



Économie rurale

Agricultures, alimentations, territoires

307 | septembre-octobre 2008

Nouvelles frontières entre les politiques rurales en Europe

Towards an Enhanced Evaluation of European Rural Development Policy Reflections on United Kingdom Experience

Vers une évaluation renforcée de la politique européenne de développement rural : quelques réflexions issues de l'expérience du Royaume-Uni

Janet Dwyer, Dylan Bradley and Berkeley Hill



Electronic version

URL: <http://journals.openedition.org/economierurale/421>

DOI: 10.4000/economierurale.421

ISSN: 2105-2581

Publisher

Société Française d'Économie Rurale (SFER)

Printed version

Date of publication: 15 December 2008

Number of pages: 53-79

ISSN: 0013-0559

Electronic reference

Janet Dwyer, Dylan Bradley et Berkeley Hill, « Towards an Enhanced Evaluation of European Rural Development Policy Reflections on United Kingdom Experience », *Économie rurale* [En ligne], 307 | septembre-octobre 2008, mis en ligne le 01 septembre 2010, consulté le 30 avril 2019. URL : <http://journals.openedition.org/economierurale/421> ; DOI : 10.4000/economierurale.421

Towards an Enhanced Evaluation of European Rural Development Policy Reflections on United Kingdom Experience

Janet DWYER • Reader in Rural Studies, University of Gloucestershire

Dylan BRADLEY • Senior Consultant, Agra CEAS Consulting Ltd

Berkeley HILL • Emeritus Professor of Policy Analysis, University of London

Introduction

The role played by policy evaluation has expanded greatly in the previous two decades, at both European Union (EU) and United Kingdom (UK) levels. Indeed, the UK can be held responsible for spurring its development in domestic policy with a particular emphasis upon “value for money” (for example, as expressed by Michael Heseltine at the Department of Trade and Industry, in the mid-1980s). In the EU, commitment to evaluation was an integral part of the expansion of the Structural Funds in 1988 (Hill, Young and Brookes, 1989). Over time the legal requirement to carry out assessments of the performance of public-funded policies has gained weight. Consequently both the Rural Development Regulation (RDR)¹ covering Programmes for 2000-2006 and the new one, covering the 2007-2013 period² require explicitly that evaluation takes place, set out the various stages in its process and state who is responsible for carrying it out. But how is this requirement being met, and how well is it achieving its aims? The authors have pooled their direct experiences gained from conducting Rural Development Plan (RDP) and related rural development policy evaluations (Agra CEAS Consulting, 2003a ; Dwyer *et al*, 2003 ; Dwyer *et al*, 2004 ; Dwyer and Kam-bites, 2006) along with insights from the contemporary evaluation literature and discussion with other UK evaluators and pol-

icy officials, to generate some lessons and reflections on these questions. It must be stressed that we are not concerned here with the performance of the individual schemes that make up RDPs, or of RDPs in their entirety. Rather, we look at the difficulties encountered in reaching an evaluation judgement, in the context of RDPs in general and with specific reference to evaluation experience in the UK.

This paper begins (section 1) by reviewing the purpose of evaluation exercises and outlines how they are perceived and organised within the EU policy framework for rural development. Section 2 then discusses some of the problems and issues inherent in this approach, illustrating these with reference mainly to the experience in the UK, although other country examples are also mentioned where relevant. The discussion is prefaced with a brief overview of key features of rural development policy in the UK, to enable readers to more fully understand the context within which this experience is set. Following the discussion of issues, a summary is drawn together in section 3 and we assess the extent to which the procedures developed for the new programming period (2007-2013) may represent progress in tackling these. Finally, in section 4 we reflect on the need for continuing reform and development in future, with a particular focus on seeking to ensure that evaluations serve a useful purpose for a range of relevant stakeholders, without giving rise to disproportionate resourcing requirements.

1. Regulation 1257/1999.

2. Regulation 1698/2005.

The purpose and process of EU policy evaluation

According to the European Commission (CEC, 2004) the main purposes for carrying out evaluations are: (a) to contribute to the design of interventions, including providing input for setting political priorities ; (b) to assist in an efficient allocation of resources ; (c) to improve the quality of the intervention ; (d) to report on the achievements of the intervention (*i.e.* accountability).

Evaluation itself is described as the *judgement of interventions according to their results, impacts and needs they aim to satisfy* (CEC, 2000a).

It is clearly seen as an important activity which merits careful and thorough treatment. The European Commission has produced a number of guides to evaluation that form a substantial plank of existing literature on this subject. The first major publication – the six-volume MEANS Collection of 1999 issued by DG XVI – contains a wealth of detailed methodology and techniques intended to improve and promote evaluation methods, especially when applied to Structural Fund operations, which would include rural development. The other general treatise is the 2004 Practical Guide to Evaluation (CEC, 2004), issued by DG Budget, from which the 2005 Evaluation of EU Activities – An Introduction (CEC, 2005a) is an extract. This incorporates the requirements of the Communication on Evaluation (CEC, 2000a), the 2002 Financial Regulation and its Implementing Rules³ and the 2002 Communication on Evaluation Standards and Good Practice⁴. DG Agriculture also produced specific guidance for the 2000-2006 programming period relating to both the *ex ante* and mid-term evaluations of Rural Development Programmes (CEC,

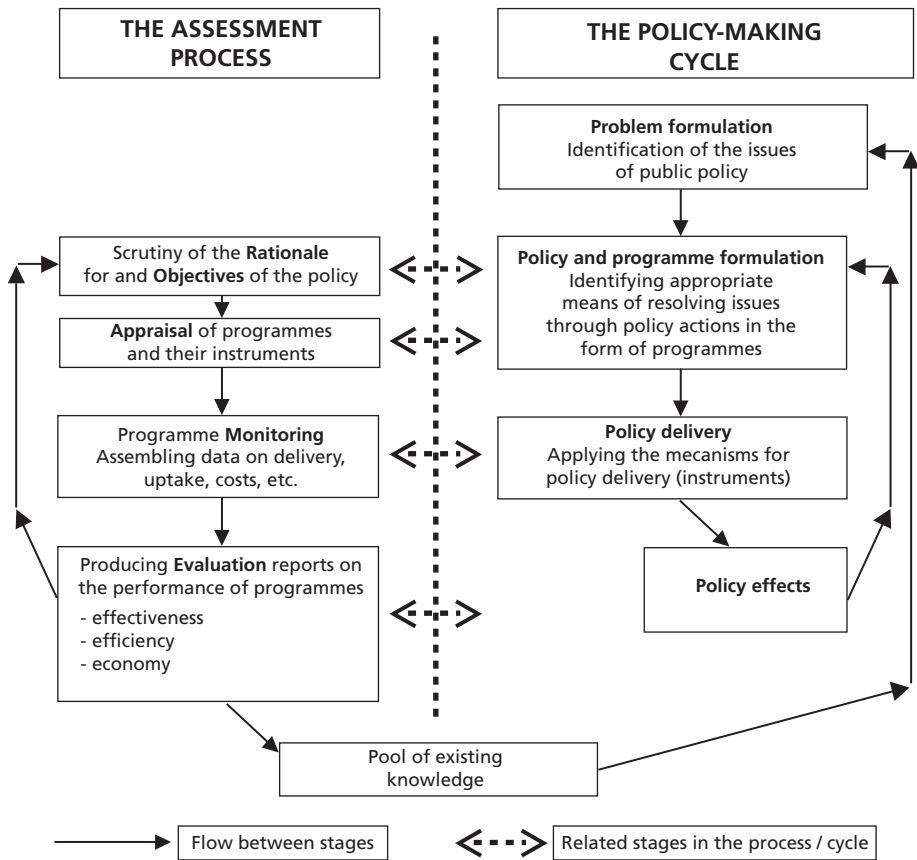
1999b, 1999c, 2001). Another document has also been important in this context. After consultation with Member States, the European Commission issued a set of Common Evaluation Questions (CEQs) with suggested criteria and indicators, to be used in evaluations (CEC, 2000b). Its purpose was to assist with synthesising the evaluations of individual RDPs into an assessment for the EU as a whole, and to achieve a degree of commonality of approach. These CEQs (listed in *Annex I*) and specified indicators will be returned to below, but in overview, groups of questions covered each chapter of the RDR 1257/1999 in turn, with an additional set of “cross-cutting” questions relating to the operation of the RDP as a whole. It is this suite of CEQs, in addition to the specific RDP guidance, which together comprise the most tangible European Commission influence upon the process of evaluating rural development programmes.

However, it is worth noting that different actors see the evaluation of RDPs from perspectives that have varied marginal benefits and costs, which can colour their attitudes and thus potentially also influence results. For national governments, evaluation is an obligatory activity that absorbs resources, so there will be an incentive to economise on evaluation expenditure. However, the extent of co-financing, as well as the likely political impact of findings, also differentially flavour national enthusiasm for these exercises and in most cases there should be a direct interest in ensuring that the findings of formal evaluations can be used to enhance policy performance. At the more local level, concurrent and informal evaluation of policy performance among those delivering the policy is almost inevitable, and this can often lead to positive developments as policy “beds down”, over the programming period. However, for these same groups the gathering of data and provision of qualitative support for formal evaluations can all too readily become an irksome burden, if the results of the exercise appear too distant or

3. Council regulation 1605/2002, articles 27, 28 and 33 and Commission regulation 2342/2002, article 21.

4. The Communication on Evaluation Standards and Good Practices (SEC, 2002/5267).

Figure 1. The assessment process in relation to the policy-making cycle



Source: adapted from Hill, Young and Brookes (1989).

the procedures too cumbersome to furnish useful lessons for direct application. In some cases, local deliverers can see formal evaluation as a threat, or a cynical exercise which will deliver little of benefit to them. Where these situations arise, the quality and consistency of the formal evaluation can be significantly compromised (Jackson, 2001). Even for the external consultants commissioned to undertake evaluations, the incentives involved in such an exercise almost inevitably shift during the tendering and execution stages of an evaluation, as its instrumental nature and its practical and political limitations become gradually more apparent.

In principle, the assessment of policy is an activity that runs in parallel with the policy process. Whilst recognising the multiple layers of specification and reinterpretation by government administrations and agents involved in policy design and delivery (Schneider and Ingram, 1997), the overall process can be presented in simplified form as a cycle which operates in parallel to the cyclical activity of policy development and implementation (figure 1). Some assessment is undertaken in the *preparatory phase* (when a policy intervention is being designed), in which the rationale for intervention is established (problem formulation), the objectives of action are clarified, appropriate indicators

are selected for monitoring performance, and the choice of instruments is determined. Some further assessment normally takes place while the instruments are being interpreted and *implemented* (e.g. monitoring against “process” targets such as how quickly or cheaply cases are dealt with, and/or ensuring that delivery mechanisms work as they are intended to). Perhaps the most widely recognised kind of assessment is once a policy has been in operation for some time, when it takes a retrospective view of the *outputs* from interventions and their impact on the basic problems at which the policy is aimed (*outcomes*), as reflected in the indicators previously determined. As a result of such assessment and its findings we usually return to the start of the cycle, in that modifications or reforms may be planned through a new “preparatory phase”, and the process goes around again. In the evaluation literature, the term *appraisal* is often used to refer to *ex ante* assessment activity in the preparatory phase, while *evaluation* is often reserved just for the *ex post* retrospective assessment. For our discussion here, *evaluation* is used in its generic sense, applying to all stages.

The European Commission’s view of the stages in the evaluation process is rather temporal and operational in nature, but broadly coincides with the cycles as depicted in *Figure 1*. It specifies *ex ante*, mid-term and *ex post* evaluations (taking place before, during and after the implementation of a programme) that are each required to answer rather different sets of questions⁵ The literature on evaluation emphasises the importance played by “feedback” to secure the effectiveness, efficiency and economy of interventions. However, the acronym ROAME(F)⁶, often used to remind UK evaluators of the areas to which they must pay

attention, places feedback at the end of a linear progression. In reality of course, it is possible for lessons to be learned before the intervention has reached its (planned) conclusion, particularly with respect to the performance of delivery systems. Indeed, many interventions may periodically alter course in response to external changes, in which case ongoing or regularly repeated feedback exercises will be necessary. So, a cycle of learning can come from all three stages of the European Commission’s scheme of evaluations, even though the experience of the interventions in hand can only be investigated through the mid-term and *ex post* stages.

The UK experience with evaluation

1. Context - rural development policy and programmes in the UK

To set the scene for discussion of the evaluation process and its limitations, we here provide some basic information about the general orientation and organisation of rural development policy and programmes in the UK. This is intended to help the reader to understand the extent to which the conclusions drawn from a UK perspective are likely to be more widely relevant.

The UK planned expenditure of around €7.6 billions on rural development measures under Reg. 1257/99, over the period 2000-2006. The budget was significantly influenced by the UK’s decision to apply voluntary modulation to direct payments under the first pillar of the CAP, to provide additional funding for the continued growth of agri-environment and woodland creation schemes through this period. This growth would have been more or less impossible within the EU budget allocation, given commitments already existing by 1999, and without making significant cuts to other RDP measures (Ward, 2002). Also importantly and reflecting a wider political agenda for devolution, the UK chose to deliver its rural development policy through four sep-

5. See annex C of CEC (2004).

6. Evaluators are reminded to probe the Rationale, Objectives, Assessment, Monitoring and Evaluation of a measure, and to ensure that Feedback takes place.

arate territorial units: the major principalities of England, Scotland, Wales and Northern Ireland. In addition, some parts of each territory qualified for Objective 1 or Objective 1 transitional status. Thus each of the four 'regions' of the UK had an RDP comprising the EAGGF⁷ Guarantee-funded measures, and each also had one or two separate (mainly more geographically-limited) Objective 1 operational programmes combining EAGGF Guidance-funded measures with other elements funded by the European Regional Development Fund (ERDF) and the European Social Fund (ESF). In contrast to France, the UK did not choose to operate additional, targeted EAGGF programmes in "Objective 2" areas.

While the 1999 RDR marked a significant shift of policy under Agenda 2000, in practice it was an amalgamation of previous instruments, supported by the new rhetoric of "integrated rural development" but without a clear strategy at EU level (Dwyer *et al.*, 2007). It is perhaps unsurprising, therefore, that the 2000-2006 RDPs in the UK were dominated by the continuation of pre-existing schemes and goals, reflecting an accumulated legacy of several decades. This meant that the individual programmes were complex, not fully coherent and sometimes difficult to translate into consistent and testable objectives. Such a situation makes evaluation difficult. We believe this situation was common to many of the EU-15 Member States, in the 2000-06 period (see also Dwyer *et al.*, 2003).

The table 1 indicates the proportionate allocation of the € 7.6 billions of total public funding (EU and UK) on different rural development measures under Regulation 1257/1999 within all the relevant UK programmes, as described above. EAGGF rural development funding within the special additional "PEACE II" structural fund programme in Northern Ireland is also included,

7. European Agricultural Guarantee and Guidance Fund.

Table 1. UK planned expenditure on EU rural development measures (2000-2006)

Measure or group of measures	% of total expenditure
Farm investment	1.4
Training	1.7
Processing and marketing	4.7
Less Favoured Area aids	28.4
Agri-environment aids	37.3
Afforestation of farmland	9.5
Other forestry measures	2.9
Farm relief services	1.8
Marketing of quality products	1.2
Basic rural services	0.7
Village renewal	1.1
Farm diversification	1.7
Farm and rural infrastructure	0.7
Tourism and crafts	1.3
Environmental improvement	1.4
LEADER +	4.0

Sources: Published RDPs and programming documents for all programmes using EAGGF funds (Guidance and Guarantee).

as well as funds for all LEADER+ programmes in the UK.

The pattern reflects a strong public commitment to two goals: the continuation of longstanding support for marginal hill and upland farms through the Less Favoured Area (LFA) compensatory allowances; and the enhancement of environmental management on farms through agri-environment schemes, targeting biodiversity, landscapes and the cultural heritage. The first goal represents a policy position within the UK which pre-dates its accession to the European Community in 1973. The second has its origins in response to strong environmental lobbying by Non-Governmental Organisations (NGOs) and agencies during the 1980s. Neither goal was identified explicitly as *rural development policy*, from the outset.

The spending plans are broadly in line with current UK *agricultural* policy priorities, namely using public funds for environmental "public goods"; with a subsidiary

focus on offering modest support to help farmers adjust to the challenge of CAP reform (by diversifying, re-skilling and adding value to their business activities). But the UK programmes did not correspond closely to the broader concept of rural development embodied in EU and European Commission policy rhetoric, during the 2000-2006 period (summarised in Dwyer *et al.*, 2007). This mismatch of policy focus between the two levels (EU and UK) has led to some difficulties in addressing the formal programme evaluations in the manner envisaged by the European Commission, in its guidance documents.

Another factor which greatly complicates rural development policy evaluation in the UK is the significance of non-CAP funded schemes and measures, which also promote or significantly affect rural development goals. Firstly, in England, Wales and Scotland, significant domestic schemes and programmes, often organised at NUTS 3 regional level⁸ or lower, targeted rural social and economic development during the 2000-2006 period. These included community planning, small business support and support for local rural services and amenities (transport, health, social housing, support for the elderly and funding to target social exclusion). Secondly, the Foot and Mouth disease epidemic across the UK in 2001-2002 gave rise to a suite of targeted "rural recovery" initiatives, again funded by UK domestic resources, in addition to the RDP funding. Thirdly, the EU Structural Funds 2000-2006 (ERDF and ESF) supported some rural areas in the UK through Objective 1 and 2 programmes – this funding was particularly important in Northern Ireland during the period. These sums are not included in *table 1* because they are independent of Pillar 2 of the CAP. However, it is likely that their combined budgets are at least comparable in scale to that of the RDPs. As such, they

8. NUTS: Nomenclature of Territorial Units for Statistics.

have undoubtedly affected rural development in the UK, over this period, and influenced the operation and targeting of the Pillar 2 programmes.

2. Experience with the evaluation process

The separate mid-term evaluations of the four RDPs in the UK, as well as the rural development elements in the five Objective 1 programmes and PEACE II, were undertaken by various independent consultants, in 2003. The precise methods varied between studies but all those evaluating RDPs included a mix of data analysis, interview surveys (postal, telephone and face-to-face) and case studies, and all attempted to answer most of the European Commission's Common Evaluation Questions (CEQs) for the measures used. Overall, the programmes were judged to be on target to deliver against their goals, with a few minor exceptions. However, most of these studies also included critical examination of the difficult methodological issues involved in the evaluation process, which raise important questions about the usefulness of the exercise as a whole.

The numbered sub-sections below discuss four key issues that have influenced the effectiveness of the RDP evaluation processes for the 2000-2006 period, in the UK at least. These are: issues of timing and learning from the evaluation cycle; balancing appropriateness and consistency in formal EU evaluations; methodological challenges and weaknesses; and data and resourcing issues.

The evaluation cycle - Timing and learning

As discussed in section 1, the opportunity to learn from evaluations is a main purpose for using public resources in this activity. However, experience with formal RDP evaluations in the UK indicates that this potential has been less than satisfactorily realised.

Firstly, there is the difficulty of timeliness, resulting in part from the design of RDPs and the evaluation process, and in

part by events transpiring to disrupt the process. Dates by which the various evaluation stages have to be completed are set out in legislation. Because programming periods run consecutively without a break, the planning and agreement of a new programme are meant to take place while the existing programme is still running. The timing of *ex post* evaluations inevitably means that their findings and recommendations can only be fed in to the next-but-one programming period. The ability to learn from experiences of RDPs in other countries depends on these being made available, and a synthesis commissioned by the European Commission from which individual Member States might benefit is necessarily further delayed, sometimes well into the period of operation of the next programmes. These factors therefore put great responsibility onto the *ex ante* and mid-term evaluations to improve the performance of policy interventions. But the mid-term evaluation itself takes place less than three years into the period of operation of the programme. While in an ideal world it is possible for such evaluations to feed into the preparation of the next RDP (*figure 3*) as well as the preparation of an evaluation strategy for the next round of programmes, in reality the learning process can be hampered if information on the performance of policy measures in the most recent programming period is not to hand when needed. This can easily happen if internal factors (such as failure to meet deadlines), or unanticipated external events, disrupt the timetable. In respect of some RDP measures, it may indeed simply be too early to gather meaningful information on performance. Forstner and Plankl (2004), discussing experience in Germany, suggest that the impact of RDP farm investment schemes applied for in 2000 is only likely to be reflected in audited farm accounts in about 2005 (bearing in mind that at least two years of data are required to establish the financial picture

reliably). This implies that even the earliest applications cannot adequately be covered by the 2003 mid-term evaluation.

In the UK, both design and accident hampered the learning process. *Ex ante* evaluations, which *inter alia* must look at basic rationale, examine evidence on past interventions and consider alternative instruments, are needed before RDPs can be approved. At the same time, this is an important period for setting a baseline for the two later monitoring exercises, so that change can be tracked effectively. When the necessary regulations are agreed relatively late, this stage can be seriously compromised. In the case of Wales, for example, the haste with which the 2000-2006 RDP was prepared following the Regulation's publication in spring and implementing regulations in early summer 1999 precluded a satisfactory *ex ante* evaluation, and no separate baseline survey was carried out. The devolution of powers to the Welsh Assembly Government on rural development issues was also a contributory factor (Agra CEAS Consulting, 2003a). Thus attention fell on making best use of existing schemes within the new framework provided by the RDP, with more innovative aspects (such as under Article 33 on the quality of life) following later.

Learning from the mid-term evaluations, both to inform the later stages of the 2000-2006 RDPs and to help in the planning phase of the 2007-2013 RDPs, has been similarly compromised. European Commission Regulation⁹ which sets out the rules for applying Council Regulation¹⁰ (the Rural Development Regulation, RDR) specifies that mid-term evaluations had to be sent to the European Commission not later than 31 December 2003. In practice, delays to some UK schemes by unanticipated external factors (for example, the Foot and Mouth Disease epidemic during 2001-2002), as well as

9. EC n° 445/2002.

10. EC n° 1257/1999.

the need to further develop some of the more ambitious new schemes meant that, by the latter half of 2003, some had only just commenced implementation. For these, 2003 was far too early for a mid-term evaluation¹¹. This situation applied in respect of the mid-term evaluations of RDPs in England and Wales and of the rural development element in both Operational Programmes in Northern Ireland (ADAS and SQW, 2003; Agra CEAS Consulting, 2003a; DTZ Pieda, 2004). On the other hand, the evaluations were arguably made too late to significantly influence policy developments at EU level. The drafting and agreement of the 2007-2013 RDR took place before the European Commission's meta-evaluation of the 2000-2006 period (built upon the mid-term evaluations of individual RDPs) could be completed, in 2004-2005.

Concerning the *ex post* evaluation of the 2000-2006 RDPs, Reg. 445/2002 specifies that this shall take place within two years of the end of the programming period (that is, by 2008). Obviously these evaluations are too late to feed in information to the design of the 2007-2013 RDP. But paradoxically, they may also be too early to capture the impacts of some elements of the RDP, such as education and training, where the impacts are expected to develop over a protracted period. Thus there must be a risk that in these circumstances, *ex post* evaluations are significantly devalued.

Because the current RDPs represent a first round of CAP second pillar programmes, there is no direct precedent whose *ex post* evaluation could feed into programming and implementation of the 2007-2013 second round, or into the 2003 mid-term evaluation process¹². Nevertheless,

given the origins of the measures marshalled under Reg. 1257/1999, relevant *ex post* evaluations would include those for previous Structural Fund programmes (Objectives 1, 5b and 6); and CAP accompanying measures. In the UK, for some of these, full *ex post* evaluations were not undertaken in time to be used in the 2003 mid-term evaluations of RDPs. The reasons for this partly relate to policy ownership: structural fund programmes were generally overseen by one part of government while the new RDPs were the responsibility of a different part. Hence the officials with direct responsibility for the old programmes would see no immediate benefit to their own operations from conducting timely evaluations which could potentially assist another, separate sphere of government activity. *Ex post* evaluation was therefore not a priority for funding, and held little political impetus. In England, a more narrowly-focused study to evaluate the main lessons from only *those elements of 5b programmes* that would be relevant to the England RDP (2000-2006) was eventually commissioned by the Department with responsibility for the new RDP (Department for Environment, Food and Rural Affairs), in 2002. But by then, with funding having ceased and many Objective 5b delivery staff having moved to other jobs, it was difficult to capture the lessons of this experience fully.

The Synthesis of Rural Development Mid-Term Evaluations (Agra CEAS Consulting, 2005b) suggests that this undervaluing of *ex post* evaluation is not restricted to the UK. Member States may significantly reduce interest in doing a thorough *ex post* evaluation, particularly if the regulations concerning new programmes have changed and therefore the comparability of performance between programming periods is not straightforward.

The result of these different factors is that to date, in several UK regions, there has been little opportunity to learn from the formal EU-directed evaluation processes

11. The Welsh Assembly Government chose not to update the mid-term evaluation at the end of 2005, as provided for in Reg. 445/2002.

12. However, evaluations of Reg. 950/97 and Reg. 951/97 did provide some insight as the measures were broadly similar, albeit implemented in a less integrated manner.

applied to the 2000-2006 period, in preparing the next RDP, covering 2007-2013, during 2006 and 2007. Instead, the process has been influenced much more by separate, domestic policy and audit reviews of the various constituent measures of the programmes, as well as by ongoing discussion and debate with stakeholders. Also, as a result of the lack of comprehensive programme evaluation, a greater responsibility has been placed on the *ex ante* evaluations for the next programming period, to convince the European Commission that the new programmes are well-founded and appropriately targeted.

Tensions between appropriate and consistent evaluation

The MEANS collection of evaluation guidance (CEC, 1999a) has been superseded by the later guides issued by DG Budget and DG Agriculture and referred to earlier. However, what MEANS described as one important phase of the evaluator's work – the setting of evaluation questions – was circumscribed in the evaluation of the 2000-2006 RDPs, by the development of the Common Evaluation Questions, set by the European Commission after discussion with Member States (CEC, 2000b). Thus the role of an *ex ante* evaluation has been somewhat constrained compared with the MEANS version in which the formulation of appropriate evaluation questions, based on perceived problems, is seen as an important step. The fact that the questions to be addressed were prescribed centrally, once most programmes had been drafted but in advance of evaluation contracts being let, might also imply a reduction in the independence of the exercise.

The CEQs issued by the European Commission were accompanied by detailed requirements for addressing them, based around common indicators and criteria for judgement. Because rural areas are heterogeneous, even within countries, it is likely that evaluators of different RDPs would

arrive at rather different sets of questions, were they to start *ab initio*. But it is equally understandable that the European Commission wishes to promote consistency in approach and measurement, given its need at a subsequent stage to synthesise an overall evaluation at EU level. As a result, between these two perspectives there is an inevitable tension.

Annex 2 shows one of the simplest sets of questions, indicators and criteria, covering the RDR Chapter on vocational training. The table also includes one UK evaluator's proposed response to these questions, as agreed with national policy officials, and set out in the Baseline Study of the England RDP (IC Consultants, 2002). From this, it is already apparent that the resulting approach to evaluation is neither straightforward, nor likely to be uniform between different programme areas. The reasons for this relate both to the appropriateness of the questions and indicators specified by the European Commission, and the methodological challenges involved in seeking to address them, as we shall seek to explain in this and the next section, respectively.

Some CEQs bear little relevance to the circumstances of particular countries or regions, nor to the ways in which they have devised the targeting and delivery of RDP measures. Also, because of the lateness of the publication of CEQs, RDP monitoring systems were not established with appropriate indicators for generating the answers, as they envisage. Thus it is perhaps unsurprising that evaluations of individual RDPs lack some of the details that the European Commission intended should be included, which gives problems when studies are compared. Alternatively, and perhaps of more concern, the evaluators may attempt to supply answers using proxy measures to avoid gaps. When this occurs, there is a danger that the information provided is of poor quality, or doubtful validity. There is also a risk of creating apparently spurious comparability between answers to CEQs

that have been derived using quite different methodologies. Two examples from the UK, in addressing CEQs for agri-environment measures, illustrate these points.

CEQ VI. 3 asks:

“To what extent have natural resources been protected (or enhanced)...in terms of the quantity of water resources, as influenced by agri-environment measures?”

This has three criteria:

1. *The utilisation of water for irrigation has been reduced (or increase avoided),*
2. *Water resources protected,*
3. *Protected water resources give benefits.*

These criteria are not relevant for most of the UK, where Wales, Scotland and Northern Ireland can all be described as regions with rather more water than is usually comfortable. Indeed it might have been more relevant, where the overarching European Commission interest is impacts upon water quantity, to focus UK attention on the extent to which the schemes helped to promote flood prevention. This is an issue which has been receiving increasing attention in the UK in recent years, particularly in the context of climate change.

CEQ VI.1 on soil conservation impacts is difficult to address because none of the UK agri-environment schemes explicitly targeted soil protection as a goal, but all are likely to have delivered some benefits for it, as a by-product of measures focused mainly on biodiversity and landscapes. As a result, in the absence of appropriately planned baseline data, the mid-term evaluations each used different *ad hoc* methods to impute values for this question. The consultants for the Scottish RDP separately measured the simple correlation between the areas in schemes and areas of particular soil erosivity, and then looked at the total area of land under various soil-conserving measures in the scheme, measuring each of these items separately (DTZ Pieda, 2003). The con-

sultants evaluating the Welsh RDP used farmer interview data to estimate reductions in sources of soil contamination by nutrients and pesticides, but did not examine soil erosion reduction at all (Agra CEAS Consulting, 2003a). In England, consultants attempted first to identify the pattern of uptake of only those specific prescriptions with a soil conserving impact and then to compare this with known erosion-sensitive areas around the country, to produce a composite measure of likely impact (ADAS and SQW, 2003). This diversity of approach significantly reduces the absolute comparability of the results obtained, such that little confidence can be placed in apparent differences between the figures for each region (Dwyer and Kambites, 2006).

The *cross-cutting* CEQs were intended to address some overall impacts of each RDP. However, there were significant issues regarding the relevance of these questions, in the UK context. For example, the CEQ *To what extent has the programme helped stabilise the rural population?* is of dubious validity, given the relatively low level of RDP aid compared with much more significant macro-economic impacts of policy and market trends. The question itself is also arguably no longer relevant in most of the UK (and increasingly in other Member States), where population growth resulting from urban-rural migration (rather than depopulation) is the dominant phenomenon in rural areas. It could have had more value if focused more clearly on identifying RDP impacts on population trends *in those areas* where decline is known to be a (continuing) concern.

Methodological challenges

• Outputs versus impacts

Policy instruments, particularly financial inducements, constitute systems with inputs and outputs (including side effects) that can, or are intended to, lead to impacts. The link between outputs and final impacts is one of rationale, shaped by cause-and-effect

relationships. Again the terminology of outputs, leading to results and then impacts, although consistent within the European Commission, is not internationally standardised, but the notion of intermediate and final objectives is clear enough. Indicators, of course, can relate to any stage of the system. Though it is often easier to measure outputs rather than results or impacts (outcomes), it is the latter that are ultimately more relevant to the problems and/or opportunities that policy interventions generally seek to address.

Several difficulties surround the linking of criteria and prescribed indicators, as set out in the CEQs. Some indicators are poorly-chosen. For example, in order to indicate how the farm investment measure may have affected farm incomes, the income measure specified is “gross farm income”. This is not a commonly used indicator within the EU-wide Farm Accountancy Data Network (FADN/RICA) system (it corresponds to Gross Value Added), and a measure such as Family Farm Income (or even Farm Net Value Added) would appear superior¹³.

Many indicators, especially relating to agri-environment and forestry measures, to training, and to support for broader community development actions under Article 33, relate to outputs rather than impacts (e.g. area of land under agreement, number of persons trained, number of projects aided). This allows consideration of issues such as scheme “reach”, but does not allow for an assessment of the benefit to, for example, biodiversity or soil quality, the level of local skills and knowledge, or the quality of life of the rural population. Given that some impacts take a long time to work through within the RDP evaluation time frame, the use of proxy indicators is inevitable. However, an over-reliance on output indicators will clearly reduce the

ability of evaluations to really grasp whether or not these schemes are effective.

- Additionality, deadweight, displacement and scale

It is clearly important for evaluators to attempt to distinguish the *additionality* offered by the RDP – *i.e.* being clear about what it has delivered which is over and above what might have happened in its absence. This is very difficult to measure with any precision, particularly over short time-periods. For example, questions about investments that can affect economic activity typically use the criterion that farm employment is maintained or increased (such as by encouraging alternative activities on the holding). In order to provide figures, a firm grasp of the counterfactual is needed; *i.e.* consideration of whether similar or alternative investments might have been made if the grant were not given, and the potential displacement effect of the grant (stimulating employment by one enterprise which merely captures market share from others that contract, without actually increasing the size of the overall market). The displacement effects of investment aids for a number of purposes, including processing and marketing, farm diversification and farm tourism, have long been recognised as an issue in the UK (ADAS and University of Reading, 2003) and more widely (Thomas *et al*, 2000). Similarly, studies in a number of countries and contexts have highlighted the potential for significant deadweight in schemes which support farm investments for modernisation (Dwyer *et al*, 2004). However, empirical information on such effects is rarely available over the medium-term, meaning that assessments of additionality in evaluation reports are usually based largely upon beneficiary and/or expert opinions. While these can provide valuable insights into how instruments have operated, they can also be open to challenge (not least by the authorities who have commissioned the evaluation, particularly in

13. See Agra CEAS Consulting, (2005b) for further comments on the appropriateness of individual indicators.

cases where external views conflict with their own). By contrast, more robust treatment of these kinds of issue can be found in some longer-term evaluation studies (Baldock *et al*, 2002, which examines the evidence for CAP impacts upon the environment over several decades).

The counterfactual is particularly problematic with LFA payments (as noted, these are a prominent component of all the UK RDPs), where the system of farming is unique to these marginal areas. The criterion that LFA payments should provide “appropriate compensation” to counteract the economic impact of the natural handicap, as reflected in higher production costs and lower output value, is almost impossible to verify. This is because of the absence of a lowland comparator, and the variation in the degree of handicap between farms within individual LFAs. It is interesting to note that on this issue, separate evaluation of the LFA policies in each of the UK countries during the 1990s came to quite different conclusions concerning their effectiveness, although all tended to support their continuation (Drew Associates, 1997 ; Midmore *et al*, 1998; Davis *et al*, 1998). In the mid-term evaluations, consultants used a variety of approaches to attempt to assess the adequacy of compensation, but the results were acknowledged as limited, in each case.

When considering the cross-cutting CEQs, there is an understandable interest in the effect of RDPs on the incomes and jobs of the non-farming rural population. However, this presents real difficulties for the evaluator in detecting any impact in rural areas where agriculture/forestry is already a very small component of the economy; and in establishing the counterfactual, where the influences on the jobs and incomes of rural residents come mainly from non-RDP, non-agricultural and (probably) non-rural factors (such as the rate of growth in the national economy, interest rates, property prices, etc.). Both these conditions apply to the UK situation.

- Evaluation of instruments versus programmes

The CEQ approach, with its suite of separate questions for each of the Chapters in the Regulation, has tended to encourage an instrument-based approach to programme evaluation. This creates difficulties for evaluators wherever programmes have taken either a multiple or an integrated approach in their design of instruments. In all of the UK RDPs, there is more than one scheme implementing certain chapters - most notably in respect of agri-environment schemes - and in some of the programmes, several measures or sub-measures of the Regulation are delivered through a single scheme. Furthermore, while the performance of individual schemes or Chapters may be satisfactory, this does not necessarily lead to a satisfactory performance for the RDP as a whole and, where areas are subject to overlapping programmes, it does not guarantee successful rural development policies, overall. This depends *inter alia* on the relationship between schemes and RDP objectives, the balance of resource allocation between them and any gaps, conflicts or overlaps in coverage, as well as opportunities for synergy. When comparing the findings of formal UK RDP evaluations with broader comparative analyses (Dwyer *et al*, 2003 ; Shucksmith *et al*, 2005), the evidence suggests that these issues were inadequately covered in the formal exercises.

The experience of UK evaluation of the LEADER Community Initiative (specifically, LEADER II and LEADER+) stands in marked contrast to this instrumentalist approach. LEADER adopts a specific methodological approach emphasising the integration of goals within joint actions and the need to work with targets and spatial strategies devised locally by Local area Action Groups. It also lays emphasis upon capacity building and empowerment among communities and between sectors, as a key element in effective delivery of its goals. Thus the measurement of crude single-

objective outputs or even outcomes has been recognised as insufficient for these kinds of action (Midmore, 1998) and a variety of alternative approaches has been used. In particular, the evaluations often focus just as much on *soft* (i.e. qualitative and attitudinal) impacts from the process of implementing LEADER and how this has been perceived by those closely involved in it, as on *hard* programme and project outcomes. Various methods have been used to measure increases in the capacity of local actors to manage and steer change effectively, including more qualitative and experiential techniques (e.g. Scott, 2004). The LEADER experience is particularly relevant to consideration of RDP evaluation in future, since this approach became part of the new EU rural development funding and programming framework, from 2007.

Data requirements and resourcing

The actual use of indicators is, of course, partly a reflection of the ease and cost with which data can be obtained. Some CEQ indicators were specified for which it is known that no data currently exist. For example, in relation to the question on payments to farmers in LFAs *To what extent have compensatory allowances contributed to the maintenance of a viable rural community?* One criterion chosen is the *fair standard of living for farmers* and the indicator specified is the ratio of (family farm income plus off-farm income of holder and/or spouse) to (average family income in NUTS 2). The absence of this sort of data is an established gap in the EU's statistical system (Hill, 2000 ; OECD, 2003, 2004). Another common problem is in trying to assess the impacts of socio-economic interventions on the wider rural economy, where the available economic data are often undifferentiated as between rural and urban areas (e.g. employment by sector, unemployment, and quality of life indicators such as access to services or general health and well-being). In the UK, the boundaries between urban

and rural areas do not commonly occur coincident with those of the administrative units (e.g. NUTS levels) within which most socio-economic data is collected; thus it is very difficult to identify and measure such rural characteristics. One government agency in England devoted considerable time and resources to try and address this data gap, over the programming period (Countryside Agency, 2004, 2005).

Where data are known to exist, evaluation can place heavy demands on administrative arrangements and processing systems that are generally designed to deliver schemes in an effective manner, rather than primarily to provide data for reporting and assessment. The use of administrative data, where possible, is an attractive alternative to expensive *ad hoc* surveys, but this often involves significant transactions costs, particularly where the use of this data was unanticipated in programme design. For example, in the 2003 mid-term evaluation of the Wales RDP, substantial resources had to be used in data retrieval, consolidation and aggregation because this task had not been planned for in advance by the authorities. Also, some indicators specified in the CEQs go beyond what could be expected to be held within monitoring systems, and would require the use of considerable financial resources on the part of the evaluator, if they were to be addressed thoroughly. In some instances, the costs to individual beneficiaries have also been a significant issue. In the case of Article 33 projects supported under the England RDP, for example, there is evidence that the burden of record-keeping and reporting placed upon individual beneficiaries, at least partly to ensure that the CEQ indicators could be measured, has acted as a major disincentive to people applying for support (Dwyer *et al.*, 2004).

Lack of consistency in data collection may render it impossible to track specific overarching goals or conditions, where these were not anticipated when programmes were conceived. In Wales, for example, the dis-

aggregated way in which the different RDP schemes were administered, reflecting a commitment by the Welsh Assembly Government to devolve scheme delivery to the relevant specialist agencies or local administrative authorities, precluded the collation of records showing which individual farm holdings were in receipt of combinations of assistance from different sources (Agra CEAS Consulting, 2003a).

Because the RDP mid-term evaluation was undertaken conforming to a separate European Commission specification than that used to specify requirements for monitoring Structural Fund programmes, separate, but far less detailed, evaluation on EAGGF-Guidance funded rural development schemes was included in the evaluations of Objective 1 programmes in the UK. Thus it has been impossible to make coherent assessments of particular groups of RDR measures against the overarching goals to which they were principally directed, because their application was governed by different funding streams and monitoring procedures. For example, in the case of forestry, the planting of trees was funded under RDPs because this was one of the four “accompanying measures” funded in all UK territories by EAGGF-Guarantee funds. However, in Objective 1 areas, the subsequent management of woodland was funded under the Operational Programme for the area. Thus in trying to assess the extent to which rural development policy has fostered sustainable forestry – one of the key goals of these particular measures in the UK – the results of both programmes in this respect should be analysed together.

Considering the issue of resources, Member States and Regions generally seek to constrain the costs of evaluation to a small proportion of total programme value – for obvious reasons of probity and public acceptability. The experience from the UK is that in very many instances it has been difficult to answer questions satisfactorily, given the resources made available. In many

instances, answers could only be provided by making reference to separate evaluation studies at national or regional levels. However, the availability of such studies depends very much on evaluation priorities, obligations and approaches at these levels, which do not necessarily correspond to those applying at the EU level.

It should be recognised that RDPs typically serve two separate but overlapping sets of public objectives. Although the EU legislation requires evaluation in relation to EU goals and priorities, RDPs are jointly funded at national and/or regional level. In the case of the UK, this makes national funding by far the main contributor to programmes (a pattern likely to be even more pronounced in the 2007-2013 RDP), particularly as a result of the country’s unique EU budget rebate. Consequently, “national” governments for the UK and its devolved regions can be expected to require their own policy interests to be well served in the evaluation process. Complying with RDR legislation on evaluation, following European Commission guidelines and responding to Common Evaluation Questions may thus be reduced to only what is adequate to secure EU funding. In Wales, for example, there was special interest in how the RDP was helping to support the family farm and affecting the concentration of production, with other priorities including equal gender opportunities and the use of information technology. Where there are potential conflicts of interest (such as between making farms more competitive, and preserving the “family farm”), it seems inevitable that the precise specification of any evaluation will reflect the interest of the main funding source (in this case, the Welsh Assembly Government).

It is also important to note that over the same period 2000-2006, numerous evaluations of various elements of rural development policy were commissioned by the same devolved administrations in UK government. In part, this reflects a longstanding

public sector commitment by which any public-funded schemes must be periodically reviewed (normally this means every 3 to 5 years)¹⁴. Additional studies were also undertaken by the government's independent auditors (the National Audit Office and its devolved equivalents). Further evaluations were commissioned independently by specialist UK agencies and NGOs, mostly to provide earlier, more comparative or more targeted analysis of different elements in programme planning and performance. In this context therefore, there is clearly scope for a high level of duplication in the full range of evaluation activities devoted to RDP measures. As a result, certain quite modest schemes within the UK have been scrutinised many times over, within the space of only a few years. For example, the Welsh Tir Gofal agri-environment scheme, launched in 1999, has already been reviewed, evaluated and/or audited at least seven times in its short lifetime¹⁵. This can lead to a perception among beneficiaries and local administrators that the weight of evaluation applied to a scheme grossly outweighs the risks associated with the resources involved in it.

14. For England, the UK Treasury stipulates that these kinds of policy should be subject to payment rate reviews every three years, and full policy evaluation and review every five years, measured from the date of the originating UK legislation. Thus, an agri-environment scheme created in 1992, for example, would be reviewed in 1997, 2003, 2008 and so on, and its payments updated in 1995, 1998, 2001, etc.

15. Mid-term evaluation review 03, internal "stock-take" review 02, payment review 04/5, WAO review 05/6, EU audit 04, socio-economic evaluation (Agra CEAS Consulting, 2005a), assessment of the role of Tir Gofal in stimulating work for land-based contractors in Wales (Agra CEAS Consulting, 2006).

Summary of weaknesses and assessment of progress in respect of evaluation (2007-2013)

In sum, we have identified a number of significant weaknesses in respect of RDP evaluations in the 2000-2006 programming period. These are :

1. The fact that the evaluation cycle does not enable a good correspondence with the policy making cycle, for these programmes, so opportunities for learning from the formal evaluations are very limited.
2. The likelihood that by prescribing a *common* approach through the CEQ system of questions, indicators and criteria for judgement, results will suffer from significant gaps, inconsistencies, spurious comparability and a lack of appreciation of the integrated and *softer* effects of measures and programmes, as well as a risk of over-reliance on output measures and short-term opinions, in reaching conclusions and making recommendations.
3. The important shortfalls in data availability and the significant resource implications of adequate data sourcing, for each specific evaluation exercise; set against the considerable duplication that is likely to exist between the data sourcing and methodological efforts applied within the various local, national and EU level evaluations of these policies.

This UK experience with the Common Evaluation Questions is unlikely to be unique. For Germany, Forstner and Plankl (2004) note that, while the questions themselves are quite extensive and relevant, criteria and (especially) indicators are incomplete¹⁶, some-

16. The example of lack of completeness given is that relating to the development of beneficiaries' income after having undertaken assisted investments. Here the questioning is confined to what happens on the farm, ignoring the impact on the household that might come from production rationalisation and the release of labour to earn non-farm income (Forstner and Plankl, 2004).

times difficult to understand and unbalanced. The EU-level synthesis of mid-term RDP evaluations found that, while 66% of the CEQs were answered across programmes where they were relevant and 55% of the European Commission's specified criteria were used, only 40% of indicators were employed (Agra CEAS Consulting, 2005b). For the schemes supporting early retirement and for Article 33 (village development / quality of life), only 29% of specified indicators were used. This level of inconsistency must raise questions as to the value of having adopted such a standardised, centrally-prescribed approach to the exercise.

Of course, many of these issues were apparent to those overseeing evaluations at both national and EU levels, and thus it is to be expected that some attempt will have been made to resolve them, in planning for the evaluation of the next round of programmes.

At the level of the European Commission, a number of positive changes have been made, for 2007-2013. These should help to address the issues covered under the first two numbered points, above.

1. Addressing consistency and timeliness

We noted in section 2.1 the fact that a lack of coherence in 2000-2006 programmes, and a lack of correspondence between national and European approaches to RDPs, hampered the effective evaluation of UK RDPs within the CEQ framework. For 2007-2013, the procedure of drawing up a proposal and having it approved by the European Commission has been somewhat modified by the new Regulation (1698/2005). Now the first stage is the submission of a National Strategy Plan (NSP) in the light of Community Strategic Guidelines proposed by the European Commission and approved by the Council¹⁷. This identi-

fies three priority axes, plus the *methodological axis* of LEADER. Among the strategic priorities adopted by the EU, simple income support is not an aim; rather the emphasis is on encouraging competitiveness of the agri-food and forestry sectors and promoting factors that encourage farmers to adapt to changing economic conditions (being "sustainable" in a dynamic sense). The aims of environmental policy are more clearly articulated than before and, while support to some types of farming system can form part of the strategy, this is increasingly seen as payment in return for public goods and services. Similar remarks could be made in respect of the third axis of the new Regulation, where experience and the influence of cohesion policy is leading to a clearer focus upon priority areas, issues and needs. While these changes do not guarantee more straightforward evaluations in future, they appear to create conditions in which this is more possible. Thus in the UK, NSPs were prepared using a template provided by the European Commission (issued 16 November 2005). There was a single NSP in the form of an envelope document containing separate NSPs for the four constituent regions. This process should have encouraged greater coherence of measures within programmes and more commonality of purpose with EU priorities (although it should be noted that the UK has retained its particular emphasis upon the environmental axis, as seen in the 2000-2006 period).

For the RDP 2007-2013, because the Regulation was agreed in September 2005 there was more time to carry out an *ex ante* evaluation, which should be reflected in an improvement in plan quality. For example, in Wales this process started in the fourth quarter of 2005. The Wales NSP (in its UK envelope) was submitted after the final adoption by the Council of the Community Strategic Guidelines (February 2006), with then some two to four months before the RDP was required to be submitted for

17. These were agreed by the Council on 9 November 2005 but await adoption by the Parliament early in 2006.

approval. Proper consultation on the NSP could take place so that the contents of the RDP better reflect the priorities of Welsh stakeholders and the problems of rural Wales, within the bounds of the EU strategy priorities and the Regulation.

This enhanced timeliness should permit the setting up of a statistical baseline and the planning of additional data collection on a regular basis that will enable the wider implications of RDPs to be measured, in so far as these are detectable. It should also permit the arrangement of administrative procedures to allow the better use of data gathered through implementation, in the evaluation of individual components of the RDPs and of their combined results. Such planning should help reduce transactions costs in the evaluation process.

2. Evaluation questions, criteria and indicators

The evaluation of the 2000-2006 RDP presented its demands for indicators too late for new data systems to be set up if they did not already exist, especially to establish baselines. In contrast, for the 2007-2013 programming period the 2005 NSP Guidance Template made reference to indicators that should be used in describing the baseline situation, some labelled as “lead” indicators and others that help set the context. Both sets are grouped according to the priority axis to which they relate, and the majority focus on measures of outcomes (termed “results” and “impacts”) rather than outputs. This suggests that it is known at an early stage which indicators will be given prominence in questions that the European Commission may wish to pose, in the mid-term and *ex post* evaluations. In consequence, for Wales the draft NSP put out to consultation at the end of 2005 already contained the list of indicators that will appear in the baseline and form the basis of quantified objectives and targets for the programme (Welsh Assembly Government, 2005). For the European Commission, its guidance on preparing NSPs

says the following about the process by which evaluation questions, criteria and indicators should be set:

– After adoption of the programmes (second stage), the national strategy should be *adapted to include all relevant common indicators* (baseline, impact, result, output) and quantified targets on the basis of the *ex ante* evaluation. This will form the basis for the strategic reporting. (CEC, 2005b) (our emphasis added).

This suggests that there will, as in 2000, be a further iterative development of evaluation questions after plans are approved. The prior articulation of clearer goals and priorities at both EU and Member State levels should enable any common evaluation questions to be more clearly justified by their relevance to these priorities. However, there are clearly some risks inherent in the process, as well. For example, the European Commission has selected a relatively small number of common impact indicators, suggesting that these are the principal aims of each axis. The indicators are: for axis 1, improved agricultural and forestry labour productivity and economic growth ; for axis 2 reversing biodiversity decline, maintaining High Nature Value areas, improving water quality and combating climate change; for axis 3 economic growth and employment creation ; and for axis 4, building capacity for successful rural development (CEC, 2006). This clearly involves a considerable level of simplification, by comparison with the individual purposes of the measures in each axis. For example, it would not generally be anticipated that support for village renewal, conservation of heritage or the preparation of local rural development strategies (all measures under axis 3) will readily generate employment or economic growth within the period of the programme – their purpose is perhaps equally for capacity building, in many cases. There must therefore be a risk that programme authorities will be overly swayed by the potential for axis 3 spending to translate

into increased jobs during the programming period, and will therefore neglect capacity building measures under this axis. Likewise, supporting enhanced environmental performance on farms using the modernisation measure under axis 1 may not produce measurable effects upon labour productivity or economic growth in the short-term. Does this system therefore discourage the use of modernisation funding for this purpose? The problems of *target-chasing* in response to overly crude performance indicators has been well documented in the evaluation literature (summarised in Jackson, 2001). If the European Commission wishes to avoid these problems it should perhaps ensure that when evaluation questions are set, they make explicit provision for programme authorities to apply a more integrated approach to assessing how different measures from across the axes might contribute to the full set of impact indicators, rather than encouraging too mechanistic a linkage between single indicators and axes.

Outstanding issues, possible solutions and conclusions

Whilst the changes discussed in section 4 are welcome, they clearly fall short of what would be required to address the range of challenges and weaknesses raised in this paper. Two particular concerns are worthy of further discussion here, to examine possible solutions and their implications for the formal EU evaluation process. Firstly, the methodological challenge of effective identification of policy impacts, within the necessary timescale of these programmes, remains. Secondly, we should consider the issue of cost-effective resourcing for EU evaluations when set within the broader context of all the various evaluation and review processes applying to rural development measures.

In respect of the methodological challenge, it would appear important to enable effective impact evaluations to take place

over a longer timescale than that which has so far been achieved. However, at the same time, approaches must be developed which will generate useful results at least before the next generation of programmes is being planned (2011-2012), if not even before the new RDR has to be agreed (this could be drafted as early as 2010). One important ingredient in this process, therefore, should surely be the *ex post* evaluation of 2000-2006 programmes. Among the formal EU exercises, only these offer the prospect of capturing some impacts of many of the more complex or investment-style measures in programmes, and thus also an important opportunity to seek to capture the degree of additionality that they have ultimately achieved, taking into account deadweight, displacement and scale effects. Assessing the discernible outcomes of former programmes against the new strategic priorities of rural development policy could also facilitate more integrated approaches which capture conflicts and synergies between different components in the programmes. At the same time, however, more effort could be devoted to learning lessons from the much longer-term existence and application of various measures within the programmes, through comparative analysis using a combination of existing local and national evaluations, and more targeted new work to complement these.

The risk, however, is that most policy attention will now focus upon establishing systems and commissioning work for the mid-term evaluation of the new programmes, since that reflects most directly upon the day-to-day concerns of programme designers and deliverers. Thus a strong steer from the European Commission for programming authorities to incorporate *ex post* analysis of former programmes, as well as more longitudinal studies of the major enduring rural development measures or approaches and more integrated assessment of programme operation and impacts, within the mid-term evaluation analytical process,

may be needed. In essence, this would focus the mid-term evaluation around the twin themes of learning from past experience as well as tracking current progress, and it would be within the first of these two themes that the major element of evaluation would be most detailed. Such an approach might appear to undermine the consistency and thus comparability of the results obtained, but as we have attempted to show in this paper, these things can all too easily compromise the quality of evaluation, which must ultimately be the most important concern. The approach might also be regarded as perverse in that hitherto, *ex post* evaluation has been considered as something entirely separate which enables a final judgement on past performance, not a comment on current progress. But as we have indicated earlier, the separation of these evaluations from current policy may render them effectively obsolete before they have even been completed. The challenge, therefore, is to design an approach to mid-term evaluation which ensures that programming authorities have to focus much more attention on learning from past experience, and on considering the integrated impact of measures at a strategic level, than they have hitherto been required to.

There is also a question about how the approach needs to adapt in response to the changes that have now been made to the policy framework. The European Commission's stated aim to promote "mainstreaming" of LEADER, in particular, suggests a need for RDP evaluation to learn from these multi-objective and less top-down evaluative approaches, in the same way that the European Commission now expects the policy to do. As noted also by Hodge and Midmore (2008), more holistic, multipurpose and locally-determined policy tools require approaches where evaluators focus more clearly on both **outcomes** (not outputs, because these become too heterogeneous between different local areas), and **processes** (because the delivery processes become an

important element influencing the quality of the programmes and their impacts). These issues are also illustrated well by Lukesch (2003). These themes complement the suggestions that we have already made here, about the value of moving towards evaluation systems that are less instrument-focused and more holistic, in future.

Moving on to the issue of cost-effectiveness in EU evaluation, we believe there is a need to clarify and co-ordinate ownership and conduct of the evaluation process and its rationale, between the various funding sources (EU, national and local) and their different evaluation requirements. These bodies should work towards ensuring a proportionate level of evaluation effort in relation to the financial resources devoted to programmes, from all different levels (EU, national, local). In order to achieve this, it may be necessary for officials to seek either to combine evaluations for different policy "clients", or for one level in the hierarchy to agree to accept the results of studies undertaken for another (lower) level, in lieu of requiring its own bespoke evaluations. This will not happen until there is a greater degree of trust between those involved in the different levels of the administration, fostered by a more fully shared agenda concerning desired outcomes and appropriate performance measures. Such trust can be promoted by improved communication and co-ordination between those responsible for commissioning evaluation exercises, within the different levels of the policy hierarchy. Evaluation consultants could also have an important role here, in that they can contribute their own insights into how exercises can be co-ordinated. By reducing the levels of duplication between separate exercises seeking answers to similar questions, the scope for more in-depth work on issues and themes of common interest to a number of clients could be increased.

However, problems remain because of the very broad nature of rural development and the way that is organised in the EU. Assess-

ing the performance of a RDP as a whole cannot avoid the matter of balance, but making judgements on the allocation of resources between the various types of activity (corresponding to RDR Chapters or Axes) is very difficult because of the problem of assessing environmental and societal goals or needs on any common axis; and also because of the problem of distinguishing RDP policy impacts from those of many other policies operating in rural areas. We have not attempted to address these issues here. Addressing whether “Community Added Value” can be found also raises further unresolved problems.

Nevertheless, our conclusion is that it is

both possible and important to learn from evaluating RDPs, and that the information gathered so far is capable of making a positive contribution not only to evaluation activity but also to the policy-making processes which it supports, at all levels. The challenge for the future will be to enable evaluation methods to be transformed and made more directly useful in this context through a greater appreciation of the strengths and weaknesses of current evaluation practice, combined with an acknowledgement of how the process of rural development policy design is changing, and how evaluation must also adapt, in that context. ■

REFERENCES

- ADAS Consulting Limited and SQW Limited (2003). *The Mid Term Evaluation of the England Rural Development Programme*. London, UK, Food and Rural Affairs, Department for Environment.
- ADAS Consulting Limited and University of Reading (2003). *An economic evaluation of the Processing and Marketing Grant Scheme*. Wolverhampton, UK, ADAS.
- Agra CEAS Consulting (2003a). *Mid-term evaluation of the Rural Development Plan for Wales*. Cardiff, National Assembly for Wales, Report for the Welsh European Funding Office.
- Agra CEAS Consulting (2003b). *Ex post evaluation of Regulation (EC) 950/97*. Report for DG Agriculture, European Commission.
- Agra CEAS Consulting (2003c). *Ex post evaluation of Regulation (EC) 951/97*. Report for DG Agriculture, European Commission.
- Agra CEAS Consulting (2005a). *Socio-economic assessment of Tir Gofal*. Report for Countryside Council for Wales and Welsh Assembly Government.
- Agra CEAS Consulting (2005b). *Meta-evaluation of the Rural Development Regulation 2000-2003 funded by EAGGF Guarantee*. Report for DG Agriculture, European Commission.
- Agra CEAS Consulting (2006). *The role of Tir Gofal in stimulating work for land-based contractors in Wales*. Report for the Countryside Council for Wales, Fforwm Tirlun and the Welsh Development Agency.
- Baldock D., Dwyer J., Sumpsi-Vinas J. (2002). *Environmental Integration and the Common Agricultural Policy*. Brussels, CEC.
- Commission of the European Communities (1999a). *The MEANS collection—“Evaluating socio-economic programmes”*. Six volumes. Directorate for the Co-ordination and Evaluation of Operations of DG XVI.
- Commission of the European Communities (1999b). *Evaluation of rural development programmes 2000-2006 supported from the European Agricultural Guidance and Guarantee Fund -Guidelines*. Directorate General for Agriculture, VI/8865/99- Rev

- Commission of the European Communities (1999c). Indicators for Monitoring and Evaluation: An indicative methodology. The New Programming period 2000-06: methodological working papers. *Working Paper 3, DG XVI, Regional Policy and Cohesion* (Coordination and evaluation of operations).
- Commission of the European Communities (2000a). *The Communication on Evaluation (SEC(2000) 1051)*.
- Commission of the European Communities (2000b). *The set of common evaluation questions with criteria and indicators*. VI/12004/00
- Commission of the European Communities (2001). *Ex ante Evaluation - A Practical Guide For Preparing Proposals For Expenditure Programmes*. DG Budget, December.
- Commission of the European Communities (2002). Guidelines for the mid term evaluation of rural development programmes 2000-2006 supported from the European Agricultural Guidance and Guarantee Fund DOC. STAR VI/43517/02.
- Commission of the European Communities (2003). *Factsheet: Overview of the implementation of rural development policy 2000-2006 - some facts and figures (11/2003)*. Brussels, CEC.
- Commission of the European Communities (2004). *Evaluating EU Activities*. A Practical Guide for the Commission Services, DG Budget Evaluation Unit, July.
- Commission of the European Communities (2005a). *Evaluation of EU Activities*. An introduction, DG Budget, January.
- Commission of the European Communities (2005b). *Establishing the National Strategy Plan: Guidance Template*. STAR, VI/197/396/05.
- Commission of the European Communities (2005c). *Draft Guidelines for Ex ante Evaluation. Rural Development 2007-2013*. Common Monitoring and Evaluation Framework, DG Agriculture and Rural Development, November.
- Commission of the European Communities (2006). *Evaluation Guidance measure fiches*.
- Countryside Agency (2004). *State of the Countryside Report*. Cheltenham, UK, CA.
- Countryside Agency (2005). *State of the Countryside Report*. Cheltenham, UK, CA.
- Davis J., Caskie P., Gilbreath G., Wallace M. (1998). *A Socio-Economic Evaluation of the Hill Livestock Compensatory Allowances Scheme in Northern Ireland*. UK, Queens University of Belfast.
- Drew Associates (1997). *Economic Evaluation of the Hill Livestock Compensatory Allowances Scheme*. London, Report published by MAFF, HMSO.
- DTZ Pieda consulting (2003). *Scottish Rural Development Programme (SRDP) Mid Term Evaluation Report*. Edinburgh, UK, SEERAD.
- DTZ Pieda consulting (2004). *Mid-term Evaluation of Measures funded by the EAGGF*. Belfast, UK, DARDNI.
- Dwyer J., Baldock D., Beaufof G., Bennett H., Lowe P., Ward N. (2003). *Europe's Rural Futures – the Nature of Rural Development II: rural development in an enlarging European Union*. London, IEEP, Pan-European Research Report for WWF Europe/LUPG agencies, January.
- Dwyer J., Slee R.-W., Buller H., Baldock D., Swales V. (2004). *Helping Farmers Adapt – comparative report*. National Audit Office study, at <http://www.nao.gov.uk>.
- Dwyer J., Kambites C. (2006). *Agri-Environment Schemes Evaluation: UK National Report*. Part of pan-EU evaluation for DG Agriculture of the European Commission led by Oreade-Breche consultants, at <http://www.europa.eu.int/agriculture/>.
- Dwyer J., Ward N., Lowe P., Baldock D., (2007). European Rural Development Under The Common Agricultural Policy's "Second Pillar": Institutional conservatism and innovation. *Regional Studies*, Vol. 41.7, October, p. 1-15.

- Forstner B. (2005). *Evaluating Rural Development Schemes in Europe: Are They Delivering on Their Goals?* Presentation at Conference "Rural Development in Europe", Germany, Institute of Farm Economics, Nov.
- Forstner B., and Plankl R., (2004). *Mid-term evaluation of single farm assistance under the Regulation (EC) No. 1257/1999: methodological problems, data shortcomings and possible solutions*. Germany, Institute of Farm Economics (FAL).
- Hill B., Young N., Brookes G. (1989). *Alternative Support Systems for Rural Areas. Vol 1 – The Report. Vol 2 – The Inventory. Report of a research project for the DoE and MAFF*. Ashford, Wye College, Department of Agricultural Economics. ISBN 0-88255-195-1. Vol 1, 201pp, Vol 2 180 pp.
- Hill B. (2000). *Farm Incomes, Wealth and Agricultural Policy*. Aldershot, 3rd ed. Ashgate Publishing, ISBN 0-7546-1132-9.
- Hodge I.-D., Midmore P. (2008). Models of rural development and approaches to analysis, evaluation and decision-making. Paris, *Economie Rurale*.
- IC Consultants Limited (2002). *Baseline Study of the England Rural Development Programme 2000-06*. London, Department for Environment, Food and Rural Affairs.
- Jackson A. (2001). An Evaluation of evaluation: problems with performance measurement in small business loan and grant schemes. *Progress in Planning*, 55, p. 1-64.
- Lukesch R. (2003). *Assessing the equilibrium between autonomy and accountability: The evaluation of LEADER II*. Paper presented at the Fifth European Conference on Evaluation of the Structural Funds, Budapest, at http://www.europa.eu.int/regional_policy), June.
- Midmore P. (1998). Rural policy reform and local development programmes: appropriate evaluation procedures. *Journal of Agricultural Economics*, 49 (3).
- Midmore P., Sherwood A-M., Hounsom B., Hughes G.-O., Jenkins T.-N., Roughley G., Russell S. (1998). *LFA Policy in Wales: A Review of the Socio-economic and Environmental Effects of the HLCA Scheme*. Aberystwyth, Report to the Welsh Office, Welsh Institute of Rural Studies, The University of Wales.
- OECD (2003). *Farm Household Income: Issues and Policy Responses*. Paris, ISBN 92-64-09965-4, Organisation for Economic Co-operation and Development.
- OECD (2004). *Farm Household Income: Towards Better Informed Policies*. Policy Brief, October.
- Schneider A.-L., Ingram H. (1997). *Policy Design for Democracy*. Kansas, Lawrence, University of Kansas press .
- Shucksmith M., Thompson K.-J., Roberts D., Eds. (2005). *The CAP and the Regions- the territorial impact of the Common Agricultural Policy*. Wallingford, UK, CABI publishing.
- Scott M. (2004). Building institutional capacity in rural Northern Ireland: the role of partnership governance in the LEADER II programme. *Journal of Rural Studies*, Volume 20, Issue 1, January, p. 49-59.
- Thomas V., Dailami M., Dareshwar A., Kaufmann D., Kishor N., Lopez R., Wang Y. (2000). *The Quality of Growth*. Oxford, OUP, Published for the World Bank.
- Ward N. (2002). Europe's Rural Futures: *The Nature of Rural Development in an Enlarging European Union*. UK National Report, Land Use Policy Group and WorldWide Fund for Nature, October.
- Welsh Assembly Government (2005). *Rural Development Plan for Wales, 2007-2013. The Strategic Approach*. Welsh Assembly, Cardiff.

ANNEX 1

The set of Common Evaluation Questions VI/12004

I. Investment in agricultural holdings

- I.1. To what extent have supported investments improved the income of beneficiary farmers?
- I.2. To what extent have supported investments contributed to a better use of production factors on holdings?
- I.3. To what extent have supported investments contributed to the reorientation of farming activities?
- I.4. To what extent have supported investments improved the quality of farm products?
- I.5. To what extent has the diversification of on-farm activities originating from supported alternative activities helped maintain employment?
- I.6. To what extent have supported investments facilitated environmentally friendly farming?
- I.7. To what extent have supported investments improved production conditions in terms of better working conditions and animal welfare?

II. Young farmers

- II.1. To what extent has the aid for setting up covered the costs arising from setting up?
- II.2. To what extent has the setting up aid contributed to the earlier transfer of farms (to relatives versus non-relatives)?
- II.2. A To what extent has the setting up aid contributed to the earlier transfer of farms (to relatives versus non-relatives)...in particular, how significant was the synergy with the aid for early retirement in achieving such an earlier transfer?
- II.3. To what extent has the aid influenced the number of young farmers of either sex setting up?
- II.4. To what extent has the setting up of young farmers contributed to safeguarding employment?

III. Training

- III.1. To what extent are the assisted training courses in accordance with needs and coherent with other measures of the programme?
- III.2. To what extent have the acquired skills/competence helped improve the situation of the trainees and of the agricultural/forestry sector?

IV. Early retirement

- IV.1. To what extent has aid for early retirement contributed to the earlier transfer of farms?
- IV.1.A. To what extent has aid for early retirement contributed to the earlier transfer of farms...in particular, to what extent has there been synergy between “early retirement” and “setting-up of young farmers” in terms of an earlier change of holders?
- IV.2. To what extent has the economic viability of the remaining agricultural holdings improved?
- IV.3. Was the income offered to the transferors appropriate in terms of encouraging them to abandon farming and subsequently offering them a fair standard of living?

V. Less Favoured Areas (LFAs)

- V.1. To what extent has the scheme contributed to: (i) offsetting the natural handicaps in LFAs in terms of high production costs and low production potential, and: (ii) compensating for costs incurred and income foregone in areas with environmental restrictions?
- V.2. To what extent have compensatory allowances helped in ensuring continued agricultural land use?
- V.3. To what extent have compensatory allowances contributed to the maintenance of a viable rural community?
- V.4.A. To what extent has the scheme con-

tributed to the protection of the environment...by maintaining or promoting sustainable farming that takes account of environmental protection requirements in LFAs? (*concerns LFA*)

V.4.B. To what extent has the scheme contributed to the protection of the environment... by increasing the implementation and respect of environmental restrictions based on Community environmental protection rules?

VI. Agri-environment

VI.1.A. To what extent have natural resources been protected... in terms of soil quality, as influenced by agrienvironmental measures?

VI.1.B. To what extent have natural resources been protected... in terms of the quality of ground and surface water, as influenced by agrienvironmental measures?

VI.1.C. To what extent have natural resources been protected (or enhanced)...in terms of the quantity of water resources, as influenced by agrienvironmental measures?

VI.2.A. To what extent has biodiversity (*species diversity*) been maintained or enhanced thanks to agri-environmental measures...through the protection of flora and fauna on farmland?

VI.2.B. To what extent has biodiversity been maintained or enhanced thanks to agri-environmental measures... through the conservation of high nature-value farmland habitats, protection or enhancement of environmental infrastructure or the protection of wetland or aquatic habitats adjacent to agricultural land (*habitat diversity*)?

VI.2.C. To what extent has biodiversity (*genetic diversity*) been maintained or enhanced thanks to agri-environmental measures...through the safeguarding of endangered animal breeds or plant varieties?

VI.3. To what extent have landscapes been maintained or enhanced by agri-environmental measures?

VII. Processing and Marketing

VII.1. To what extent have the supported investments helped to increase the competitiveness of agricultural products through improved and rationalised processing and marketing of agricultural products?

VII.2. To what extent have the supported investments helped to increase the added value and competitiveness of agricultural products by improving their quality?

VII.3. To what extent have the supported investments improved the situation of the basic agricultural production sector?

VII.4. To what extent have the supported investments improved health and welfare?

VII.5. To what extent have the supported investments protected the environment?

VIII. Forestry measures

VIII.1.A. To what extent are forest resources being maintained and enhanced through the programme... particularly by influencing land use and the structure and quality of growing stock?

VIII.1.B. To what extent are forest resources being maintained and enhanced through the programme... particularly by influencing the total carbon storage in forest stands?

VIII.2.A. To what extent have the assisted actions enabled forestry to contribute to the economic and social aspects of rural development... by maintenance and encouragement of the productive functions on forests holdings?

VIII.2.B. To what extent have the assisted actions enabled forestry to contribute to the economic and social aspects of rural development...by maintenance and development of employment and other socio-economic functions and conditions?

VIII.2.C. To what extent have the assisted actions enabled forestry to contribute to the economic and social aspects of rural development...by maintenance and appropriate enhancement of protective functions of forest management?

VIII.3.A. To what extent have the assisted actions contributed to the ecological func-

tions of forests...by maintenance, conservation and appropriate enhancement of biological diversity?

VIII.3.B. To what extent have the assisted actions contributed to the ecological functions of forests...by maintenance of their health and vitality?

IX. Article 33 – Promoting the adaptation and development of rural areas

IX.1. To what extent has the income of the rural population been maintained or improved?

IX.2. To what extent have the living conditions and welfare of the rural population been maintained as a result of social and cultural activities, better amenities or by the alleviation of remoteness?

IX.3. To what extent has employment in rural areas been maintained?

IX.4. To what extent have the structural characteristics of the rural economy been maintained or improved?

IX.5. To what extent has the rural environment been protected or improved?

TRANSV. Cross-cutting issues

Transv.1. To what extent has the programme helped stabilising the rural population?

Transv.2. To what extent has the programme been conducive to securing employment both on and off holdings?

Transv.3. To what extent has the programme been conducive to maintaining or improving the income level of the rural community?

Transv.4. To what extent has the programme improved the market situation for basic agricultural/forestry products?

Transv.5. To what extent has the programme been conducive to the protection and improvement of the environment?

Transv.6. To what extent have the implementing arrangements contributed to maximising the intended effects of the programme?

ANNEX 2

Example of Common Evaluation Questions (CEQ), Indicators and proposed data sources as they were interpreted for evaluation in England (adapted from the Baseline Study of the England Rural Development Programme 2000-2006). Source: IC Consultants Limited (2002)

CHAPTER III. TRAINING (ARTICLE 9)

Name of Scheme in England: Vocational Training Scheme (VTS)

ERDP Objective: *to broaden the skills base of the agricultural and forestry workforce to enable it to meet the challenges of the re-orientation of agriculture and forestry and so contribute to the new demands of the rural economy.*

Criteria (based upon EC criteria)	Indicator (as suggested in CEQ guidance)	Q. Is this indicator to be used?	Source/Methods
CEQ III.1. To what extent are the assisted training courses in accordance with needs and coherent with other measures of the programme?			
1. The training responds to the needs and potential adaption (conversion, re-orientation, improvement) at the level of individuals, sectors or regions (including gaps/weaknesses or potential/opportunities identified during programming or <i>ex ante</i> evaluation)	1.1. Share of assisted training accommodating [addressing] issues identified as gaps/weaknesses or potential/opportunities during programming/ <i>ex ante</i> evaluation (%) of which... (a) thanks to [as a result of] the type/mix of participants (<i>e.g.</i> , young people, women...) (%) (b) thanks to [as a result of] the topic/contents of the courses (%) (c) related to co-financed actions of other chapters of the programme (%)	Yes ? Yes Yes Yes	Scheme economic evaluation, scheme application data, farmer/trainer surveys. Relevance and if responding to needs has to be assessed at regional level. Links to other English RDP projects must be tracked.
<p>Comments (by the Baseline Study consultant) on proposed data sources, gaps, etc.</p> <ul style="list-style-type: none"> • <i>The proposed data sources appear to be robust for this purpose.</i> • <i>Contact DEFRA for information on the Economic Evaluation of VTS (due March 2003).</i> • <i>See project documentation for data content of scheme application forms. Application data is held on PROBIS database (note generic problems of using this database and possible incompleteness of coverage; see notes of discussions with scheme managers in Annex 2.1).</i> • <i>To assess "Responding to needs", each region has its own needs assessment in the form of "Regional Targeting Statements" for this and other measures.</i> • <i>Each course participant has a "Training Need Assessment" form prepared for them, prior to the course. TNAs are held by the course organizers and should be available to the evaluators.</i> • <i>VTS operation links particularly with PMG (processing and marketing) and RES (adaptation and diversification of rural economy) schemes under the English RDP – these links should be noted.</i> 			

CEQ III.2. To what extent have the acquired skills/competence helped improve the situation of the trainees and of the agricultural/forestry sector?			
1. The skills/competence acquired by the trainees help improve their employment conditions	1.1. Share of assisted trainees (both holders and employees) experiencing job improvements related to the training (%) (a) of which farm/forest holders (%) (b) of which employees (%) (c) of which thanks to [as a result of] better remuneration (%) (d) of which thanks to [as a result of] non-pecuniary job quality (e.g., seasonal/contractual work security, exposure to risk and adverse conditions, job-variation/enrichment...) (%)	Yes Yes Yes Yes ? Yes	A post- training questionnaire is completed by all beneficiaries. Further surveys may be needed. Income data may be sensitive and therefore not comprehensive.
<p>Comments (by the Baseline Study consultant) on proposed data sources, gaps, etc.</p> <ul style="list-style-type: none"> • Course organizers hold the post-training questionnaires, and these should be available to evaluators – see PROBIS for details of course organizers. • Questionnaires are completed immediately post-training and then (by employers) after another three months. It is unlikely that the full impact of training will have been felt at this stage. • We suggest that further follow-ups are needed to catch the longer-term impacts. This implies additional surveys of trainees at (say) one year and three years. <p>Questions/problems:</p> <ul style="list-style-type: none"> • How to separate out exogenous factors? • How much of any income change is attributable to having completed the course? 			
2. The skills/competence acquired by the trainee facilitates the adaptation of agriculture and forestry (conversion/re-orientation /improvement)	2.1. Share of holdings with an assisted trainee, initiating conversion/ reorientation /improvement related to the assisted training (%) (a) of which new/additional activities (%) (b) of which improved quality/hygiene/added value concerning existing activities (%) (c) of which management related (%) (d) of which environmental benign methods/practices (%) (e) of which farming (%) (f) of which forestry (%)	Yes Yes Yes Yes Yes Yes Yes	By survey.
<p>Comments by the Baseline Study on proposed data sources, gaps, etc.</p> <ul style="list-style-type: none"> • Minutes of meetings with scheme managers note that there is currently no data for regional training priorities, targeted gender and age groups, targeted special interest groups. • The proposal for data collection “by survey” may under-estimate what data already exist. The minutes mentioned above imply that it will be possible from applications received and case-specific monitoring to collect (by region): number of training days, number of workshops, number of qualifications, gender, age, training subject area, sector of employment and measure of impact (conversion/reorientation/improvement). • This does not negate the desirability of follow-up surveys of individual cases to assess the extent of the initiation of conversion/reorientation/improvement referred to. • Contact Lantra (the industry training agency) for alternative data on the wider implications of participation in training schemes. 			