EUROPEAN COMMISSION
DIRECTORATE-GENERAL
REGIONAL POLICY

A Report on

The Performance Reserve and Mid Term Evaluation

in Objective 1 and 2 Regions

27 July 2004

Table of Contents

Part 1

1. A First Analysis of the Implementation of the Performance Reserve and Mid Term Evaluation in Objective 1 and 2 Regions

The reform of the Structural Funds in 1999 introduced a stronger emphasis on quality and performance through the development of a number of new and mutually re-inforcing instruments. The regulatory deadlines and the link with the availability of additional financial allocations if those deadlines were met and if performance was achieved were new elements in the Structural Funds delivery system. While the N+2 rule introduced an incentive to ensure that programmes advanced to schedule, it was complemented by initiatives which aimed to ensure that those resources were spent as effectively as possible. This report concerns these latter initiatives, specifically the performance reserve and the mid term evaluation.

The performance reserve and the greater emphasis on evaluation in general are a reflection of a growing focus internationally on results-oriented management. The intention is to build on the experience gained and to further strengthen the focus on quality and performance in the management and implementation of the Structural Funds in the future. It is appropriate therefore at this stage to consider how both instruments performed and to draw lessons for the future. Part One of this report draws conclusions on the performance reserve and mid term evaluation, based on the technical analysis which is contained in Part Two. The report is based on analysis of the experience in Objective 1 and 2 regions.

1.1. The Performance Reserve in Practice

The performance reserve was implemented in partnership between the Member States and the Commission. The Commission recognises and welcomes the positive attitude on the part of the Member States to this new approach which for the first time linked performance to financial allocations. In working together to identify successful programmes and to agree financial allocations which reflect performance, the Commission believes that the partnership between it and the Member States has been strengthened.

The major strength of the performance reserve was its ability to act as **an incentive for capacity building** in good management practices. The fact that targets and deadlines were involved meant that regions and Member States took action to ensure that:

- Resources were spent;
- Evaluations were produced on time;
- Monitoring systems were put in place;
- Financial control systems were implemented; and
- Project selection was increasingly transparent.

Three types of criteria were involved in judging performance: financial, management and effectiveness criteria. In relation to financial absorption, the N+2 rule provided the greater impetus to increased expenditure, re-inforcing the beneficial effect of the mid-term review and performance reserve allocation.

In relation to management criteria, the performance reserve stimulated the more rapid establishment of sound management systems for project selection, financial control, monitoring and evaluation. In some cases these practices already existed and the performance reserve ensured they were applied in good time. For other programmes and regions – especially in Objective 1 – the performance reserve encouraged the introduction or more rigorous application of these practices. This is a particularly important conclusion in the context of an enlarged EU and the need for capacity building in management and administrative capacity in the new Member States.

As regards effectiveness indicators, linking financial resources to the achievement of targets for outputs and results concentrated the attention of managing authorities on the need to develop their capacity to appropriately identify indicators and set targets. However, future mechanisms linking performance and financial allocations need to take a more realistic view of the capacity for target setting in the current and future Member States, with perhaps more guidance and appropriate benchmarks proposed by the Commission.

A second conclusion is that there was **great diversity in the detail of the methods** used for assessing performance and making allocations. While this is to be expected for the first implementation of an innovative instrument, any future such instrument would benefit from more clarity in its definition at the outset.

In relation to the principles the Commission proposed for the allocation of the reserve, the following can be concluded:

- **Transparency** was enhanced where there was clarity from the outset on indicators to be used, the targets and how performance would be judged. Equally, clarity on the principles by which the resources would be allocated also assisted transparency. Future performance related allocations should ensure that indicators and mechanisms are agreed at an early stage in the programming process with more uniformity in approach if possible across Member States.

- Judgements were made on an **equitable basis** across programmes and objectives with the same standards being applied within Member States. Because Community legislation did not impose one system, the systems used varied across Member States, both in content and timing.

- The decisions on the scale of allocations of the reserve to programmes were **in line with performance** and, appropriately, took account of **absorption capacity**. Some programmes received larger amounts because of the priority accorded to their content. Programmes judged to be successful all received larger amounts from the reserve than the basic 4%.

- **Objectively measured performance** led to pro rata allocations being made in a number of countries. Where all programmes met their targets, the Commission agreed that they should receive a full allocation from the reserve. In such cases in the future, it may be necessary to benchmark targets to ensure that they are set at an appropriate level, based on international experience.

Further consideration will be give to these findings in the context of the development of cohesion policy for the 2007-2013 period. Such a mechanism has the potential to act as an incentive and to add value to the development of managerial and administrative capacity in the implementation of the Structural Funds.

1.2. The Mid Term Evaluation process and results

The mid term evaluation was carried out in all cases and **two thirds of reports were of good or excellent quality.** This is a significant improvement on previous performance and reflects growing evaluation capacity in the Member States. It also reflects the linkage with the performance reserve, which created the financial reward for completing the evaluation on time and to an acceptable quality standard.

Feedback from the Member States suggests that managing authorities and monitoring committees found the process useful, although there were some difficulties. The fixed deadline and the link with the performance reserve ensured that the evaluation was done and in time to influence decisions on programme allocations for the rest of the programming period. However, in some cases where programmes were agreed late or there was a slow start up, there was very little activity to evaluate. This was particularly the case for INTERREG III and URBAN II and for some of the Objective 2 programmes (e.g., Luxembourg Objective 2, Italy Objective 2).

The Commission has already begun to consider how these issues could be addressed in the future. The Third Report on Economic and Social Cohesion suggests that the mid term evaluation should be replaced by a process of ongoing evaluation. It suggests that more consideration needs to be given to designing programmes of evaluation which are adapted to regional and national needs. This should be done in a way which ensures that developing evaluation capacity and the systematic use of the evaluations are further strengthened.

1.2.1. Next Steps – the Mid Term Review and Mid Term Evaluation Update

The ultimate purpose of the mid term evaluation was to provide a critical assessment of the progress of the programmes and suggestions for how to improve them and maximise their impact. The consideration of evaluation findings forms part of the mid term review which is underway in all Member States. Through the mid term review, evaluation recommendations are considered and decisions are made on changes to make to programmes. The mid term review decision will also allocate the performance reserve amounts by programme to the relevant priorities.

Article 42(4) of Regulation (EC) No 1260.1999 requires that the mid term evaluation is updated by the end of 2005 in order to prepare for future interventions. The Commission has proposed a draft Working Paper on the Update of the Mid Term Evaluation which has been finalised following discussion at the Technical Group on Evaluation.

1.3. Final Comments

The performance reserve and the mid term evaluation have contributed to a results-oriented management of the Structural Funds. Management by results is a more complete approach than the exclusive focus on spend and it has a longer term perspective. The first steps have been taken in this regard. The long-term effect of this greater focus on results should be evident in the greater impact achieved by Structural Fund programmes in the future.

Part 2

2. TECHNICAL REPORT

2.1. Introduction

Part Two of this report comprises a technical analysis of the implementation of the performance reserve and mid term evaluation in Objective 1 and 2 regions. At a later date, the Commission will produce a more detailed report on the processes, methodologies used and content of the mid term evaluations. That report will form the basis for a seminar with the Member States which will seek to draw lessons for the future.

2.1.1. The Provisions of the Regulation

The performance reserve is governed by Article 44 of Regulation 1260/1999 (see below), which states that 4% of resources should be held back at the initial programming stage for allocation at mid term to performing programmes. The performance reserve does not apply to the Community Initiatives. In total, over €8 billion was allocated to Structural Fund programmes in March 2004 through the performance reserve mechanism, €6 billion under Objective 1, €1 billion each under Objectives 2 and 3 and €50 million under the Financial Instrument for Fisheries Guidance (FIFG) outside Objective 1.

Article 44 of Council Regulation 1260/1999 – The Performance Reserve

1. Each Member State, in close consultation with the Commission, shall assess under each objective and not later than 31 December 2003 the performance of each of their operational programmes or single programming documents on the basis of a limited number of monitoring indicators reflecting effectiveness, management and financial implementation and measuring the mid-term results in relation to their specific initial targets.

These indicators shall be decided by the Member State in close consultation with the Commission taking account of all or part of an indicative list of indicators proposed by the Commission and shall be quantified in the existing different annual implementation reports as well as the mid-term evaluation report. The Member States shall be responsible for their application.

2. At mid-term and not later than 31 March 2004, the Commission shall allocate, in close consultation with the Member States concerned, under each objective, on the basis of proposals from each Member State, taking account of its specific institutional features and their corresponding programming, the commitment appropriations referred to in Article 7(5) to the operational programmes or single programming documents or their priorities which are considered to be successful. The operational programmes or single programming documents shall be adapted in accordance with Articles 14 and 15.

It should be noted that the "specific institutional features" of Article 44(2) refer to the federal regions of Germany, Austria and Belgium and the devolved administrative regions of the UK. It was not proposed to allocate resources across these regions in these Member States.

The mid term evaluation is required by Article 42 of the Regulation. The linkage between the performance reserve and the mid term evaluation lies in the fact that the evaluation should present the results for the performance reserve indicators and also that the Commission has a role in assessing the quality and relevance of the mid term evaluation in the context of making its decision on the allocation of the reserve. In most cases, the quality of the mid term evaluation was one of the criteria for the allocation of the reserve.

Article 42 of Council Regulation 1260/1999 – The Mid Term Evaluation

1. Mid-term evaluation shall examine, in the light of the ex-ante evaluation, the initial results of the assistance, their relevance the extent to which the targets have been attained. It shall also assess the use made of financial resources and the operation of monitoring and implementation.

2. Mid-term evaluation shall be carried out under the responsibility of the managing authority, in cooperation with the Commission and the Member State. It shall cover each Community support framework and each assistance. It shall be carried out by an independent assessor, be submitted to the Monitoring Committee for the Community support framework or assistance concerned in accordance with Article 35(3), and then sent to the Commission, as a general rule three years after adoption of the Community support framework or assistance, and no later than 31 December 2003, with a view to the revision referred to in Article 14(2).

3. The Commission shall examine the relevance and quality of the evaluation on the basis of criteria defined beforehand by the Commission and the Member State in partnership, with a view to reviewing the assistance and allocating the reserve referred to in Article 44.

4. As a continuation of mid-term evaluation, it shall be updated for each Community support framework and assistance and completed no later than 31 December 2005 in order to prepare for subsequent assistance operations.

2.1.2. The Commission's Approach

In relation to the performance reserve, the Commission produced a working paper in 1999 which described the instrument in more detail and proposed management and financial indicators to be used[1]. In October 2003, it outlined guiding principles for the allocation of the reserve in a note to the Committee for the Development and Conversion of the Regions. In that note, the Commission proposed a number of principles which should govern the allocation of the reserve, as follows:

- Transparency in decisions on the allocation;
- Equity in treating programmes and objectives;
- An alignment between objectively measured performance and the scale of allocations;
- Avoiding pro rata allocations unless clearly justified;
- Rewarding programmes which are more successful; and
- Taking account of absorption capacity in making allocations.

The Commission's approach to the mid term evaluation was outlined in a working paper[2] which provided guidance on the evaluation which was discussed with the Member States during 2000 and finalised at the end of that year.

[1] Working Paper No. 4 on the Implementation of the Performance Reserve
[2] Working Paper No. 8 on the Mid Term Evaluation of Structural Interventions

2.1.3. Scale of the Task

The mid term evaluations were submitted to the Commission by the end of December 2003. In total, there were 120 evaluations of Objective 1 programmes, nearly 100 for Objective 2 and 120 for the Community Initiatives, INTERREG III and URBAN II. As noted above, the performance reserve did not apply to the Community Initiatives. In total 59 proposals for the allocation of the performance reserve in Objective 1 and 2 regions were received towards the end of 2003 and in early 2004. All evaluations were assessed and proposals for the performance reserve were examined and discussed with the Member States to allow a decision to be taken on the allocation of the reserve by the deadline of 31 March 2004.

Because of the tight timescale involved, the Commission decided to have one global decision on the allocation of the performance reserve, covering all funds and all Objectives. This decision allocates the resources to programmes or priorities. The breakdown of the allocation between years and priorities (where global decision allocations were to programmes) and the inclusion of national co-financing is undertaken in the context of individual decisions on programme modifications which also give effect to the decisions of the mid term review.

2.2. The Performance Reserve

The following analysis relates to the experience of the performance reserve for Objectives 1 and 2.

2.2.1. The Performance Reserve Process in Practice

The performance reserve involved two stages, as follows:

- Assessment by each Member State in close consultation with the Commission by 31 December 2003 of the performance of each programme on the basis of a limited number of indicators and targets reflecting effectiveness, management and financial implementation (Article 44 (1)).
- By 31 March 2004 at the latest, allocation of the Performance Reserve to successful programmes or priorities by the Commission in close consultation with the Member State on the basis of a proposal by the Member State (Article 44 (2)).

While these two stages were concentrated in the period towards the end of 2003 and early 2004, an important element of the mechanism was the preparatory work in identifying the indicators, setting targets and developing a method for judging successful programmes and deciding how much each programme should receive. This work started with the drawing up of Operational Programmes and Single Programming Documents in 2000 and 2001. The programmes included management and financial indicators and targets and effectiveness indicators, which were then quantified in the Programme Complement. Further work on developing the mechanism and defining it in each national or regional context continued throughout the period up to the submission of formal proposals by the Member States. The Commission worked closely with Member States throughout this exercise.

The Identification of Performance Indicators

Three types of indicators were proposed in Regulation (EC) No 1260/1999 to allow the performance of programmes to be assessed. Management and financial indicators were relatively straightforward and the same indicators were used in nearly all Member States with similar targets. Effectiveness indicators were more complex.

In most cases Member States used the list of management indicators recommended by the Commission in Working Paper No. 4, i.e., the quality of the monitoring system, financial control, project selection and evaluation. These indicators applied usually to the entire programme; therefore, they were more suited to those cases where decisions were being made on allocations by programme. The variations ranged from Denmark which used just one management indicator – financial control of projects - to Portugal with six, with the monitoring indicator broken down between physical and financial monitoring (for which the target was 80%, as compared to 100% for most Member States) and the financial control indicator broken down between a financial control system and financial control of projects. A further refinement was introduced in Italy where most indicators were obligatory, while project selection was not.

The Commission proposed two financial indicators – financial absorption and leverage, with the target for financial absorption depending on whether the programme was agreed in 2000 or 2001. The target level for payments as a percentage of commitments for the first two years was set at 100% in nearly all countries. The performance threshold was 75% in most countries, but was 100% in Austria, Belgium, Italy and UK–England.

Financial leverage was defined differently across Member States (e.g., private funding as a percentage of the total financing plan; or total private funding realised; or the percentage of planned private funding achieved with the plan representing 100%). The percentage targets are therefore not comparable across countries. Leverage was not used as an indicator in Belgium, Germany–transport, Finland, Italy, Spain and Sweden.

As noted above, effectiveness indicators were more complex. The Commission's proposal was that they should comprise a subset of the output and result indicators of the programme, representing at least 50% of expenditure. The following summarises the result of analysis of the effectiveness indicators used across Member States:

- All countries used both output and result indicators, except Greece, Italy Objective 2 and Denmark, which used only outputs, and Sweden, which used only results.

- Most Member States deemed that a target would be reached if 75% of the absolute figure for the mid term target identified in the programme complement was achieved. 100% was to be achieved in Denmark, some German regions, Italy and UK–England. 80% was proposed for Spain while 60% was set in Portugal.

- There was a significant variation in the number of indicators used for effectiveness, varying from two output indicators in Denmark to 28 indicators (14 output and 14 result) in UK Eastern Scotland Objective 2. Originally there were more indicators in some Member States (up to 100 for one programme) and sometimes unrealistic targets were set. As part of the simplification exercise[3], the Commission proposed in 2002 that effectiveness indicators for the performance reserve should be streamlined and simplified. This was completed in 2002 for the UK while the exercise continued to 2003 in Greece, Portugal and Spain.

In general, effectiveness indicators were less robust than the management and financial indicators for a number of reasons. Some targets were set very low (with mid term achievement sometimes several hundred per cent of target); others were set too high. Clearly, the reliability of the effectiveness indicators was particularly important for decisions on allocations within programmes, where management and financial criteria were less relevant. While there were weaknesses in the effectiveness indicators in some Member States, much useful learning has taken place in relation to the issue of identifying indicators and setting appropriate targets. This needs to be built upon on the future, in further developing monitoring systems.

The Definition of Methods to Transparently Identify Performance

It is evident from the above analysis that methods to identify performance varied across Member States. We can identify two main approaches:

- Six Member States set out their methods for identifying performance and allocating the reserve in their programming documents (Belgium, Denmark, Finland, Luxembourg and Sweden) or at the beginning of the programming process (Italy). Thus the method was defined at the start of the process.

- The remaining Member States developed methods through proposals and bilateral discussions with the Commission during 2002 and 2003. Ireland and the Netherlands developed their methods earlier – during 2002, while Austria, France, Germany, Greece, Portugal, Spain and the UK undertook the process during 2003. Although the method to be used in France was not agreed in advance, it was discussed during 2003. The UK incorporated their method in their proposal which was discussed in draft form with the Commission. The Commission was concerned for reasons of transparency when methods were discussed very late in 2003 and only very shortly before performance was to be assessed.

Even where methods were identified early, the actual process of identifying programmes with higher performance sometimes raised further questions which needed clarification after the proposals were received. The Commission sought clarifications on the methods used before agreeing to Member State proposals.

An example of good practice is demonstrated by Italy which followed the advice of the Commission in Working Paper No. 4 to appoint a group of experts in an advisory role to assist in assessing performance. The working group for Italy Objective 1 was chaired by the Italian Evaluation Unit and included representatives from the Evaluation Unit, regional evaluation units and two independent experts nominated by the Commission. The working group produced an in-depth report on performance of

[3] Commission's proposals to the Ministerial Meeting on 7 October 2002 on the Simplification, Clarification, Co-ordination and Flexibility of Structural Policy Management in the period 2000-2006

each programme against the agreed criteria. A working group was also established for Objective 2.

Methods to Allocate Resources

Having identified performing programmes, the next stage was the allocation of the performance reserve in a manner which reflected the performance demonstrated. The methods proposed varied between Member States, with some again deciding not to propose a method in advance of making their proposals. Where Member States did propose and agree methods with the Commission in advance, these varied in clarity and detail. Some methods which were proposed in advance were not entirely reflected in practice when the proposals were made.

The analysis reveals the following variations in allocation mechanisms:

- Where all programmes clearly performed, they all received a full allocation from the performance reserve. There was therefore no need for a method for the allocation of additional resources.

- A further method was to first eliminate non-performing programmes and then distribute the available resources to the highest performing programmes. This method was used in Greece and in the Austrian regions to identify the performing priorities. In Spain the allocation was pro rata across all programmes, having excluded three. In Portugal, having excluded three programmes, the remaining programmes were divided into two groups, according to performance.[4]

- Some Member States proposed an allocation to all programmes with amounts less than 4% to those programmes which only partially achieved targets, with the highest allocations to those which had highest performance. This method was used in Italy Objective 1, UK–England and France[5].

- The method used in the UK–England introduced a cap of 5% because of absorption concerns.

- Weighting was used in some German regions (e.g., Brandenburg) and in the UK – Scotland Objective 2. The Scottish proposal was a good example of a rigorous and transparent system. It introduced weighting for effectiveness indicators and capped the actual performance ratio at 200% of the target.

- Finally, in Ireland and Portugal higher amounts were allocated to performing programmes with a high priority.

2.2.2. Reaching Agreement with the Member States

All proposals were received by early January 2004. The Commission engaged in speedy discussions with the Member States to seek clarifications on the methods used

[4] In France, Greece, Portugal and Spain, the technical assistance programmes were excluded from the performance reserve exercise.
[5] France had a slightly different approach in that it excluded the two technical assistance programmes from the performance reserve.

to assess performance and to decide on the proposed allocations. The regulatory deadlines reduced the time available for detailed consideration and discussion on proposals. Most proposals were received in early January 2004 and decisions in principle on the amounts to be allocated had to be concluded by early February.

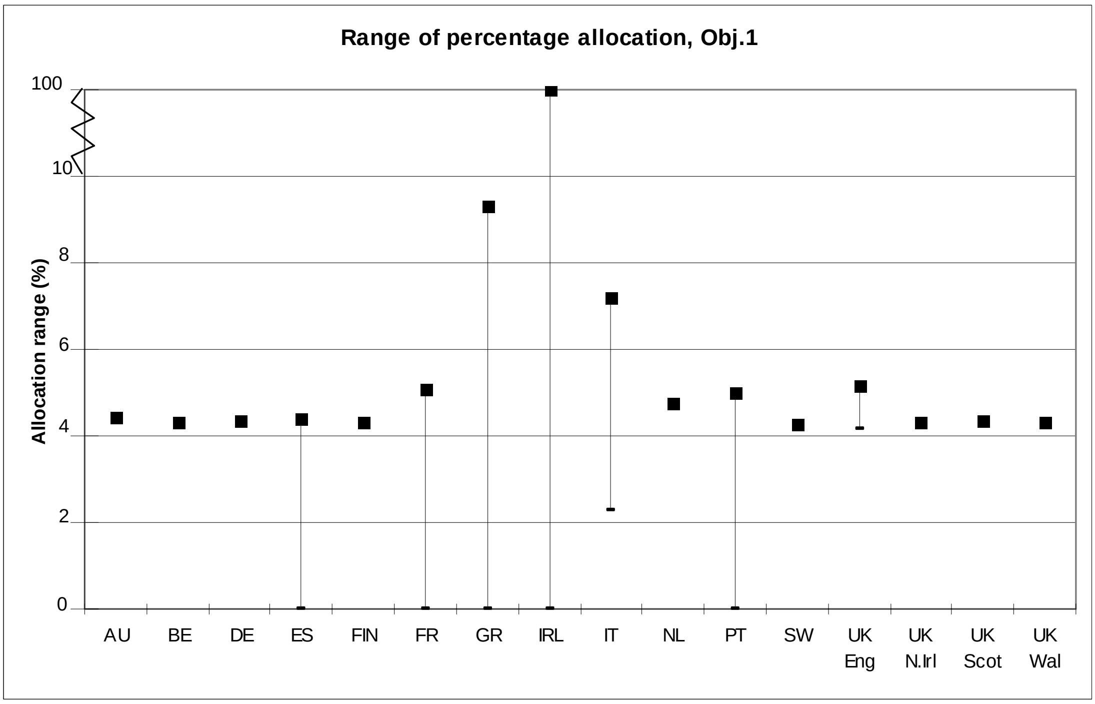
We can identify two groups of countries or regions:

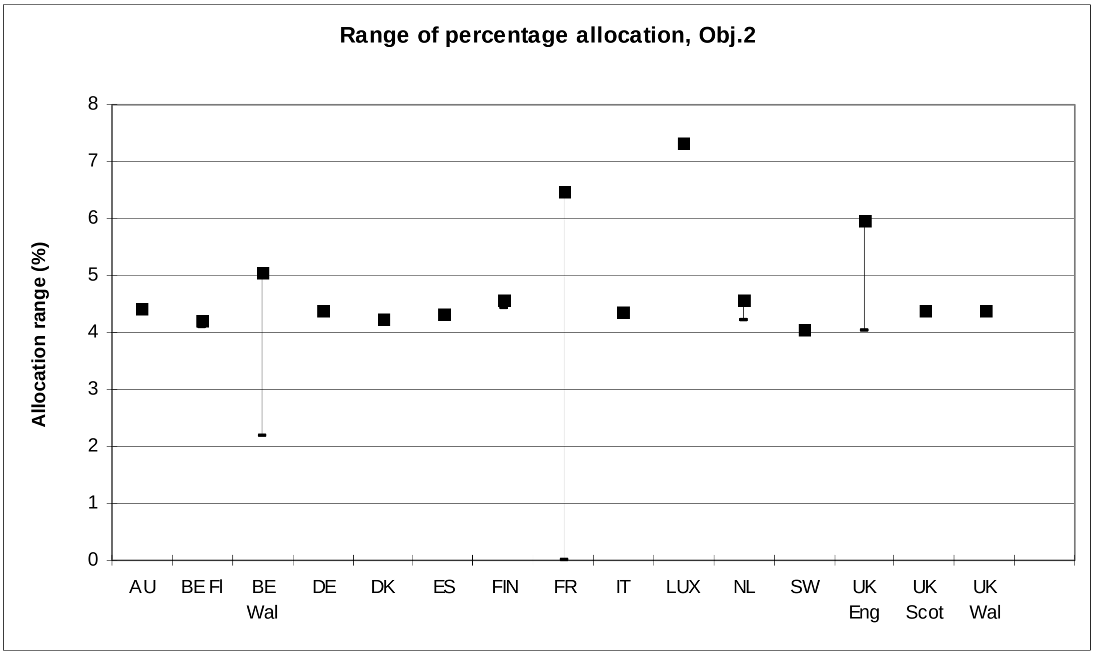
- In the majority of cases, the proposals made by the Member States could be immediately accepted by the Commission. These proposals transparently identified performance and the proposals for allocation were logically aligned with it. This was the situation for Austria, Denmark, some German regions, Greece, Finland, Ireland, Italy Objective 2 and Objective 1 in transition, Luxembourg, the Netherlands, Spain, Sweden and UK–Northern Ireland, Scotland and Wales Objective 2. A special case was France where a draft proposal was received and discussed with the Commission and some changes were incorporated into a final proposal which was accepted by the Commission.

- In other cases, the Member State proposal was accepted after clarification of the methodology used or some change in the initial proposal. The proposals for Belgium (except Flanders Objective 2 regions), several German regions and UK–Wales Objective 1 could not be accepted until the methodology was clear. For Italy Objective 1, Portugal and UK–England clarifications were sought and discussions with the Member States led to changes in the allocations.

2.2.3. The Result – the Allocation

The figures below summarise the range of allocations for Objectives 1 and 2 from the performance reserve based on the tables at Annex 1 which compare performance reserve allocations to the financial allocations agreed in the programming documents. It should be noted that for Germany, Austria, UK and Belgium, because of their institutional structures, the allocation of resources does not take place across federal or devolved administrative regions. Where there is only one programme under an Objective within such a region, the allocation was made to priorities within the programme. Equally, where there is only one programme under an Objective in a Member State, allocations are also by priority (Denmark, Luxembourg). Allocations in Member States within programmes are represented by a single point on the figures below. A single point also appears where there is a pro rata allocation across programmes within a country (Finland, Netherlands, Sweden).





A number of observations can be made on the outcome of the allocation of the reserve:

- Nearly 80% of Objective 1 programmes (93 in total) received allocations from the reserve; all Objective 2 programmes, except the technical assistance programme for France received allocations.

- For Objective 1 the range of allocations was greater than for Objective 2. For programmes which received an allocation, the greatest range was in Greece, where the amounts represent from 4% to 9.33% of total commitments with 11 programmes not receiving any allocation (including the technical assistance OP which was excluded in advance). In Italy, the allocations range from 2.3% to 7.2%, while in Ireland the entire allocation goes to one programme (infrastructure). Similar ranges are evident in France (3.2% to 5.12%) and Portugal (3.71% to 5.01%) with two technical assistance programmes in France not receiving an allocation, while three programmes were excluded in Portugal. Pro rata allocations were made in Finland and Sweden, where all programmes performed, while for Spain a pro rata allocation was made to the 20 programmes which received an allocation (two not receiving any allocation).

- Excluding technical assistance programmes from the analysis, in Greece, Portugal, Spain and Ireland some programmes received no allocation from the performance reserve, with the largest number of such programmes (10) in Greece. While the Italian method had proposed the exclusion of programmes which did not meet obligatory criteria, the proposal from the Italian authorities was for allocations to all programmes. This was agreed by the Commission on the basis that the Italian method had been more rigorous than those for some other Member States. The Commission accepted the argument that the performance reserve should operate more as an incentive than as a sanction and therefore partial performance was recognized and rewarded.

- For Objective 2, the allocations are more uniform with a less extensive range, partly due to a higher level of performance. The largest range is evident in France, with allocations ranging from 2% to 6.5%, while a significant range is also evident in England – from 4.02% to 5.96%.

- In the case of allocations between priorities there is a more extensive range. In many cases, allocations are concentrated on a limited number of priorities.

<u>Sectoral Allocation</u>

Where allocations were made to programmes, the sectoral breakdown of allocations is not available for regional programmes. However, the following table provides an analysis based on the information which is available. Clearly this table presents only a partial picture, given that 51% of the Objective 1 resources and 79% of the Objective 2 resources are allocated to regional programmes. The full picture will only be known when the mid-term review is complete and the allocations by priority within programmes are agreed.

Allocations of the Performance Reserve by Sector

	Objective 1		**Objective 2**	
Sector	**Amount €M**	**%**	**Amount €M**	**%**
Regional programmes	3,091	50.7	804	79.3
Infrastructure	866	14.2	75	7.4
Productive environment	1,249	20.5	131	13.0
Human resources	874	14.3	2	0.2
Technical assistance	19	0.3	1	0.1
Total	6,099	100	1,013	100

National Reserves

A final important point to mention is the experience of the national reserves, which went further than the requirements of the Commission. These reserves operated in Italy, Portugal and Greece. In Portugal and Greece the reserves amount to 2.6% and 2% respectively and they will be allocated to national strategic priorities in the context of the mid term review. In Italy, the national reserve amount was 6% of the original commitment and additional criteria, mostly related to management and administrative capacity, were identified, which provided further incentives for performance.

2.3. The Mid Term Evaluation

This section of the report is based on the quality assessments of the mid term evaluations and a preliminary analysis of the content of the evaluations. The analysis will be developed over the coming months to produce a report which draws lessons and identifies best practice in the mid term evaluation process. The report will form the basis for a seminar with the Member States later in 2004.

2.3.1. The Process

The Commission produced its guidance on the mid term evaluation in 2000 and consulted with the Member States on its content before finalising Working Paper No. 8 at the end of the year[6]. In the Working Paper the Commission defined its understanding of its role in the mid term evaluation process. The Commission asked to be consulted on the budget for the evaluation, the terms of reference, the inception report and the draft report. In fact, in most cases the Commission was represented on the Steering Groups and participated actively throughout the entire evaluation process.

Article 42(3) of Regulation (EC) No 1260/1999 requires the Commission to examine the quality and relevance of the mid term evaluations on the basis of criteria defined beforehand by the Commission and the Member State. The criteria used were those developed under the MEANS[7] programme of methodological guidance to evaluation of the Structural Funds. In order to contribute to the raising of the quality of the evaluations, the Commission decided to undertake quality assessments of Objective 1 and 2 reports at draft stage where they were available and to feed back views to managing authorities in time for evaluators to be able to improve the quality of the final reports.

Partnership Working

Partnership worked at several different levels during the mid term evaluation process and in different ways depending on national and regional practice, budgets involved and the scale of the programmes. In most cases steering groups were established which comprised members of the Monitoring Committee. The role of the steering groups varied across programmes and countries and they were not always involved

[6] *Guidelines for Systems and Monitoring of ESF Assistance in the Period 2000-2006* (1999) provided guidance specific to ESF interventions.

[7] European Commission, (1999). *Evaluating Socio-Economic Programmes. Volumes 1-6.* Luxembourg: Office for Official Publications of the European Communities. (The MEANS collection has been updated and a new GUIDE to the Evaluation of Socio-Economic Development can now be accessed through the website of the Directorate General for Regional Policy or directly via the website www.evalsed.info).

from the beginning of the process. In general, the role of the steering groups resulted in enhanced ownership and involvement in the evaluation by the partnership, though in some cases there seemed to be some attempt to steer the evaluation too much towards particular conclusions and recommendations. The presentation of the report to the monitoring committee broadened the involvement of the partnership in the process. In some cases, there was keen media and political interest in the findings of the evaluation.

The evaluation process was also characterised by partnership with the European Commission. The Commission provided an overall framework for the evaluation in consultation with the Member States. It participated in the steering groups and focused particular efforts on the terms of reference and the inception and draft reports.

As the Commission pointed out in its Working Paper, the mid term evaluation process was a complex exercise. In general, organisation of the mid term evaluation contributed to higher quality. Core terms of reference were drawn up in a number of countries (e.g., Austria, France, Italy, UK) which assisted in ensuring the evaluations were well designed. In France and Italy a particularly close support to the evaluation process was provided by DATAR and the Italian Evaluation Unit respectively, with quality reviews of evaluation drafts being undertaken in partnership with the Commission.

The Commission underlined in its guidance on the mid term evaluation that sufficient time should be devoted to the exercise. In most cases, planning started appropriately early and evaluators were in place in time to allow them to produce high quality work. Delays occurred in Greece and Luxembourg (Objective 2). In Luxembourg this was due to the late start and small scale of the programme. In Greece intensive work once the evaluators were in place ensured that any harmful impact on the quality of the work was minimised.

<u>The Quality of the Mid Term Evaluations</u>

As noted above, the Commission examined the Objective 1 and 2 mid term evaluations at draft and final report stage. At draft report stage, suggestions were made as to how reports could be improved. The following table demonstrates the very considerable improvement made in these evaluations between draft and final report stage.

Quality Assessments – Objective 1 and 2 Reports		
	Draft Reports %	**Final Reports %**
Good/Excellent	34	65
Acceptable	48	35
Unacceptable	18	0
Total	100	100

The final outcome, with nearly two thirds of the evaluation reports assessed as being good or excellent, is extremely positive. At draft stage, the relatively high number of unacceptable reports was due to many reports being unfinished. It is clear that the partnership between the Commission, the managing authorities and the steering groups contributed significantly to improve the quality of the reports.

The Commission is currently gathering information on the cost of the evaluation and will provide analysis on its cost effectiveness in the more in-depth report it is preparing. Significant resources were allocated to the exercise, which contributed to its higher quality compared to the past. Most large programme evaluations cost between €150,000 and €250,000, but higher amounts were used in general in Italy and Germany.

The Evaluators

In total approximately 170 evaluation companies were involved in the evaluation of Objective 1 and 2 programmes. In many cases consortia of evaluators were established. This is particularly the case for Germany, which had the largest number of evaluators (28) involved in the mid term evaluation process, compared to the number of evaluations (19). By contrast, in Sweden one evaluation company undertook all 6 evaluations while another company carried out two evaluations concerning management issues for European Social Fund measures under Objective 1. The largest number of evaluations undertaken by one evaluation company was 10 in Spain, in several cases in consortium with other companies. In Portugal one evaluation company was involved in 9 evaluations while another was involved in 7, but in both cases the companies were in consortium with each other and/or other companies. In France, one company undertook 6 evaluations. For Germany, Belgium, Greece, Italy and the UK, the maximum number of evaluations undertaken by one company was 4. It was less than this in Austria, Denmark, Finland, Ireland, Luxembourg and the Netherlands.

In terms of involvement in evaluations across national boundaries, there is some limited evidence of transnational involvement with evaluators from Ireland, Austria, UK and Denmark involved in the German evaluations, Swiss evaluators undertaking one Austrian evaluation, a Belgian evaluator undertaking the Luxembourg evaluation and a Spanish evaluator involved in one Portuguese evaluation.

An innovative approach to the mid-term evaluation based on a transnational exchange of experience was demonstrated in the Objective 2 regions of Nordrhein-Westfalen, East of Scotland and West of Scotland. Building on long term cooperation links, the regional authorities decided to launch a comparative study on the mainstreaming of horizontal issues in these Objective 2 programmes. The objective was to stimulate debate on experience of implementation of the Structural Funds and to strengthen the learning process. The results of the study were discussed in workshop of representatives of the three regions.

For URBAN II, approximately half of the programmes were evaluated individually, while the evaluation of others was organized on a national level, as in Italy, Spain, Greece and England.

Overall, the mid term evaluation process demonstrates the increasing size of the marketplace of evaluators. However, there is evidence that there was pressure on that market with the 31 December 2003 deadline for all mid term evaluations while at the same time some evaluators were also working on ex ante evaluations for the new

Member States. Contacts with the Member States during the process suggest that those responsible for managing the evaluations had to insist on obtaining quality work as scheduled, particularly from evaluators engaged in a large number of evaluations, some of whom appeared to be over-stretched. This underlines the importance of evaluation capacity in the management of evaluation being available in the managing authorities to ensure the quality of the process.

Methodologies

In general, the evaluations adopted a mix of desk research, including analysis of monitoring data, and some fieldwork. Fieldwork involved key stakeholder consultations, surveys of beneficiaries in some cases and in others case studies or focus groups. For larger Objective 1 programmes or Community Support Frameworks (CSFs), some macro-economic modelling was carried out.

Some excellent evaluations had a mix of methodologies specifically designed to answer the evaluation questions and to build on the available information. Examples of such evaluations include UK–Merseyside, France–Guyane, Italy–Toscana. For example, the Merseyside evaluation, in addition to desk research and key stakeholder consultations, was based on 128 project reviews explicitly designed to provide a "reality check" on the views of implementing bodies and two beneficiary surveys targeted at businesses and individuals.

In general, the examination of implementation systems provides a useful input to improve management of Structural Fund programmes. This element of the evaluation was generally based on desk research and interviews with those responsible for implementing the programmes. It took on greater importance than foreseen in the Commission's working paper because of the late or slow start-up of many programmes, leading to limited information on outputs and results on which to base judgements on the performance of programmes. This focus of the evaluations on implementation systems was particularly notable in Italy and France and the INTERREG and URBAN programmes.

Methodological concerns arose for the following reasons:

- In some cases, there was limited primary research which was focused only on those responsible for implementing the programme. An evaluation of a Structural Fund programme should incorporate the "on-the-ground" reality and use this to explain the findings of desk research and results of interviews with managing authorities and implementing bodies. Even where programmes were slow or late in starting up, once projects are selected there should be information available on the perceptions of the beneficiaries.

- In many cases monitoring data was weak. Some programmes were late starting and had only limited data gathered on activity, while for others the monitoring system is not yet functioning and providing information on physical performance to the monitoring committee and to the evaluators for this exercise. This created problems for evaluators in reporting on what has actually been achieved to date in the programmes.

- Linked to the unavailability of monitoring data but also to the strong focus on financial performance due to the demands of N+2, some evaluators over-emphasised financial analysis with a limited focus on what was actually being achieved by the programmes. This led to recommendations which were purely finance driven, rather than reviewing the strategic implications of such shifts of allocations. A further conceptual concern was that many evaluators defined "efficiency" as the absorption of funds, rather than the more appropriate definition as the unit cost of achieving outputs and results.

- Evaluations often did not produce sufficiently operational conclusions and recommendations and some did not clearly differentiate between the two. Conclusions should be specific and clearly based on evidence presented while recommendations should be practical and actionable. Vague recommendations were a feature of some evaluations.

A final point relates to the length of evaluations and in some cases their lack of analytical content. Many of the evaluations were too long and their length militated against the dissemination of their findings. There is a need for evaluators to develop a capacity for more synthetic judgement and to focus more on analysis rather than including very long descriptions.

2.3.2. Brief Summary of Mid Term Evaluation Findings

The findings of the mid term evaluation will be synthesised in more detail in the next few months and the Commission will draw conclusions in relation to the major themes of the Lisbon Strategy. The following summarises the preliminary findings.

<u>Need for Strategy Shifts</u>

In nearly all cases, the evaluators concluded that the strategies were still appropriate. The global downturn in the world economy was analysed across the different programmes, but the conclusion was that the programmes were still appropriate to meet the development challenges. Some adjustments or refinements to strategies were recommended, however:

- The German evaluations emphasised the need for greater concentration on important priorities with decisions on actions to be funded to be made on the basis of the potential for a contribution to long term growth. All Objective 1 programmes will need to be fine-tuned in 2004 to reflect new labour market legislation.

- The Irish evaluations examined the fact that allocations for enterprises and research in particular were not being spent and suggested a shift towards transport infrastructure, further strengthening the emphasis on this area.

- The Portuguese evaluations recommended a better articulation with the European Employment Strategy, a greater focus on the modernisation of the services sector and public administration, and the need for a stronger contribution to diffusing innovative public policies and developing programmes of training for SMEs.

- The Spanish evaluations recommended some modifications as a result of the "Prestige" catastrophe as well as the adaptation of programmes to focus more on research, development and innovation.

- The UK evaluations in general found strategies to be still sound. For Merseyside, refinements to the strategy were recommended including a new measure on innovation and technology transfer.

- For almost all the URBAN II programmes, the socio-economic circumstances had not changed significantly. Changes to the strategy were only recommended in seven out of seventy programmes, although over half of the evaluations recommended the shifting of some funds between priorities or measures.

Often the evaluations accepted the strategy as set down in the programming document and only considered change in the context of changing socio-economic circumstances. The evaluation was also supposed to give a view on whether or not the programme was still the correct response to the development needs of the region or sector concerned.

<u>Achievements to Date</u>

Achievements measured in terms of *financial* performance were presented in all cases, although recommendations were sometimes weak on actions to take in relation to under- or over-spends. Analysis of *physical* performance presented problems in many Member States. In a number of cases, the monitoring system provided little or only partial information on achievements. It is clear that a substantial amount of activity has already taken place under the 2000-2006 period, although the full impact of that activity will only be evident over time:

- In Germany in Nordrhein-Westfalen, 38% of the resources are committed and 14% are paid. 20,000 new jobs have been created, 20% of the target, while 46,000 jobs have been saved, well above the target of 29,000. In Rheinland-Pfalz, 2,066 new jobs had been created, 36% of the target and 385 jobs had been saved. In R&D and business start-ups, 70 new jobs had been created and 213 jobs saved.

- In Spain the CSF evaluation estimates that the rate of annual growth was 0.41% higher in 2000 to 2002 as a result of CSF expenditures. Improved accessibility has been achieved through 476 km of roads/motorways built or improved, 173 km of rail built or improved. In addition, 810 km of energy distribution networks have been constructed and 250 km of gas pipeline built. Support for the productive environment has resulted in 4,600 SMEs being assisted with a leverage effect on investment estimated at approximately €12.2 billion. A contribution to the development of the knowledge society is evident in 1,503 research centres and 48,199 researchers supported and the installation of 26,864 ADSL lines. Approximately 7 million beneficiaries have been supported by the ESF, of which 57% received training.

- In Portugal, 38,000 jobs have been created in SMEs. The gas distribution network has been extended by nearly 3,000 km and 25,000 consumers have converted to natural gas.

- In Eastern Finland, 15,900 jobs had been created, representing 90% of the target and 3,200 businesses had been created (105% of target). In Northern Finland, 6,700 new jobs had been created, 62% of the target, and over 900 new companies had started up (94% of target). For Objective 2, 8,300 new jobs and 1,000 companies were created in Southern Finland (55% and 20% of target respectively) while 16,000 jobs and 2,200 companies were created in Western Finland (46% and 22% of targets).

- In Sweden, the two Objective 1 programmes have created 1,390 jobs, supported the establishment of almost 5,000 SMEs and the training of over 18,000 people. The two Objective 2 programmes supported the creation of 2,400 jobs and 7,500 SMEs.

Quantification of Objectives

The quantification of objectives is an area where most evaluations found room for considerable improvement, although there were some very good examples of quantification of objectives feeding into an effective monitoring system (e.g., Germany - Nordrhein-Westfalen). In some cases, (e.g., Greece, France) baselines were missing which made evaluation work difficult. Recommendations to revise and simplify the indicator system were made in Belgium, France, Germany and Greece, while in Denmark some additional indicators were recommended. In Portugal the evaluators recommended further changes to improve the indicator system which had already been simplified during 2003. In many cases, inadequate definition of indicators was identified as a problem and recommendations were made to address this issue.

A further issue raised by the evaluations is where performance was far in excess of targets, suggesting that the target setting exercise was flawed at least in some areas (e.g., Finland, the Netherlands, Spain, Sweden). The explanation may be extremely good performance, but the evaluators do not always analyse the reasons for such over-performance. The fact that it is often accompanied by expenditure patterns in line with target suggests that in these cases targets were set too low. In Spain, the evaluators have recommended changes to targets to set them at more realistic levels.

In the UK, quantification of objectives was generally of a good quality since most programmes were based on a methodology developed during the last programming period[8]. This work provided a good starting point and the mid term evaluations have recommended further improvements to bring targets more in line with regional and programme specificities.

INTERREG presented particular challenges for the quantification of objectives and the Commission asked that the mid term evaluations should focus specifically on this issue. Evaluations demonstrate that the cross-border/transnational benefit in terms of outputs, results and impacts is combined with an institutional aspect, i.e., the degree involvement of partners from different countries. Some evaluations have suggested that apart from certain basic quality indicators, scope should be given to projects to propose project specific output and result indicators. As learning about co-operation processes is seen by many participants as a major result and added value of

[8] EKOS Limited, (1998): *Baselines and Quantification in the United Kingdom Objective 2 Programmes*

INTERREG self assessment by participants on progress in this field could represent a pathway towards better measurement of such qualitative results.

Given the small geographical areas covered by the URBAN II programmes which often cut across municipal boundaries, obtaining reliable and frequently updated baseline indicators is particularly difficult. The evaluations found in some cases that the quantification of objectives was inadequate and poorly structured and that data collection was poorly organised. Two thirds of the URBAN II evaluations included recommendations for improvements to the structure and definition of output, result and impact indicators. There were also examples of good practice highlighted in the evaluations. The URBAN II programme of the Grenoble agglomeration, which covers parts of five municipalities, commissioned a specially designed telephone survey to obtain good baseline indicators.

Implementation Systems

As noted above, the evaluations provided extensive material on the implementation systems and made many recommendations for changes to improve and streamline practice. However, in nearly all cases, the evaluations found that implementation systems had improved considerably compared to the past, with more rigorous project selection procedures, better financial control and evaluation and improving monitoring systems.

The UK evaluations demonstrate a strong partnership basis for programming and implementation, which is complemented by an appropriate degree of top-down co-ordination. The evaluations also highlight a high private sector involvement in the programmes which is also evident in the Netherlands.

In some cases (e.g., Greece) systems were found to be complex and recommendations were made to simplify project selection or other decision making procedures. In Portugal the mid term evaluations recommended a stronger co-ordination role in managing the CSF and a more strategic role in managing and monitoring the CSF and OPs. In Germany–Rheinland-Pfalz among other regions and Member States, recommendations were made for improved integration between the authorities responsible for different funds.

While the existence of an integrated monitoring system is identified as a positive element for Greece, the system for collecting and processing information is identified as a weakness although progress has been made compared to the past. The evaluations conclude that a continuing emphasis is needed on this area to bring about further improvements.

The evaluations generally found that INTERREG III demonstrates a significant advance in relation to the implementation of co-operation on the basis of truly joint structures. Such co-operation arrangements involve all parties concerned, select the projects and co-ordinate and monitor the implementation of programming. However, even though the joint implementing arrangements in most cases are found to work well, others experience problems. Several reports stress the need to have a legal instrument (for example in the form of a European co-operation structure), in order to help Member States, regions and local authorities to address the legal and administrative problems encountered in the management of transeuropean programmes and projects.

For URBAN II, just over half of the mid term evaluations recommend changes to the implementation system and in 18 programmes the modifications are considered to be crucial to the success of the programme. The recommendations cover such aspects as project selection procedures, support for project managers, staff and resources, communication and better involvement of local partners.

ANNEX

Performance Reserve
Range of Percentage Allocations, Objective 1

	Commitment 2000-2006, €M [1]	Performance Reserve, €M [2]	%[3]	Allocation Range (%)	Comment
Between Programmes					
Finland	948	41	4,32	4,32	Between the 2 programmes
France[4]	3.948	171	4,33	3,2 - 5,1	Between 6 of 8 programmes; 2 TA programmes excluded
Greece	21.389	945	4,42	4 – 9,33	Between 14 of 25 programmes; 11 programmes excluded, including 1 TA programme
Ireland	3.061	134	4,38	100	To 1 programme out of 6; 5 programmes excluded
Italy	21.638	996	4,60	2,3 - 7,2	Between 13 programmes; between priorities for 1 programme phasing out
Portugal	19.177	855	4,46	3,66 – 5,01	Between 16 out of 19 programmes; 3 programmes excluded, including 1 TA programme
Spain	39.548	1.717	4,34	4,41	Between 20 of 23 programmes; 3 programmes excluded, including 1 TA programme
Sweden	748	32	4,28	4,25 - 4,30	Between the 2 programmes
UK England	3.003	130	4,34	4,18 - 5,15	Between all 3 programmes
Between Priorities					
Austria	271	12	4,43		Between all priorities of the only programme (lowest 8%, highest 36% of total amount)
Belgium	645	28	4,34		Between all priorities of the only programme (lowest 7%, highest 35% of total amount)
Germany	20.602	899	4,36		Between priorities of 9 programmes : 2 programmes - 100% to one priority, 2 programmes - to all priorities, the remaining 5 programmes to the best performing priorities (lowest 2,3%, highest 93% of total amount).
Netherlands	126	6	4,76		Between 3 of 4 priorities of the only programme (lowest 18%, highest 46% of total amount)
UK N.Ireland	890	39	4,34		Between 3 of 5 priorities of the only programme (lowest 13%, highest 65% of total amount)
UK Scotland	306	13	4,37		Between 3 of 4 priorities of the only programme (lowest 17%, highest 64% of total amount)
UK Wales	1.853	81	4,34		Between 5 of 6 priorities of the only programme (lowest 7%, highest 34% of total amount)

[1] - 1999 prices; [2] - Current prices

[3] - The percentage is greater than 4% because the commitment figure is in 1999 prices and the performance reserve is in current prices and it differs between countries because DG BUDG made its calculation in millions and rounded figures up or down.

[4] In France, there were two proposals for the allocation of the performance reserve under Objective 1 – one for the 4 programmes in Objective 1 and one for the 2 programmes in Objective 1 phasing out.

Performance Reserve Range of Percentage Allocations, Objective 2					
	Commitment 2000-2006, €M [1]	**Performance Reserve €M** [2]	**%** [3]	**Allocation Range (%)**	**Comment**
Between Programmes					
Belgium Flanders	186	8	4,17	4,09 - 4,20	Between all 4 programmes
Belgium Wallonia	217	9	4,28	2,19 - 5,05	Between all priorities of the 2 programmes
Finland	507	23	4,54	4,41 - 4,58	Between all 3 programmes
France	6.262	273	4,36	2 – 6,5	Between 22 of 23 programmes; TA excluded
Italy	2.608	113	4,33	4,32 - 4,35	Between all 14 programmes
Netherlands	823	36	4,37	4,2 - 4,58	Between all 4 programmes
Spain	2.748	119	4,33	4,31 – 4,34	Between all 7 programmes
Sweden	423	17	4,02	3,99 - 4,05	Between all 4 programmes
UK England	3.774	170,9	4,52	4,02 - 5,96	Between all 9 programmes
UK Scotland	807	35	4,38	4,35 - 4,39	Between all 3 programmes
Between Priorities					
Austria	703	31	4,41	Between priorities of 8 programmes : 2 programmes-100% to 1 priority, 5 programmes - to all priorities (lowest 5%, highest 76% of total amount)	
Belgium Brussels	44	2	4,44	Between all priorities of the only programme (lowest 25%, highest 75% of total amount)	
Denmark	189	8	4,23	Between 3 of 4 priorities of the only programme (lowest 20%, highest 75% of total amount)	
Germany	3.626	159	4,38	Between priorities of 11 programmes : 3 programmes - 100% to one priority, 4 programmes - to all priorities, the remaining 4 programmes - to the best performing priorities (lowest 6%, highest 88% of total amount)	
Luxembourg	41	3	7,32	To 1 of 4 priorities of the only programme	
UK Gibraltar	8,3	0,36	4,34	Between priorities of the only programme	
UK Wales	121	5	4,39	Between priorities 2 of 3 of the only programme (allocation 28% and 72%)	
[1] - 1999 prices; [2] - Current prices					
[3] - The percentage is greater than 4% because the commitment figure is in 1999 prices and the performance reserve is in current prices and it differs between countries because DG BUDG made its calculation in millions and rounded figures up or down.					