% to1.3% of non-industrial employment. Notwithstanding the fact that the reconversion process was already well underway by the end of the 1980s, it was clear from the evaluation that Structural Fund measures played a major role in sustaining this process through the recession of the early 1990s and helping to bring about convergence with the rest of the EU.

Considerable emphasis was placed on quantification - in particular, the need to estimate the employment effects attributable to Objective 2 interventions - but we accepted that there would be limits to this type of analysis and that the evaluations would also need to rely heavily on qualitative
judgements in arriving at overall conclusions. Amongst the most significant non-measurable Objective 2 impacts was the way in which the interventions stimulated the development of regional partnerships and their capacity to plan and efficiently manage economic development strategies. The evaluation points to this as being one of the most significant and lasting effects of Objective 2 intervention during the 1989-93 period. The repercussions of other Structural Fund reforms, especially the new emphasis on the principles of concentration, additionality and programming, were more difficult to assess but the study’s conclusions were basically positive.

In the final report, we concluded with an assessment of best practices, drawing mainly on the thematic research, and examined policy implications stemming from the evaluation. Policy recommendations included:

• improving the concentration of Objective 2 interventions;
• making project sponsors compete more for funding;
• enhancing the role of the private sector in regional partnerships;
• linking decisions on future funding more closely to past results (conditionality);
• improving financial sustainability by achieving a better balance between grant aid and loans;
• further improving appraisal, monitoring and evaluation methods and practices.

It is interesting to note that some of these ideas are now being discussed in the run-up to the 1999 Structural Fund reforms. Whilst we would certainly not claim that this stems directly from the Objective 2 evaluation, the research has perhaps added weight to the arguments being put forward to support change, particularly with regard to the principle of conditionality and financial sustainability.

Responsibility for synthesising the evaluation findings and for the impact assessments lay in the first instance with the nine regional co-ordinators. There was a very mixed performance in this respect.

3. Conclusions

So, overall how successful was the Objective 2 evaluation and what lessons can be learnt with regard to carrying out future studies?

On the positive side, the research was successful its key aim of providing a quantified estimate of Objective 2 net employment effects and the impact on industrial structures in assisted regions. Despite the imperfections of the ERDF and ESF output data, the study pushed forwards the Commission’s so-called ‘quantification exercise’ and from a methodological perspective, demonstrated the feasibility of applying a standardised approach to Structural Fund ex-post evaluation work reasonably consistently across a large number of regions with very differing traditions and levels of expertise in evaluation. Less positively, the thematic and best practice dimensions of the research were generally weak, perhaps underlining the difficulty in combining a macro-level impact assessment with detailed insights into activities at a project and programme level.

What could have been done to improve the quality of the work? Here, it is helpful to distinguish between the role of the evaluators, Commission and national authorities, and factors that were external to the evaluation process as a whole.
Role of the Evaluators

Turning first to the performance of the evaluators, Table 2 provides a summary analysis of how successfully evaluators from different countries completed the various methodological steps outlined in the previous sections. As can be seen, very few evaluators - only seven of the 25 organisations that worked on the Objective 2 study - successfully implemented all aspects of the methodology.

<table>
<thead>
<tr>
<th>Methodology/Regions</th>
<th>B</th>
<th>D</th>
<th>DK</th>
<th>E</th>
<th>F</th>
<th>I</th>
<th>L</th>
<th>NL</th>
<th>UK</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of General Assessments:</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>7</td>
<td>17</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>9</td>
<td>60</td>
</tr>
<tr>
<td>Number of In-Depth Studies:</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Step 1: Gross outputs based on monitoring data</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>10</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>9</td>
<td>41</td>
</tr>
<tr>
<td>Step 2: Survey of sample of projects</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Step 3: Scaling up of sample outputs</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Step 4: Adjustments for additionality</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Step 5: Adjustments for displacement</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Step 6: Adjustments for indirect effects</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Step 7: Comparison with targets</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Step 8: Assessment of synergy effects</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Step 9: Estimate of cost-effectiveness</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Step 10: Assessment of impact on regional trends</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>

The best results, from a methodological point of view, were demonstrated by the evaluators from Italy, Netherlands and the UK. Factors that are significant in explaining differences in performance include the type of organisation appointed to carry out the evaluation in each region, the willingness of regional and national authorities to collaborate, and whether or not interim evaluations had been carried out. Thus, taking the first of these factors, university-based evaluators generally produced the best general assessments whereas professional consultants proved better at tackling the in-depth studies, in particular the survey work and best practice analysis. However, overall, the analysis contained in Table 2 does not point to any straight-forward patterns in the performance of evaluators, least of all the notion of a North-South divide in evaluation expertise and traditions that has been postulated by some observers.

Given the very varied evaluation expertise, the role of co-ordinators of the nine regional groupings was critical. We estimate that around a quarter of the study budget was spent on co-ordination activities, which should have been more than sufficient to perform this function effectively. However, in retrospect, it is clear that some co-ordinators added considerably more value than others to the research carried out by their sub-contractors. In many cases, the co-ordinators’ final
synthesis reports did little more than summarise the key evaluation findings from the various Objective 2 regions with no attempt to provide a comparative analysis or overall synthesis. In contrast, others - the co-ordinators for the North Sea grouping being a good example - added real value by using their expertise to make up for deficiencies in the analysis provided by subcontractors, especially on more complex aspects of the methodology such as the calculation of net impacts, and in pulling together a coherent overall synthesis.

In light of the key role of co-ordinators, more could perhaps have been done to define their responsibilities at the outset of the study. Looking back at the co-ordinators’ inception reports, and the original terms of reference, this was in many respects a neglected aspect of the overall evaluation plan.

**Role of the Commission and Member States**

The Commission played a very important role in guiding the research and using its influence to secure the collaboration of regional and national authorities. Helpful contributions were also made with the interpretation of research findings and in feeding in information from other sources such as Eurostat.

Less positively, the merits of organising the 60 Objective 2 regions into multi-national groupings for the purposes of the research is doubtful. In theory, this approach might have been expected to have a number of benefits - facilitating cross-country comparisons, the transfer of evaluation know-how from regions with strong evaluation expertise to weaker ones, etc. In practice, however, there is little evidence of this having been the case and the complications almost certainly outweighed the advantages. This is especially so since the trans-national regional groupings were ultimately abandoned as a framework for the analysis with the final report presenting results on the more familiar country-by-country basis. The arrangement of having the regional co-ordinators appointed by and responsible to the Commission, rather than to Ernst & Young, also complicated the overall co-ordination.

As far as the role of Member States is concerned, many of the complications encountered in the evaluation - especially the lack of good-quality monitoring data - stemmed from weak programme management practices although as observed in the evaluation itself, steps were being (and continue to be) taken to rectify this situation. The experience of carrying out the Objective 2 evaluation also underlined the importance of interim evaluation work, which is of course the responsibility of Member States. Interim assessments of the 1989-93 Structural Fund programmes - some 19 in total - were undertaken in Denmark, Germany, Italy, The Netherlands and the UK and it was by and large in these countries that the ex-post evaluation work was carried out most successfully. Surprisingly, none of the evaluators selected by national authorities to carry out the interim assessments appear to have been used for the ex-post evaluation. Involving them in the Objective 2 study would have probably improved the results given their familiarity with the situation in the various regions.

Finally, although the Commission has responsibility for ex-post evaluations, the Structural Fund regulations emphasise the need for a partnership approach and yet the direct involvement of Member States in the study was limited. In some countries, where Objective 2 assistance
accounted for a very small proportion of the overall regional development effort, this lack of involvement is perhaps understandable. Greater Member State ‘ownership’ of the evaluation exercise would, however, have been helpful and in retrospect more could perhaps have been done at the outset of the study to secure their buy-in, perhaps through more extensive consultation over the terms of reference and appointment of evaluators.

Other Factors and Future Priorities

There were a number of factors that influenced the way in which the Objective 2 evaluation was carried out, and the results achieved, which were largely beyond the control of the evaluators, Commission and Member States.

These factors included the very tight times-scales for the study - about nine months for the main research - which was dictated by the need for results to be available in time to inform the negotiations with Member States over the new 1997-99 Objective 2 programmes. The complication noted earlier of often not being able disaggregate statistical information on regional trends down to the required level clearly applies more to Objective 2 regions than those covered by Objective 1 which are much larger and in some cases cover whole countries. Similarly, the fact that Objective 2 programmes generally represent a smaller proportion of expenditure on regional development than is the case in Objective 1 areas made the task of attribution more complicated.

Nothing much could have been done about these factors but the Objective 2 study does point to other ways of improving the way in which Structural Fund evaluation work is carried out and used. From a theoretical point of view, the ex-post evaluation model described at the outset of this paper has considerable merits. However, to improve practical implementation, more attention needs to be given to developing synthesis methodologies, with regard to both the horizontal (cross-regional) and the vertical (top-down/bottom-up) dimensions of the model. At present, the concept of synthesis evaluation seems to be relatively under-developed in the thinking on how Structural Funds programmes should be evaluated.

The need to develop synthesis methods also applies in a broader sense. Ideally, the results of the Objective 2 study would have been compared with evaluation findings for earlier periods and with other evaluations such of those for the Objective 1 programmes and the Community Initiatives. Here, the challenge is one of devising methods for bringing about a synthesis or cumulation of evaluation findings where the source material is often dispersed and there are diverse underlying methodologies. The Commission’s recent publication ‘Learning through Best Practice’ is a step in the right direction. Experience elsewhere is also interesting in this respect, the Evaluation Knowledge System (EKSYST) being developed by the International Fund for Agricultural Development (IFAD) being one example of how this problem is being tackled. This system, which is managed by the IFAD’s Office of Evaluation (OE), is a computerised database containing the best practice lessons and other findings from all the evaluation studies carried out by the Fund. It can be searched by theme, region or project type.

Priority clearly needs to be given to improving evaluation methods if the results are to have any credibility and be used by policy makers. Beyond this, however, other steps might be considered. Again, drawing on experience elsewhere, perhaps we need to consider including a more explicit requirement in the terms of reference for Commission studies that evaluators specify precisely how their findings might be used by programme managers and policy makers. A possible model in
this respect is provided by the American Evaluation Association (AEA) which has developed a set of ‘utility standards’ - constructive orientation, clear definition of users, functional reporting, follow up and impact, etc - as part of its attempt to develop a professional code. Likewise, more rigorous and systematic monitoring of the utilisation of evaluation research might also be considered. Here, there could be some benefit from examining the format adopted by the United Nations which has a compliance officer in its Office for Internal Oversight Services (OIOS) whose function is to monitor the implementation of recommendations made in evaluation reports. This is done on an annual basis, the results - or ‘implementation rates’ - being reported directly to the Secretary General.

Jack Malan is responsible at Ernst & Young for evaluation work. He has worked extensively in this field for regional authorities, UK Government Departments and the European Commission. He was the co-author of the Ernst & Young synthesis report for the European Commission on the ex-post evaluation of the 1989-93 Objective 2 programmes and is currently leading an evaluation of Structural Fund impacts on SMEs.