

EUROPEAN PARLIAMENT



Directorate General for Research

WORKING PAPER

**THE PURPOSE AND METHODOLOGY
OF EVALUATION IN REGARD
TO EU AGRICULTURAL EXPENDITURE**

Budgetary Series

BUDG-102 EN

01 - 1999

This study was carried out for the European Parliament's Directorate General for Research by:

AUTHORS: Prof Stefan TANGERMANN
Institute of Agricultural Economics-University of Göttingen
&
Prof Allan BUCKWELL
Wye College - University of London

SUPERVISOR: Anthony COMFORT
Principal Administrator

PUBLISHER: **EUROPEAN PARLIAMENT**
DIRECTORATE-GENERAL FOR RESEARCH
Division for Budgetary Affairs, Civil Liberties and Internal Affairs, Rules of Procedure,
Petitions and Comparative Law
Tel.: 00 352 4300-22167
Fax : 00 352 4300-27723
E-mail: acomfort@europarl.eu.int

The opinions expressed in this document are the sole responsibility of the authors and do not necessarily represent the official position of the European Parliament.

Reproduction and translation for non-commercial purposes are authorized provided the source is acknowledged and the publisher is given prior notice and sent a copy.

An abridged edition available in all languages of the European Union under the number BUDG 102 A XX

Original in english

Manuscript completed in January 1999

CONTENTS

Pages

EXECUTIVE SUMMARY	i
1. INTRODUCTION	1
2. EVALUATION OF PUBLIC POLICY: CONCEPTUAL ISSUES AND OPERATIONAL APPROACHES	2
3. CURRENT PRACTICE OF CAP EVALUATION AT THE LEVEL OF THE UNION	6
3.1. CAP Evaluation in the European Commission	6
3.2. CAP Evaluation in the European Court of Auditors	14
4. CURRENT PRACTICE OF CAP EVALUATION IN THE MEMBER STATES	17
4.1. Germany	17
4.2. The Netherlands	21
4.3. The United Kingdom	24
4.4. Sweden	30
5. DEFINING THE OBJECTIVES AND COSTS OF THE CAP	34
5.1. Objectives of the CAP	34
5.2. The Nature of Agricultural Policy Costs	53
6. TOWARDS A SYSTEM OF INDICATORS OF EFFECTIVENESS OF THE CAP	58
6.1. Indicators for a Strategic Evaluation of the CAP	59
6.2. Indicators for Evaluating Changes to the CAP: The Example of the MacSharry Reform	67
7. METHODOLOGICAL ISSUES IN EVALUATING THE CAP	71
7.1. Determining Policy Effects	72
7.2. Quantitative Versus Qualitative Approaches	78
7.3. Methods of Quantitative Analysis	80
7.4. Treatment of Short Term and External Factors	86
8. RECOMMENDATIONS: TOWARDS IMPLEMENTING A COMPREHENSIVE SYSTEM OF CAP EVALUATION	87
9. CONCLUSIONS	94
REFERENCES	97
OFFICIALS INTERVIEWED	103

* * *

EXECUTIVE SUMMARY

1. This study discusses the **purpose and methodology of evaluation in regard to the Common Agricultural Policy (CAP)**, with an emphasis on evaluation of the Common Organisations of the Markets. It has been requested by the Committee on Budgetary Control of the European Parliament. Evaluation has to deal with all objectives and costs of the policy concerned. As costs include not only budgetary expenditure but also other elements (for example consumer burden), the study goes beyond an evaluation of agricultural expenditure from the Union budget and adopts a broad perspective of policy objectives and costs. The aim of the study is not to engage in an actual evaluation of the CAP, but to consider conceptual, methodological and institutional issues that arise in this context.
2. The **conceptual issues and operational approaches** relevant for this study are closely related to the fact that the **evaluation of public policy** is based on the notion of government efficiency. This issue has gained importance as a consequence of the limited scope for tax increase in OECD countries since the 1980s. Evaluation analyses the way in which public policy pursues its objectives. Hence, one of the principal requirements of evaluation is to have a clear policy objective. Where this is not the case, evaluation will always be ambiguous. Much of the literature on evaluation and the auditing of public policy refers to the "Three Es", i.e. economy, efficiency and effectiveness. Among them, effectiveness is the term which is of most importance to this study. Unfortunately, though, it also causes major problems of definition.

Evaluations are important at every stage of policy implementation, *ex ante*, ongoing and *ex post*. Evaluations should be based on certain criteria using a set of indicators relevant to the particular policy. Not all policies can be evaluated under the same assumptions and with the same methods. In different parts of Europe there are different traditions in both policy making and policy evaluation. Moreover, each of the fifteen Member States has different opinions and priorities. This may create problems if evaluations at EU level are not considered by some Member States as having been performed under acceptable criteria.

- 3.1. **Evaluation of the CAP in the European Commission**, defined in a broad sense, is a complex process, involving various parts of the Commission in different responsibilities, and taking a wide variety of forms. The state of affairs as it stood in 1995 was described in a study for DG XIX by Levy (1995) which provides a large amount of information. At that time, the Commission had not yet established a formal process of evaluating EU policies. Meanwhile the situation has changed noticeably, through the introduction of a new process of evaluating all EU expenditure programmes in the Commission, under the "Sound and Efficient Management 2000" initiative, known as SEM 2000. In this context, DG VI has established a new unit for CAP evaluation (DG VI.A.I.4) which became fully operational in February 1998. At the time of writing none of the evaluations initiated by the new unit had yet been completed, but the approaches adopted hold the promise of generating useful results. It is appropriate that ample involvement of operational staff in DG VI in the evaluation exercises is envisaged. However, the staff of the new evaluation unit should probably be increased.

Independent from formal evaluation procedures, evaluation of the CAP has a long tradition in the Commission, based particularly on analytical work done in the Studies Unit of DG VI, but also on contributions from DG II and on studies commissioned from external experts. In addition, there is a large number of independent academic analyses of the cost-effectiveness of the CAP. The present study does not review all this work. However, it is argued that the effective influence of the formal and informal CAP evaluation that has been provided in the past is evidenced by the sequence of bold reform proposals made by the Commission since 1991, starting with the MacSharry reform and continuing in the Agenda 2000 proposals.

- 3.2. **CAP evaluation in the European Court of Auditors** mostly falls in the category of auditing financial management and technical implementation of agricultural measures and expenditure, in particular as the Court's Annual Reports are concerned. However, the Court also occasionally presents assessments that go in the direction of more general policy evaluation. In its most recent Opinion on the Commission's CAP reform proposals under Agenda 2000 the Court has adopted a fairly general political perspective. If the Court's assessments go beyond pure audits, they are particularly strong where they build on the information regarding CAP implementation collected by the Court in Member States' administrations and companies. The Court's economic arguments are not always convincing, and its work on the CAP does not, and is not expected to, fulfil the need for full-scale policy evaluation in the general sense of the term.
- 4.1. **The current practice of CAP evaluation in Germany** is not based on a formal evaluation process like that now established in the Commission. However, important policy measures are regularly reviewed, by the Planning Group of the Federal Ministry of Food, Agriculture and Forestry, which does the conceptual groundwork for policy decisions to be taken in the Ministry. The Planning Group produces documents that assess the merits and drawbacks of individual measures. There is no provision which would make sure that all measures are reviewed on a regular basis, in a more or less standardised format. Only loosely related to evaluation in a formal sense are the activities of the Scientific Advisory Council of the Federal Ministry of Agriculture, which provides independent advice to the Minister. The Council's reports, generally arguing for more market orientation in agricultural policy, are occasionally hotly debated in the general public, not the least because they often deviate fundamentally from the policy orientation of the Minister.
- 4.2. Until recently, **CAP evaluation in the Netherlands** also did not rely on a formalised routine process. However, from time to time analysts in the Dutch Ministry of Agriculture produced reports and commissioned external studies on strategic issues in agricultural policy, and the Ministry used the results in the preparation of policy decisions. An area that was studied with particular emphasis in recent years is the relationship between the CAP and the environment. When the Dutch Minister of Agriculture presents his draft budget to the Parliament, the document submitted also has one chapter devoted to the evaluation of existing policies, typically drawing on the results of *ad hoc* studies. More recently, the Dutch government has embarked on the process of establishing routine policy evaluation. In 1998, the Ministry of Agriculture began to plan its evaluation activities and envisages to programme the policy evaluation process on a yearly basis starting in 1999

The intention is that all policy areas are evaluated at least once every five years. Moreover, an *ex ante* evaluation is foreseen for all newly designed policies. Both domestic policies pursued in the Netherlands and agricultural policies administered at the level of the Union are potential candidates for evaluation.

- 4.3. The current systematic and formal process of **policy evaluation in the United Kingdom** was established in 1988. In the UK, appraisal and evaluation are seen as important stages in the policy cycle. Evaluation of agricultural policies is one of the activities of the UK National Audit Office (NAO). The NAO has a team of five full time staff engaged in investigations of agriculture expenditure, including that under CAP schemes administered by the Ministry of Agriculture, Fisheries and Food (MAFF). Since 1988 the NAO has published 15 reports on agriculture, on subjects as different as fraud prevention to the protection of environmentally sensitive areas. NAO nearly always employs advisory panels whose members are drawn from different relevant areas. These panels normally consult trade and other representative associations. Currently all published reports are subject to quality reviews by the London School of Economics.

The UK MAFF had a systematic rolling programme of *ex post* evaluation of specific programmes under the CAP (and UK policy instruments) for the last ten years. These evaluations are commissioned by MAFF from independent research institutions following a system of competitive tenders. The use of external organisations for this evaluation is justified on the grounds that it provides an independent and objective analysis of policies. The MAFF quite explicitly include in their evaluation programme EU policies which have significance for UK interests. There is clear reference to possible divergence between EU and UK objectives for policies. When evaluations are completed, the reports are published (usually nowadays on the internet) and frequently they feed into a formal policy review procedure. *Ex ante* evaluation is conducted on a less formal basis and done mostly in-house. It is seen as an integral part of the on-going debate on the CAP.

- 4.4. **Agricultural policy evaluation in Sweden** has played an important role in the extraordinary series of switches in agricultural policy in the last decade. Evaluation involves the Swedish National Audit Office (RRV) and the Ministry of Agriculture, but also agricultural interest groups and independent researchers. The RRV does not have a department specifically working on agriculture, but has units which for a period of time concentrate on specific government departments and policies. In 1997 the Swedish National Audit Office joined a parallel audit with the UK NAO and the Dutch auditing body of the Arable Area Payments scheme which was introduced by the 1992 Reform of the CAP. Work on more detailed studies was to be conducted by staff of the RRV, with help from external consultants, especially on some of the modelling required.

In the Swedish Ministry of Agriculture, the amount of internal resources available for evaluations is still very small, but this is a consequence of the structure of the Swedish public sector and reliance on agencies and committees. During the past years the ambitions to perform long-term CAP analysis is enhanced and co-ordinated by the Ministry, which organises and participates in evaluation exercises of the CAP. For large evaluations a committee is set up, consisting of a wide selection of representatives from different relevant organisations. Usually the committee consists of members from the Swedish Board of Agriculture, from the Parliament, researchers, representatives of trade and farmers unions, as well as members from other governmental and private organisations.

The Swedish Ministry of Agriculture has been deeply involved in a series of studies on the Common Agricultural Policy, an exercise that began with the EU accession negotiations and has been reinforced since accession.

The research is performed with the aim of clarifying the Swedish position on the CAP. The strong international orientation is particularly interesting. What can be concluded is that Sweden is one of the EU members at the forefront of CAP evaluation.

- 5.1. **Defining the Objectives of the CAP** is less straightforward than it might appear at first glance. Moreover, decisions on the CAP are taken in a political process in which the participants have their own political motivations and objectives. Hence, whilst formal evaluation of the CAP can only proceed on the basis of stated, agreed objectives, it should not be too surprising to find that the results of such evaluations do not always appear to have a very strong connection with observed policy decisions.

Analysis of CAP objectives is complicated by the fact that other EU policies and their objectives may be in conflict with the CAP, in particular macro-economic policies, policies regarding international trade and external relations, and environmental policies.

The instruments of the CAP, though supposed, *inter alia*, to stimulate agricultural productivity improvement (defined in terms of technical efficiency), have in reality focussed on stabilising and raising prices of farm products. They have thus tended to contradict the achievement of economic efficiency. Ensuring a fair standard is an objective that has multiple interpretations. Defining 'fair', 'standards of living' and 'persons engaged in agriculture' each poses difficulty, and different Member States have different perspectives of the resulting policy implications. The objective of market stability is often confused with that of raising the level of prices. Assuring the availability of supplies does not translate easily into an unambiguous policy. The objective to ensure that supplies reach consumers at reasonable prices is clearly in conflict with other CAP objectives.

The relative weights attached to the objectives of the CAP and their interpretation have changed over time. Improving technical productivity is now seen as less important than enhancing the competitiveness of EU agriculture. The change in policy instruments pursued since 1992 has significantly changed the terms of the debate about fair standards of living in agriculture. The emphasis given to market stability has declined in the process of integrating EU markets more fully into international trade. Significantly more weight is now given to environmental protection and rural development, though these objectives still lack clarity. New, quite explicit, objectives of the CAP were formulated with the goal to reduce problems caused in pursuit of the original objectives. Improving market balance is one such second-order objective of the CAP. Another one is redistribution of support towards more vulnerable holdings. Because of such changes, the real objectives of the CAP have become even less clear. As direct payments are now an important part of the CAP, their longer-term objectives, beyond immediate compensation for price cuts, should be defined more clearly.

- 5.2. **In assessing the nature of agricultural policy costs**, it is important to go beyond budget expenditure and to include all forms of policy costs. More specifically, agricultural policy costs include economic or resource costs, budget costs, international costs, and environmental costs.

An important element of the economic or resource costs of agricultural policy are the extra resources employed in agriculture as a consequence of policy interventions. The income that these resources could earn in other sectors of the economy is an element of agricultural policy cost. Price support places a burden on consumers, which also has to be included in an assessment of agricultural policy costs. Other costs of agricultural policy include the implications for the structure of the rest of the economy and hence for GDP growth, as well as for macro-economic variables such as the trade balance and fiscal deficit.

The budgetary costs of agricultural policy are relatively easily defined and measured. However, they may go beyond expenditure institutionally defined as expenditure on the CAP. In particular, administrative expenditure and expenditure by the Member States may have to be included. An increase in budget expenditure resulting from a policy change does not necessarily indicate an overall increase in policy costs if it is offset by savings in other areas, in particular savings of economic and resource costs such as consumer burden and distortions of resource use. Indeed, improving the (overall) cost-effectiveness of agricultural policy may require policy changes (such as the substitution of decoupled payments for price support) that involve larger budget expenditure.

The international costs of the CAP result from the fact that the EU is a large player in international agricultural trade. Price support in the EU contributes to depressing world market prices for agricultural commodities. For the EU this results in larger expenditure on export restitutions, and exporting third countries are hurt. This is relevant not only because the EU has a responsibility for global developments, but also because it influences the Union's position in international trade talks and has feedback effects on the pursuit of the CAP.

The environmental costs of the CAP are not easily measured. However, their growing significance is evidenced by the fact that environmental concerns have become an important driving force for changes in the CAP.

- 6.1. For practical policy evaluation, empirically observable indicators are required which allow to determine the extent to which the policy has achieved its objectives and caused costs. **A number of indicators for a strategic evaluation of the CAP can be suggested.**

The improvement of productivity can be measured by various indicators of technical and economic productivity, at both the farm and the sectoral level. For measuring the standard of living in agriculture, an aggregate indicator is net value added per annual work unit. Farming incomes can be derived across the Union from the Farm Accountancy Data Network (FADN), including data on the distribution of income between farms of different sizes, types and regions.

Measuring market stability requires indicators of price and income variability and comparative analysis of market stability inside and outside Europe. To evaluate the ability of 'normally' viable farms to survive under extraordinary conditions, indicators such as the frequency of insolvency, bankruptcy or other involuntary changes in the status of farmers can be used.

To discover how well the objective of reasonable consumer prices is reached requires indicators of relative prices at all stages of the food chain from the farm gate through to the consumer. These can be developed for international comparisons or for analysis of the actual policy compared to alternative price policies under the CAP.

Indicators to measure the extent to which competitiveness of EU agriculture is improving include nominal and effective rates of protection, and producer subsidy equivalents (PSEs). Cost comparisons between producers in Europe and those in other parts of the world can also be made, though they involve considerable practical and conceptual problems. In particular, a comparison of costs between heavily protected and unprotected producers does not provide information about the long run competitiveness after input prices and asset values have had time to adjust.

There is a long way to go to develop practicably operational indicators for environmental objectives. Indicators would be required for resource protection for soil, water and atmosphere; of the extent and health of natural and semi-natural habitats; of bio-diversity on both farmed land and land surrounding agriculture; landscape features and landscape change; historical and heritage features. Indicators of rural development include demographic features, the employment structure, income levels, access to transport, service provision, and social, education and health attainment indicators.

Budgetary and economic or resource costs can be measured in a joint conceptual framework based on welfare economics. Producer benefits and consumer burden are measured through producer and consumer rents. The burden on taxpayers is equivalent to net budget expenditure. The balance of these effects provides an estimate of the (static) loss in overall economic welfare. This "true" economic cost of agricultural policy is usually only a fraction of the budget expenditure involved, though the size of that fraction depends on the instrumentation of the policy.

- 6.2. The present study also presents a concrete list of **indicators for evaluating changes to the CAP for the example of the MacSharry reform**. These indicators are suggested to measure the improvement in market balance; the effects on farm incomes and their distribution; the impact on public expenditure; the effect on overall economic welfare; the implications for international trade; and the impact on the environment and rural development.

The many indicators that can be used to evaluate the outcome of a policy change such as the MacSharry reform cannot be formally aggregated into one single indicator of success or failure. Aggregation has to occur in a very informal, subjective way in the brain of the policy maker or the evaluator. However, there are some subsets of indicators which can be aggregated, into measures such as cost-benefit ratios or cost-effectiveness. If aggregation of the overall set of indicators is considered an absolute must, a crude approach based on subjective rankings on a point scale and weights for individual indicators can be suggested.

- 7.1. Among the **methodological issues in evaluating the CAP**, a fundamental one is **how to determine policy effects**. The problem is that the effects of the policy concerned must be filtered out from the effects of the many other factors that shape reality. For example, cereals use in the EU increased by around 18 mio tons between 1991/92 and 1996/97.

However, it cannot be said that this was the effect of the MacSharry reform, which may have been larger or smaller, depending on the joint impact of all the other things that have happened in the same period.

Indeed, looking at observed changes over time cannot only overestimate or underestimate the actual effect of the policy concerned, it can even yield a wrong impression of the direction of the policy effect. For example, the observed decline in EU beef consumption after the MacSharry reform was certainly not the effect of that policy change, but a result of, among others, the BSE crisis.

The only logically sound way to solve this problem is to construct a counterfactual reference situation that would have prevailed under an alternative policy. Because this reference situation is only hypothetical, there is often considerable resistance to using this approach in the practice of policy evaluation. Indeed, the single most frequently committed mistake in policy evaluations is negligence in constructing an appropriate counterfactual reference situation (and a number of examples are cited in the present study). In *ex post* policy evaluations, the before-after approach is often applied, in spite of its logical and empirical deficiency. In *ex ante* evaluations, on the other hand, there is less resistance against constructing counterfactuals because the future is hypothetical anyhow.

There are various ways to construct a reference situation, ranging from the use of econometric models to simple guesswork. The appropriate approach depends on the type of analysis and the resources available. However, the evaluator should always make it explicit which reference situation was used in the evaluation. This also includes the requirement to specify the policy that is assumed to be pursued in that reference situation.

- 7.2. The choice of **quantitative versus qualitative approaches** in policy evaluations depends on the issues to be analysed. Where only the direction of the policy effect is of interest, or where the objectives pursued cannot be quantified, a qualitative analysis is appropriate. However, in most cases policy evaluations will have to involve quantitative analysis.
- 7.3. The **methods of quantitative analysis** include a wide variety of approaches, ranging from simple accounting frameworks to fully behavioural models. Among the latter, there is a wide choice of both partial and general equilibrium models, with or without econometrically estimated parameters. The DG VI study on the impact of the Agenda 2000 proposals published in October 1998 provides a good example of the use of different types of quantitative models in the analysis of agricultural policy changes.
- 7.4. In *ex post* evaluations, the appropriate **treatment of short term and external factors** is important in order to determine the systematic influence of the policy concerned. The major approaches available are use of quantitative models and averaging of data.
8. In the present study, a number of **recommendations** are made for **implementing a comprehensive system of CAP evaluation**. Because some of them are political in nature, and potentially sensitive, they should be interpreted as a starting point for discussion rather than a blueprint for a holistic regime.

Four different types of CAP evaluation should be distinguished, i.e. evaluation of A) the strategic orientation of the CAP; B) significant changes in the pre-existing policy regime; C) specific ongoing programmes; D) operational implementation of individual policy measures.

Evaluations of type A) should be done around every five years, and mainly with an *ex post* perspective, but also including some *ex ante* assessment of policy alternatives. Type B) evaluations should be done *ex ante* whenever policy changes are considered, and definitely before they are proposed by the Commission. Around three years after the policy change there should also be an *ex post* evaluation. Evaluations of types C) and D) should be done every three years, on an *ex post* basis.

Strategic evaluations (type A), which are supposed to address the fundamental justification of policies, are best done by external experts. The Commission should have funds earmarked for commissioning such evaluations. Evaluations of policy changes (type B) and of specific ongoing programmes (type C) should be done in-house in the Commission. From time to time the in-house evaluations provided in the past should be assessed by outside experts, to make sure that the approaches adopted and results achieved are not too much constrained by considerations regarding political "feasibility" of alternative policies. The Commission should have funds available for that purpose. Operational evaluation of the administration of policy measures to make sure they are cost-effective and safe from corruption (type D) is the domain of the European Court of Auditors.

Evaluation of Union policies such as market and price policies under the CAP should be the responsibility of the Commission. The results of these evaluations should, though, be available to the Member States, and the Member States should be provided an opportunity to comment on the results of these evaluations, on a regular basis and in a structured process. However, it is desirable that the Member States also engage in evaluations of these policies, such that an atmosphere of "competitive evaluation" is generated.

The European Parliament should have access to the results of all CAP evaluations. Moreover, the European Parliament should have the right to request specific evaluations to be done either by the Commission or by outside experts. For this purpose, it is suggested that a given share of the total budget made available for CAP evaluation (say, ten per cent of that sum) is earmarked for evaluations specifically requested by the European Parliament.

In its proposals for regulations and directives, the Commission should in future specify objectives more clearly than is currently the case in the *whereas* clauses. Moreover, all policy proposals should in future be accompanied by a list of indicators to be used in later evaluation the policy concerned. Commission proposals for policy changes (type B) should be accompanied by the results of prior *ex ante* evaluation, and Commission proposals for changes of policies falling in category C should also explicitly make reference to results of *ex post* evaluations of the policies concerned.

In policy evaluations, both *ex post* and *ex ante*, one element should always be an explicit description of what the reference situation is meant to be, i.e. which policy is the benchmark and which outcome would have been or is expected to prevail under that benchmark policy.

The EU should invest considerably more resources than in the past into the analysis of agricultural policies and related issues, both in a general sense and in the context of formal policy evaluation. An explicit decision should be taken on the magnitude of resources to be made available for CAP evaluation. As a starting point for discussion, a magnitude of one half of one tenth of one per cent of EU expenditure on the CAP is suggested as the budget to be made available for CAP evaluation (around Euro 20 million).

9. In **conclusion**, the picture emerging is that a lot of work has always been done in the broad area of analysing and evaluating the CAP. The accelerating speed of CAP reform to an extent mirrors the results of these efforts. However, compared to what is done in other parts of the world in terms of laying the analytical groundwork for agricultural policies, the EU should intensify its work in this area. Moreover, the formal process of CAP evaluation can be improved in a number of regards. A public policy that consumes budgetary and other resources at a level found in the CAP should be well justified and thoroughly scrutinised. In that context, evaluation of the CAP can make a significant contribution.

* * *

1. Introduction

The present study discusses the purpose and methodology of evaluating the Common Agricultural Policy (CAP) of the European Union. It has been requested by the Committee on Budgetary Control of the European Parliament and is, as specified the terms of references for this study, intended to assist the members of the Committee in their discussions on the control of expenditure from the EU budget in the field of agriculture, especially in the light of current discussions on the reform of the CAP. The study is expected to emphasise evaluation in regard to Common Organisations of the Market, rather than structural spending. It is not supposed to engage in an actual evaluation of the CAP, but to review the nature of any evaluation systems currently in place, and the purpose and methodology of evaluation in this sector and the potential impact on the question of evaluation of the reforms to the CAP.

Policy evaluation stems from the notion of government efficiency. It has always been the case that once Government takes on to itself the role of providing goods or services, or steering or regulating the private sector that delivers such goods and services, it is reasonable to inquire how well the government policy is performing. Policy evaluation, in a general sense, is therefore as old as policy-making. However, public policies have become increasingly complex, and operate in an increasingly complex environment. Evaluation of public policies is, therefore, an equally complex business, requiring specific methodology and expertise. Moreover, the limited scope for tax increase and a growing awareness of the necessity of reducing the deficit in public budgets have given rise to calls for more efficiency in government action. The demand for sound and comprehensive policy evaluation has therefore grown strongly in the recent past, and has triggered the establishment of formal evaluation procedures and related institutional arrangements in many governments and international institutions.

Policies pursued by the European Union in general, and the CAP in particular, are no exception to this trend. Analysis of the functioning of the CAP has been provided, by various authors and institutions, inside and outside the European Commission, since that policy was introduced in the early 1960s. However, a formal process of evaluating the CAP on a regular basis has been set in motion only recently, in the context of a general process of evaluating EU expenditure programmes now established in the European Commission, under the "Sound and Efficient Management 2000" initiative, known as SEM 2000. Given that roughly one half of EU expenditure is used for the CAP it is imperative that CAP evaluation becomes an integral element in the process of planning and implementing EU policies for agriculture. It is therefore timely to consider the purpose and methodology of CAP evaluation. This is what the present study attempts to do.

To provide some general background, the study begins with a brief survey of the concept of policy evaluation (Chapter 2). The study then reviews the current practice of CAP evaluation, both at the level of the European Union (Chapter 3) and in four selected Member States, i.e. Germany, the Netherlands, the United Kingdom and Sweden (Chapter 4). One fundamental requirement of policy evaluation is the definition of the objectives the policy is supposed to achieve, and measurement of the costs involved in pursuing the policy. Chapter 5, therefore, provides an in-depth discussion of the objectives and costs of the CAP.

It explains, among others, why the costs of the CAP transcend budget expenditure and include items such as consumer burden and resource costs. CAP evaluation would, therefore, not provide a full picture if the assessment of its cost-effectiveness were limited to budgetary expenditure. This is why this study adopts a broad perspective of all objectives and costs of the CAP.

Once the objectives and costs of the CAP have been defined in a general way, they have to be translated into empirically observable indicators that can be used in practical evaluation exercises. This is discussed in Chapter 6, both in general form and, to provide an illustration of how this can be done in practice, in regard to a potential evaluation of the MacSharry reform. The analysis of how the CAP and its individual instruments affect such indicators poses a number of methodological issues, which are discussed in Chapter 7. After all this has been considered, a number of issues emerge that should be considered in future efforts to progress towards implementing a comprehensive system of CAP evaluation. Such recommendations are presented in Chapter 8, before conclusions are drawn in Chapter 9¹.

2. Evaluation of Public Policy: Conceptual Issues and Operational Approaches

Policy evaluation is based on the notion of government efficiency. This issue has gained importance as a consequence of the limited scope for tax increase in OECD countries since the 1980s (Oxley and Martin, 1991). There has been a clear move to improve the efficiency of the public sector in an attempt to curb the waste of resources. This has led to widespread privatisation programmes and the downsizing of the public sector, the tightening of budgetary resources for government departments and the creation of internal markets² in the health and education sectors (*ibid.*). From the literature on public policy performance measures it is reasonable to claim that Britain is amongst the leaders in the evaluation of public policies (Carter, 1991). The overriding concern of the Thatcher administration was to control expenditure by minimising the administrative cost and waste of the policies. The intention was to introduce management control and product quality in government. These concepts were common in private enterprises, but scarcely heard in the public sector until then. A systematic system of public policy and performance evaluation emerged (Bovaird and Gregory, 1996). New more up-to-date efficiency indicators were introduced, and the UK's National Audit Office (NAO) soon began using the measures and causing political discomfort to government departments (Gray *et al.*, 1992).

A number of other countries also began to improve their management and policy auditing offices, either as a consequence of the British success and/or because of the realisation that their public finances were reaching the limit. The pressures did not spare the European Union's budget, on the contrary, it created a strong demand for the Union's official bodies to introduce monitoring and evaluation techniques. Member states have called for a system of evaluating the benefits and costs of EU regulations in the Maastricht and Amsterdam treaties. Definitions, concepts, indicators and practices in evaluation multiplied as well as the literature on this subject. The basic definitions, needs and intentions of evaluations, however, are similar and will be described here.

¹ We are grateful to Jorge Nuñez Ferrer for research assistance provided in working on this study.

² Introducing competition between the institutions.

One of the principal requirements needed to perform a successful evaluation is to have a clear policy objective or target. This simple, crucial and reasonable requirement is often absent for public policies. Agricultural policy objectives are discussed in Chapter 5 where their vagueness, generality and conflicting interpretations are pointed out. Clear objectives are so crucial that when discussing the problems of performance indicators, Grainer (1996) quotes the US Congressional Budget Office's warning: "If a program's objectives cannot be determined, performance measures will always be ambiguous, if not superfluous" (Congressional Budget Office, 1995).

Much of the literature on evaluation and the auditing of public policy refers to the "Three Es", i.e. economy, efficiency and effectiveness. For example, they are explicitly referred to by the UK and Swedish National Audit Offices as the central concerns in trying to get value-for-money from the administration of public policy.

Economy refers to the need to minimise the use of resources in achieving the objectives of a policy.

Efficiency is directed towards the management of the public administration and has prompted the gradual introduction of private sector managerial practices.

Effectiveness is the term which will be of most importance to this research and which causes the most problems of definition. In the UK alone the definitions used by different government and private auditing agencies differ in their emphasis. Gray *et al.* (1992, p. 603) lists five different ones, but finds that all share one common view of effectiveness 'the value given to the relationship between an activity and its effects' (p. 604). The biggest problem is to award value to a public policy. To understand the effectiveness it is necessary to understand the nature of the effects of the policy. Gray *et al.* (1992) categorise the effects in three groups: outputs, outcomes and impacts. Outputs are units of products and services provided. Outcomes are the direct measurable consequences of the particular policy on the stakeholders (target groups and interested parties). Impact assessment extends the analysis on the outputs by including net and gross changes in values due to the policy alone or in combination with other activities. It goes beyond the output research looking for indirect effects too. Gray *et al.* (1992) distinguish two types of effectiveness concepts: substantive and evaluative. Substantive effectiveness is further subdivided into managerial and policy effectiveness. These are concerned with the relationship between outputs and inputs of policy. Managerial effectiveness is concerned with the effective management of the economy in terms of its efficient use of labour, materials, skills and funds. This is the aspect dealing with the input effectiveness of a programme, project or a policy. Policy effectiveness is concerned with the outcome and impacts the policy produces. Government auditors generally avoid this latter type of analysis because of the difficulty of assessing the value of the outcomes or impacts of certain types of policy, in particular social policies that have political importance and whose outcomes are not easily quantified. Evaluative effectiveness refers to the capability of a programme, project or policy to measure its own effectiveness and improve it. Managerial freedom and internal audit systems can enhance evaluative effectiveness.

Different authorities interpret effectiveness in different ways. Three examples can illustrate this: Effectiveness is the "extent to which the objectives of a policy are achieved. The most effective policy is the one which achieves all its objectives" (Treasury, 1988).

"Effectiveness is concerned with the relationship between the intended results and the actual results of projects, programmes or other activities. How successfully do outputs of goods, services or other results achieve policy objectives, operational goals and other intended effects?" (National Audit Office n.d., p. 5).

Tomkins (1987, p. 49) suggests a broader definition of policy effectiveness, as the "value which society desires from given inputs and outputs of a particular public service".

Despite the attempts in the literature of public administration to define generally accepted concepts and terms in the field of policy evaluation, there is no universal acceptance of one set of such terms. Thus, for example, in the European Commission's own document summarising best practice in evaluation across the Commission, it is suggested to use just two Es. Effectiveness is defined as the degree of achievement of a programme's or project's objectives (either in terms of output or outcome), and efficiency is defined as the productivity of an intervention given the resources and means used (European Commission, 1996, Annex 1.2). No authors have yet succeeded in imposing standard definitions of these ideas and there is no attempt in this report to do so.

Evaluations are important at every stage of a policy implementation. *Ex ante* evaluations should be performed on proposed policies to understand what are their needs, costs and the future likely impacts. Anticipating costs and benefits, and particularly assessing the risk of failure is very important aspect of policy analysis. *On-going* evaluations are also useful to control the progress of the policy implementation and to understand if there are any new factors needing redressing. *Ex post* evaluations are important to check the output and impact effectiveness of the policies. This type of analysis is particularly helpful if it follows the same structure and can be directly linked to an *ex ante* analysis. It is rare, however, to find government departments which follow a consistent and linked system of *ex ante*, *ex post* and on-going evaluations.

Evaluations should be based on certain criteria using a set of indicators relevant for the particular policy. Not all policies can be evaluated under the same assumptions and with the same methods. It is important to construct a proper hierarchy of objectives and select the reference (a scenario in the absence of intervention) against which a policy is evaluated. The hierarchy of objectives determines which impacts of the policy are more relevant. A policy might seem wasteful in economic terms, yet there may not be a politically acceptable alternative, as its existence may be important for reasons as varied as national welfare and security, equality, social and political stability. This is why conventional evaluative tools like a Cost-Benefit analysis (CBA) may be inappropriate (Pelkmans and Labory, 1998). Such problems are particularly acute for the evaluation of European policies. How can common EU policies be evaluated? What are the underlying criteria and the hierarchy of objectives? The CAP is a multiple objective policy, which one is the most crucial? All depends on the point of view and the fifteen Member States may have different opinions and priorities.

OECD (1994) indicates that the performance indicators in different countries are based on different guiding principles. Auditing methods and the underlying philosophy of policy effectiveness can and do differ across countries. Buckaert (1996) lists important causes for the heterogeneity of systems. In all cases the base for the discrepancies originates in the different historical developments of the state. In his view, there are fundamental differences in the underlying values about the role of the state

as a provider of services and implicitly in the criteria on which evaluations of the policies should be based. In Europe, Buckaert distinguishes three broad public policy traditions: the Anglo-Saxon, the Latin and the Eastern European tradition (the last has to be taken into account in view of the future EU membership of Central and Eastern European Countries).

Buckaert argues that in countries under the Latin tradition many services are considered as a necessary provision by the state. The citizen has a right to enjoy these services and the three Es should not constitute the main basis for their existence, whereas in Britain the government movement towards private-sector style management systems has changed the status of the citizens in the direction of government "clients". In this view, effectiveness and efficiency in the public sector will be fostered by competitiveness and free market style systems. A service not able to survive under this system may then be discontinued. This behaviour and the philosophy underlying it, prompted Buckaert (p. 231) to suggest that the British Citizens Charter should be renamed the "Clients Charter". This is not considered acceptable in a number of other European countries where delivering public services under market conditions is simply unacceptable and considered a violation of the citizen's basic rights.

For the Eastern European countries, the central planning tradition has left a different legacy in the approach of the state on public policy monitoring. "(...), there is a tendency in Western Europe to focus on efficiency and effectiveness and how to control this, whereas in Eastern Europe was (is) more a systematic tendency to focus on control and planning, and how to do this in an efficient and effective way." (Buckaert, 1996, p. 231)

Another important point stressed by Buckaert is the meaning of subsidiarity as a criterion to allocate competencies to levels and institutions at lower hierarchical levels if possible and higher if necessary. A problem may arise when subsidiarity is interpreted differently by the Member States. The interpretations on the level of competencies of the EU as regulator or evaluator may differ. This may create problems if evaluations at EU level are not considered by some Member States as having been performed under acceptable criteria. An attempt has been made by Pelkmans and Labory (1998) to give guidance on how to interpret subsidiarity and at what level regulations should be implemented, monitored and evaluated. However, there is no universally accepted formula for determining the correct degree of subsidiarity and the criteria on which to base the analysis.

These few remarks on evaluation serve merely to indicate to the reader that there is a large literature in the field of public policy evaluation. The bulk of this report is now focused on why and how to conduct evaluation in the field of agricultural policy. This report deliberately steers a different course through the subject than taken by Prof. Levy in his 1995 report on the same subject conducted for DG XIX of the European Commission. Prof. Levy is an expert on Public Administration and this is reflected in his approach. The authors of this report are agricultural economists specialising on the CAP and therefore the approach here emphasises the economic impacts and costs of the Common Agricultural Policy. Coming three years later than the Levy report we are able to reflect further developments both in the policy itself and in the evaluation activity in the field of agriculture.

3. Current Practice of CAP Evaluation at the Level of the Union

3.1. CAP Evaluation in the European Commission

In a broad sense, evaluation of the CAP in the European Commission is a complex process, involving various parts of the Commission in different responsibilities, and taking a wide variety of forms. It ranges all the way from day-to-day monitoring of the effects of market interventions, through *ad hoc* analysis of individual policy measures and more formal evaluation of measures on a regular basis, to quasi-academic studies into the more strategic aspects of the longer-run orientation of the CAP. It is impossible in this study to describe and assess this process comprehensively and in detail. As a matter of fact, the state of affairs in evaluating the CAP in the European Commission, as it stood in 1995, was described in a study for DG XIX by Levy (1995). That study provides a large amount of useful factual information on what is done where and how in the Commission, as far as evaluation of the CAP is concerned.

However, when Levy did his study in 1995, the Commission did not yet have a formal procedure of evaluating EU policies. Levy hence concentrated his assessment of internal Commission evaluation of the CAP mainly on work done in the Studies Unit of DG VI. That Unit (now DG VI A1) is still engaged in major analytical work on EU agricultural policies and markets and provides, as a "think tank", important input into the design of reforms of the CAP. There is no doubt that much of this work also contributes to policy evaluation in a general sense. However, as far as formal evaluation procedures are concerned, the situation has changed noticeably since the Levy study, through the introduction of a new process of evaluating all EU expenditure programmes in the Commission, under the "Sound and Efficient Management 2000" initiative, known as SEM 2000.

Under SEM 2000, responding among others to Article 2 of the Financial Regulation and Declaration 18 of the Treaty on European Union which signalled the Commission's responsibility for strengthening its systems for evaluating Community legislation, the Commission has established a fully-fledged procedure for evaluating EU expenditure programmes³. The design of this procedure was first outlined in a Communication to the Commission by Commissioners Liikanen and Gradin, adopted by the Commission on 8 May 1996 (document SEC 96/659 final). This Communication, after reviewing the (then) present state of evaluation in the Commission, recommended concrete steps towards evaluation. It emphasised that operational services in the Commission, as opposed to DG XIX (responsible for budgets) and DG XX (responsible for financial control), have the responsibility to carry out evaluations of their programmes. It defined the goals of evaluation to be improving programming and management; accountability to the Budgetary Authority and public opinion at large; and resource allocation. It also made the point that "evaluation is not a substitute for the political process where wider considerations need to come into play as well. Evaluation merely serves to generate insights to enlighten the discussion preceding political decisions."⁴ On that basis, the Communication outlined the evaluation procedures to be adopted. Every operational service (i.e. DG) was expected to designate an official, sector or unit with lead

³ In Commission terminology, the CAP generally does not use programmes, but actions, measures or policies. However, as far as fundamental issues of evaluation are concerned it appears that this distinction is of only limited relevance.

⁴ European Commission (1996a), p. 9.

responsibility for evaluation, and to establish an annual evaluation plan, to be updated in mid-December of each year. The first plans were to be presented in mid-December 1996. Evaluations are required within a timescale corresponding with financial and legislative planning, in practice at least once every six years. Evaluation results are to be documented in reports, to be made available well before the adoption of the proposal or decision which the evaluation is meant to influence. Executive summaries of the reports are to be transmitted to DGs XIX and XX. Brief accounts of evaluation studies are expected to be submitted as an accompaniment to the annual revenue and expenditure account (analysis of budget execution), associated with Article 80 of the Financial Regulation. DG XIX has prime responsibility for assessing the evaluations, and DG XX for ensuring that evaluation and monitoring techniques are used properly. DG XIX was to establish a Network of Evaluators, to be a forum for the exchange of ideas and expertise, and a number of support instruments were to be established, including instructional manuals and training seminars. DGs XIX and XX were charged with the duty to provide an Annual Evaluation Review.

In order to improve the relationship between evaluation, decisionmaking and budgeting, the Communication suggested that operational services should submit an *ex ante* evaluation report when proposing a new legal basis for any action incurring expenditure from the general budget, setting out the logic of the proposed programme, its objectives, the outputs of activities to achieve the objectives, and the resources necessary. Where possible, the objective, results and outputs should be expressed in the form of objectively verifiable indicators. The findings of these *ex ante* evaluations should be reflected in the explanatory memorandum accompanying the draft legislative proposal. It was suggested that DG XIX will take account of the availability and quality of findings of evaluation reports when negotiating the budget with the operational services. The Communication acknowledged that this new evaluation procedure could only be implemented gradually, and had to be based on a learning-by-doing process.

The process of implementing the new SEM 2000 evaluation procedures is fully underway in the Commission. DG XIX/02 has produced a guide for *ex post* and intermediate evaluation in January 1997 (European Commission, 1997a), describing in detail, and usable by non-specialists, the concepts and methods of evaluation. In co-operation with the operational services, DG XIX has provided the first Annual Report of evaluation studies done in the years 1996 and 1997, in the context of submitting the preliminary draft budget for 1999 (Bilan d'Évaluation, Commission Européenne 1998). The operational services establish their evaluation plans and submit them to DG XIX, where they are compiled into one overall evaluation plan for the whole of the Commission.

As far as evaluation of the CAP is concerned, DG VI has established a new unit for just that purpose (DG VI.A.I.4). Staffing began in the middle of 1997 and was completed in February 1998. The unit has three professionals and an annual budget for external studies of 1 mio ECU.⁵ During the first months of its existence, the new unit has developed its conceptual approach and strategy, supported by an outside specialist consulting company. The Commission's Evaluation Plan for 1998 foresees 13 evaluation studies in the area of agricultural policy, relating to measures financed under FEOGA

⁵ In DG VI it is pointed out that their new evaluation unit is small compared to those in other DGs, for example DG XVI.

Guarantee.⁶ While 12 of these evaluations are to be done by DG VI, one is to be conducted by DG III, with support from DG VI. Among the three alternative types of evaluation considered by the Commission (*ex post*, intermediate, *ex ante*)⁷, three of the 13 agricultural policy evaluations planned for 1998 are *ex ante*, eight evaluations are intermediate, and two are listed as both intermediate and *ex ante*. Eight evaluations are to be done by external experts, while five are executed internally by DG VI staff. An internal DG VI document lists the budget expenditure foreseen for the 13 measures to be evaluated in 1998, showing that, in aggregate, they account for 72.1 per cent of total FEOGA Guarantee expenditure budgeted for 1998.

At the time of writing (fall 1998) none of the evaluations planned for 1998 had been completed. Hence no comments can be offered here on the outcome of the evaluations conducted by the new unit in DG VI. However, a few remarks may be appropriate regarding plans for some of the evaluations included in the plan for 1998. Of particular relevance to the policy domain to be covered in this study, i.e. the common market organisations under the CAP, and in an interesting way related to evaluation approaches that will be suggested in this study, are three *ex ante* evaluations closely related to the Commission proposals in Agenda 2000. In these three evaluation exercises, to be conducted in-house, DG VI deals with the new regimes proposed for (i) arable crops, (ii) beef and (iii) milk. The idea appears to be, fundamentally, to create the conceptual basis on which the results of the policy reforms now proposed by the Commission can be evaluated in the future.

For each of these three studies, a steering group has been established, comprising staff from operational units. As far as conceptual and methodological approaches are concerned, these groups receive, through the evaluation unit, advice from an external consulting company helping the new evaluation unit in DG VI to develop instruments and methodology and to manage the evaluation. For each of the three new market regimes proposed, the evaluation starts by considering in detail which objectives are being pursued. Important sources of information in this context are the "whereas" clauses in the preamble of the (proposed) market regimes and the explanatory memoranda and other official documents relating to Agenda 2000. The next step then is to consider the ways in which the individual measures to be pursued under the market regime can potentially affect the various objectives. For this purpose, a flow chart is generated that sketches the causal links through which individual measures can potentially influence the different objectives. Based on this analysis, it is planned to come up with a list of criteria and indicators necessary to evaluate, in the future, the results of the policy reforms. The process of working on these three studies is scheduled such that in an initial stage they are done on the basis of the Commission proposals, while in a second stage they will be adjusted to take account of the final decisions taken by the Council.

⁶ The evaluations scheduled for 1998 are summarised in the Evaluation Plan 1998, pp. 1-5, provided by DG XIX. According to information received from DG VI, in 1998 seven external and three internal *ex ante* evaluations were launched by DG VI and one external evaluation was commissioned by DG III. In addition to the 13 evaluations listed in the Commission plan under the heading of agricultural policy, an internal DG VI document lists seven evaluations of measures financed under FEOGA Guidance.

⁷ See the Communication, European Commission (1996a), p. 17.

The major purpose of these *ex ante* studies is not so much to evaluate the policy proposal tabled by the Commission, but to establish a firmer basis on which intermediate and *ex post* evaluations of the programmes concerned can be conducted at a later stage. A central product of these studies will therefore be the list of criteria and indicators to be used in future evaluations of the respective measures.

It is definitely too early to assess the usefulness of these three studies, as they are not yet completed. However, the basic approach adopted appears to hold promise. As this study will argue below, a clear definition of the objectives pursued by individual policies is a fundamental prerequisite of any evaluation. Moreover, in order to make evaluation practicable, well defined criteria and indicators are necessary that can be empirically observed, if not measured. These criteria and indicators must be derived from the objectives pursued, and they should reflect the mechanisms through which the measures pursued work in the reality of the economic activities concerned. To do this, it will be necessary to consider, in detail, the logic that connects the objectives and the concrete measures concerned, and to establish criteria and indicators for future evaluations on that basis. This seems to be exactly the approach envisaged in the three *ex ante* studies related to Agenda 2000, which therefore seems to be an appropriate approach to creating the basis for future evaluations. At the same time, as will also be argued below, one needs to consider carefully the methods to be used to filter from the many events that occur in reality those that are due to the policies pursued. If this is not done well, it is impossible to say whether anything that has happened in the past was a result of the policy measures adopted, or whether it might have happened in any case. It appears that this aspect is not yet fully covered in the approach now adopted by DG VI. For example, the *ex ante* evaluations now being pursued appear not to deal with the issue of whether any tools of quantitative analysis should be used to assess the extent to which any future changes in the criteria and indicators suggested can be attributed to the measures evaluated.

As manpower available for evaluation in DG VI is not sufficient to conduct all studies in-house, most evaluations will be provided by external experts. Each of these external evaluations is commissioned on the basis of a tender. A working group of DG VI staff involved in the measures concerned is established for each study. Detailed terms of reference are provided to guide the external evaluations. While the new formal evaluation procedures, as noticeable in the establishment of a separate evaluation unit in DG VI, have not yet yielded completed evaluation studies, earlier work in DG VI has resulted in documents that fall in the category of evaluation. For the two years of 1996 and 1997, these evaluations are listed in the Bilan d'Évaluation 1996/97.⁸ Of the nine evaluations listed under the heading of agricultural policy (FEOGA Guarantee), eight dealt with market promotion for various products. The one remaining study assessed the extent of overcompensation in the cereals sector. The study, done in the first half of 1997 and thought to help in the preparation of the Agenda 2000 proposals, compared wholesale prices actually observed in the first four years following the MacSharry reform with the market price foreseen by the Council when it decided that reform in 1992. The result, widely publicised and hotly debated among agricultural policy makers at the time, was that actual market prices had been significantly above prices foreseen. On that basis it was concluded that direct payments, set such that a full decline of market prices to the price level foreseen by the Council

⁸ Commission Européenne (1998). Evaluations regarding agricultural policy (FEOGA Guarantee) are listed on pp. 15-16, and their results are summarised on pp. 43-48.

would be compensated, were in reality largely overcompensating cereal growers. Based on that result the Commission proposed only partial compensation of the further reduction in the intervention price for cereals that would follow from the Agenda 2000 proposals. This study was a good example of a partial evaluation of one particular policy measure. It was partial only in the sense that only one objective (revenue compensation) was considered, at the aggregate level of the Union. A more comprehensive evaluation would have considered a broader range of objectives (e.g. including the effects on farm incomes, as opposed to revenue, and effects on the environment) and would have analysed the results at a more disaggregate level (e.g. individual Member States, different farming systems). It may also be noticed that the measure concerned was relatively easy to evaluate, at the aggregate level chosen, as the cause-effect links are straightforward in this case, and the objective (compensation of the price cut) was well defined. However, such comments do not question the value of evaluation exercises of this type.

In addition to such studies listed more recently under the heading of evaluation by the Commission, there is a long history of studies that have contributed to evaluating the CAP. As far as the role of the Commission is concerned, many of these studies have been done in-house, mainly in DG VI and its Studies Unit. Other studies were formally commissioned from external experts.⁹ The nature of such studies, the process through which they are initiated, and the different approaches adopted have been described in much detail by Levy (1995), who has also reviewed a considerable number of selected studies. There is no need to repeat this survey here. Though generally falling in the category of policy analyses and projections, many of the studies and regular activities reviewed by Levy would not appear to come directly under the heading of formal policy evaluation, in the sense of assessing the extent to which given measures contribute to the objectives pursued, in a cost-effective manner.

Some of the recommendations made by Levy in 1995, in order to improve the process of evaluating the CAP, are now outdated, given the more recent establishment of formal evaluation procedures by the Commission. Other recommendations still merit consideration, and may not have been made redundant by the procedures now adopted by the Commission. Without going into detail, the following observations on CAP evaluation done by and for the Commission can be offered.

- 1) Fundamentally, a far reaching process of evaluating the CAP is obviously going on in the Commission, in a more or less formal way. The strong evidence of this is the sequence of bold reform proposals made since 1991, starting with the MacSharry reform and continuing in the Agenda 2000 proposals. Irrespective of the merits of these reform proposals, their very existence proves that the Commission has come to the conclusion that the CAP as it stood in 1990 did not achieve its objectives, and the broader goals of the Union, in a satisfactory way. Such conclusions can only result from an evaluation of past policies, conducted in whichever way.¹⁰ Drawing

⁹ In addition, there is the constant flow of unsolicited external studies, mainly done by academics. As this study is focussed on formal evaluation procedures, these unsolicited studies will not be considered here.

¹⁰ The Agenda 2000 proposals for agricultural policies, more apparently than the MacSharry reform proposals, were indeed based on a significant number of studies done by and for DG VI, including the reports on the situation of agriculture in the countries of Central Europe and the related summary report (European Commission, 1995a), the Agricultural Strategy Paper (European Commission, 1995b), and the Working Documents under the heading of CAP 2000 (e.g. European Commission, 1997b). These studies and their impact on the Agenda 2000 proposals

constructive conclusions from past policy failures, or from the insight that current policies will not suffice in future, is the most important and desirable end result of any policy evaluation. Formal procedures of policy evaluation may be technically as good as conceivable, they will never substitute for the willingness to change a policy that no longer functions well.

- 2) The new formal process of evaluation established by the Commission under the SEM 2000 initiative is a significant step forward. It is based on proper concepts, well organised, methodologically sound, and comprehensive. As far as its structure is concerned, if anything it may run the risk of being over-administered and resulting in a machinery that could turn out to work more for itself than for the process of preparing policy proposals for the future. This risk is the better avoided the more this process involves staff in the operational units, as opposed to being constrained to "evaluation specialists". This said, there are some counter-arguments concerning the desirability of independence between the evaluation unit and the operational units responsible for execution of policy. This issue is discussed further in Chapter 8.
- 3) As the new formal process of evaluation has only recently started in DG VI it is too early to assess its outcome. As far as possible to see at this stage, the approaches adopted by DG VI hold the promise of generating useful results. In particular, the significant involvement of staff from the operational services responsible for the measures concerned creates a sense of ownership of the evaluation and thus the potential for feeding the results of evaluation activities directly into the process of administering and developing the concrete policies. If the "spirit of evaluation" is spread as far as possible among those directly responsible for individual measures, a major benefit of evaluation has been achieved. In this sense, the process of evaluation and the thinking it provokes among all those participating may, in the end, turn out to be more important than the formal evaluation results as appearing in the form of evaluation reports.
- 4) Because of the importance of involving operational staff responsible for the measures concerned in the process of evaluation, it may appear recommendable to keep the size of evaluation units reasonably small. On the other hand, too small a staff in the new evaluation unit in DG VI could be a major constraint on the effectiveness of CAP evaluation in the Commission, given the large amount and complex nature of the measures pursued under the CAP, and the big volume of budget expenditure involved in these measures. It may well be appropriate to see the evaluation unit as the initiator and leader of evaluation activities involving also staff from the operational units, and not as the sole provider of evaluation. The role of the unit would, from that perspective, be that of a "spearhead" of evaluation to be continuously conducted throughout DG VI, but not that of a body engaging in an activity pursued independently from the day-to-day business of administering and developing the CAP. However, evaluation also requires a good deal of methodological experience that has to be provided by specialists, and a certain amount of independence from the implementation of the policies concerned. Moreover, sound evaluation takes time that may not always be sufficiently available among the operational staff. The staff engaged exclusively in evaluation should, therefore, not be too small.

cannot be discussed here in any detail.

Hence, some increase in the size of the evaluation unit in DG VI may well be appropriate. Some suggestions on the overall amount of resources to be made available for evaluation of the CAP will be offered below in Chapter 8.

- 5) Though the CAP measures to be evaluated according to the evaluation plan for 1998 (part of which will not be finished before well into 1999) appear to cover a large part of FEOGA Guarantee expenditure, the majority of the number of evaluations planned, and all of the intermediate (as opposed to *ex ante*) evaluations, deal with measures that are relatively peripheral to the functioning of CAP market policies (such as promotion and programmes for islands [POSEICAN, POSEIMA, POSEIDOM]). The large coverage of FEOGA Guarantee expenditure results from the *ex ante* evaluations in process for the Agenda 2000 reform proposals in the sectors of arable crops, beef and milk (which amongst them account for 64.2 per cent of FEOGA Guarantee expenditure budgeted for 1998). Important as these *ex ante* evaluations may turn out to be in the future, it cannot yet be said that the functioning of the majority of existing CAP measures in the recent past is being made subject to formal evaluation at this time.
- 6) More generally, it is desirable that evaluations not only deal with the more "technical" aspects of the implementation of individual measures, but also with the fundamental orientation of the overall policy regime. Eventually, the purpose of evaluation should not only be to improve the management of given policies, but also the policies themselves. In an informal and rather general sense, a more strategically oriented evaluation of the CAP has been provided by the Commission, for example, in the Agricultural Strategy Paper (European Commission, 1995b). However, in future it would be desirable that such more fundamental evaluations of the CAP are also provided in a more detailed and quantified manner, on a regular basis, as part of the formal evaluation process now established in the Commission. Given the political sensitivity of such more strategic evaluations of broader policy domains, it may be necessary to rely mainly on external experts for them (see below, Chapter 8). In the same context, it is important that in evaluating the cost-effectiveness of measures, costs are defined in a broad sense, not only including budgetary expenditure (see below, Chapters 5 and 6).
- 7) In addition to studies done in DG VI, other DGs, in particular DG II, occasionally engage in, or commission, analyses of the functioning of the CAP. One example is the study on "EC Agricultural Policy for the 21st Century" (Larsen *et al.* 1994) commissioned by DG II from academic experts. Though not falling in the category of evaluation in a narrow sense, such studies (can) contribute a lot to understanding and assessing the effectiveness of the CAP. However, looked at from the perspective of evaluation done in the Commission, it is not exactly clear what the relationship is which such studies have with the process of evaluating measures in the Commission. Moreover, as evaluation of measures under the SEM 2000 initiative is to be conducted within the DGs responsible for the measure concerned, it is not clear what the role, if any, is given to economic analysis done by DG II. It would be desirable to establish more direct links between such studies done by other DGs and the evaluation process in DG VI.
- 8) The approach adopted for the *ex ante* evaluations of the reform proposals in DG VI holds the promise of creating an extremely useful basis for future intermediate and *ex post* evaluations of the measures concerned. In particular, the establishment of criteria and indicators which can later be used to evaluate the effectiveness of the policies covered has the potential of providing exactly

those tools which will be necessary for assessing the success, or failure, of the new policies proposed by the Commission. In the longer run it would be desirable to involve the Council and the European Parliament in the discussion of these criteria and indicators. Fundamentally it would be ideal if decision makers responsible for given policies were to provide the general public with indications of what exactly it was they were trying to achieve, and how the outcome of their decisions can be assessed. There is no doubt that this is a very demanding request. However, transparency and accountability can only be fully generated if it is established, *ex ante*, how policy success or failure can be assessed. This issue will be taken up again below in Chapter 8.

- 9) Ideally, the results of evaluation exercises should be available at the time changes to past policies are being considered and proposed. Given the recent establishment of the new formal evaluation procedures in the Commission, the evaluations resulting from this process could not be used as a basis for the Agenda 2000 proposals. To be sure, some form of evaluation lay behind these proposals, because otherwise they would not have been made (as argued above, with reference to some studies). In the future it would be desirable to require that proposals for policy changes and for new measures make explicit reference to the results of evaluations of the functioning of past policies (see below, Chapter 8). The 1996 Communication by Commissioners Liikanen and Gradin has indeed established this principle, and it is to be hoped that it will find its way into day-to-day Commission practice.
- 10) As a final point on the Commission's role in CAP evaluation, it is interesting to note that in some, though by far not all, Commission proposals for Council Regulations on agricultural policies under Agenda 2000, explicit reference is made to evaluation. In the proposal for a Regulation establishing common rules for direct support schemes under the common agricultural policy, Article 9 reads: "In order to gauge their effectiveness, payments under support schemes shall be subject to evaluation designed to appraise their impact with respect of their objectives and to analyse their effects on the relevant markets." The proposal for a Regulation regarding support for rural development from FEOGA requests that rural development plans should include "a quantified description of the current situation showing ... the main results of operations undertaken in the previous programming period with regard to the evaluation results available" and "provisions to ensure an effective and correct implementation, including monitoring and evaluation with the definition of quantified indicators for evaluation" (Article 41). Equally, the proposal for a Regulation regarding Community support for pre-accession measures for agriculture and rural development in the applicant countries of Central and Eastern Europe in the pre-accession period suggests that "in order to assess their effectiveness, support for rural development measures shall be subject to prior and mid-term appraisal, on-going monitoring and *ex post* evaluation designed to appraise the success and impact with respect to the defined objectives" (Article 5). All of these proposed Regulations relate to measures that are implemented by the (current and prospective) Member States, in part with significant latitude for their national decisions. It is certainly not a bad idea to require evaluation in these cases. However, it may also not be a bad idea to include the requirement for evaluation also in regulations relating to measures administered more directly by the Commission. In a way that may not appear to be necessary, because of the existence of the SEM 2000 evaluation procedures. However, if the evaluation requirement is included in regulations eventually decided by the Council, the Council is implicated in the process of evaluation, and may take the results of evaluations better into account in its decision making.

3.2. CAP Evaluation in the European Court of Auditors

The European Court of Auditors (ECA) audits expenditure from the EU budget, including expenditure under FEOGA Guarantee, on a regular basis. These audits, a major part of which is presented in the Court's Annual Report, generally address financial management, legality and regularity of EU expenditure and therefore do not directly fall in the category of policy evaluation as considered in this study. The 1990 Audit Manual of the ECA distinguishes audits clearly from assessments of "bonne gestion financière", a concept which, according to Levy (1995), is closely equivalent to "value for money". The ECA Audit Manual describes this latter type of assessment as being concerned with the "Three Es", i.e. effectiveness, economy and efficiency. Analysis of "bonne gestion financière" therefore comes much closer to the concept of policy evaluation. Though some of the Court's statements in the Annual Reports can also be classified as falling in that category, assessments of "bonne gestion financière" find their place mainly in the Court's occasional Special Reports, Opinions and Observations in which the ECA comments in detail on selected EU policies, including measures pursued under the CAP. These Special Reports, Opinions and Observations are usually published in the Official Journal and therefore available to the general public. The following brief observations address only these types of ECA assessments.

Special Reports issued by the ECA generally deal with measures as currently pursued by the Union. In terms of evaluation categories, they therefore mainly fall in the category of *ex post* or intermediate studies. In its Opinions and Observations the ECA mainly comments on Commission proposals or plans for future policies or policy reforms, either in response to a request from another Union institution (e.g. the Council) or on its own initiative. These Opinions and Observations, therefore, come closer to *ex ante* evaluations, though in many cases they are also based on results of earlier audits as presented in Annual and Special Reports. An internal list made available by the ECA records 22 Special Reports and three Opinions adopted by the ECA between 1989 and 1998 concerning EU agricultural measures. The topics addressed in these documents cover a wide range of measures and policies under the CAP such as, for example, the agri-monetary system, the CMOs for various agricultural products or individual elements of them (e.g. the dairy quotas), the management and control of export refunds, controls of irregularities and frauds in agriculture, and the clearance of accounts for guarantee expenditure under the EAGGF. In addition, in 1998 the Court adopted Observations on the Commission plans for a reform of the CMO for olive oil, and an Opinion on the Commission proposals under Agenda 2000. It would go far beyond the scope of this study to engage in a detailed discussion of how the ECA dealt with the CAP in these many documents. Instead, only a few deliberately selective observations will be offered here.

A first point to be made about the Court's Special Reports is that many of them are exclusively or mainly technical audits of financial management, legality and regularity of EU policy measures and related expenditure, and as such cannot be called evaluation reports in the sense considered in this study. That is the case not only for reports dealing with topics such as irregularities and fraud, or clearance of accounts, but also for reports dealing with seemingly broader policy issues under the CAP. For example, the Court's Special Report No. 5/97 on the management of EU cereals trade involving export refunds, special import arrangements and regional schemes (OJ C 159, 26/5/1997) does not deal at all with the economics of export restitutions etc., but provides a technical audit of the administration and implementation of the measures concerned, dealing with issues such as the appropriateness of the conversion coefficients used by the Commission and with the accuracy of the

physical controls of the quantities shipped. Similarly, the recent Special Report No. 4/98 concerning the importation at reduced duties of certain dairy products from New Zealand and Switzerland (OJ C 127, 24/4/1998) addresses only the technicalities of administering the measures concerned. In such cases, the Court restricts its comments to technical matters even though it would be easy, if not tempting, to add a few more fundamental comments on the economics of the measures concerned, for example regarding the optimal timing of EU cereals exports with restitutions, or the economic rationality (or lack of it) of setting minimum import prices for duty-reduced imports of Swiss cheese.

Some other Special Reports, though, go further in the direction of analysing CAP measures from the perspective of the "Three Es". However, even where that is the case, such analysis is often combined, in a very specific way, with elements of more technical audits. Moreover, fundamental economic issues related to the policies considered are often not discussed, and the Court's comments are frequently confined to improving the working of the measures already in existence, rather than comparing them to alternative policy designs. Moreover, occasionally the economic analysis presented is somewhat surprising. For example, Special Report 3/94 on the CMO for beef (OJ C 356, 14/12/1994) discusses at length technical, though financially relevant, details such as deboning and control of intervention stocks. It (rightly) criticises some absurdities in the implementation of intervention measures (such as reduction coefficients which result in excess offers for intervention). But in that report the Court also discusses the more fundamental orientation of the 1992 reform measures, though it is somewhat difficult to accept the Court's comment that they aimed primarily at improving the income of beef producers, rather than addressing the surplus problem (which the 1992 reform did in various ways, though eventually with limited success). It also is somewhat surprising that the Court uses statistics showing that the income of beef producers is around half that of milk producers to suggest that the CAP objective of a fair standard of living has not been met by the CMO for beef. Based on the statement, made in an earlier audit of the beef regime, that some specific payments to beef producers do not raise their incomes, the Court argues that the new payments introduced as part of the reform have only marginal or no effect on producer incomes, without providing any analytical details. Most surprising from an economic point of view is the fact that the Court calculates a value of the surplus on the beef market by multiplying the surplus quantity by the EU intervention price, though it is not at all clear what the intended use of that information is, or ever could be.

Even where a Special Report very strongly criticises an element of the CAP on seemingly fundamental economic grounds, and hence where it appears to engage in policy evaluation as considered in this study, the analysis provided is somewhat selective. A case in point is Special Report No. 4/91 on the CMO for sugar (OJ C 290, 7/11/1991). In this report the Court goes as far as speaking of the "sugar lobbyists", and of window dressing in relation to the so-called financial neutrality of the sugar regime, which the Court also rightly shows to be achieved, in reality, by placing the economic burden on sugar consumers. The report strongly criticises that sugar quotas are used to satisfy the interests of producers and Member States rather than to control production and that high world market prices at times of revisions of the regime were used as arguments for generous quotas and prices. The Court also makes a number of convincing critical comments on various aspects of the regime, including the economic implications of C sugar and the provisions for ACP sugar exports to the EU. At the same time, though, the sugar report also includes a long discussion of financial management and administration of various technical details of the regime. Surprising is the fact that the Court's bottom line under all its well justified and strong criticism of the EU sugar policy is not a comparison with

(if not an argument for) a fundamentally different and much more market oriented policy¹¹, but the comment that the level of guarantees, both price and quota quantity, is too high. In the report, the Court calculates the area under sugar beets that would have sufficed to produce enough sugar to cover EU consumption, implicitly suggesting that a quota set along such lines would have been appropriate. The more fundamental economic issue of whether self-sufficiency in sugar is a desirable orientation for EU sugar policy is not debated, and the Court also does not discuss whether and how the EU could possibly live without any quota regime for sugar at all.

Less frequent than the Special Reports are the Court's Opinions and Observations providing *ex ante* assessments of policies proposed or considered by the Commission. The way the Court deals with future policies in these statements again differs significantly from case to case. For example, in its Observations on the February 1997 Commission paper on options for a reform of the CMO for olives, presented on its own initiative by the Court in March 1998 (European Court of Auditors, 1998), the Court makes a number of rather critical comments on the way the Commission has analysed the situation and estimated future market developments. The Court also criticises, in its words, a lack of analysis, based on appropriate information, of the possible impact of some elements of the options considered by the Commission. Moreover, the Court comments that the Commission paper did not indicate the specific objectives of the alternative options presented. On the other hand, in these Observations the Court essentially does not engage in its own analysis of what the implications of the policy options considered might be, nor does it provide a detailed discussion of the objectives of the COM for olive oil. Most of the report deals with administrative aspects of the COM for olive oil, with the accuracy of data (or rather the lack of it), and with various aspects of financial control.

Of a rather different nature is the Court's most recent Opinion on Agenda 2000, adopted on 29 October 1998 (Cour des Comptes Européenne, 1998). In this document, presented on invitation by the Council, the Court comments, among others, on the Commission's proposals for CAP reform of March 1998.¹² The approach adopted is that of a fairly general evaluation of major political and economic aspects of the reform proposals, going far beyond financial aspects and administrative implementation. The Court welcomes the Commission's reform initiative in general, but is rather critical of a large number of specific aspects. The Court finds the Commission's budget projections too optimistic and confronts them with its own estimates in which the agricultural guideline is projected to grow less and FEOGA expenditure to grow more, making expenditure exceed the guideline significantly in the years 2002 to 2005. The Court also questions the Commission's projections of world market prices for cereals, and argues that the validity of the whole reform initiative could be seriously undermined if the price cuts proposed by the Commission are not sufficient to allow the EU to export cereals without subsidies, and therefore do not relieve the EU from the WTO constraints on export subsidies. Regarding the next round of WTO negotiations, the Court comments that the blue box is likely to be examined, with implications for the way the EU can implement its direct payments, which according to the Court should be changed such that they can

¹¹ A proposal for a fundamentally different sugar market policy in the EU, based on an evaluation of the current regime, has for example been made by the Scientific Advisory Council of the German Ministry of Agriculture, see Wissenschaftlicher Beirat beim Bundesministerium für Ernährung, Landwirtschaft und Forsten (1994).

¹² The three other parts of the Court's Opinion on Agenda 2000 (regarding structural funds, pre-accession measures and external actions) will not be considered here.

be brought in the green box. The Court is generally rather critical of the direct payments. It makes the point that undifferentiated subsidies to agricultural activities which might be profitable without government support are not justified. The Court also suggests a much more pronounced degressivity of payments with increasing farm size (with an absolute ceiling at Euro 100,000 per farm), and calculates the savings that could be made if the Court's parameters were chosen. On the Commission's proposals for the milk market, the Court criticises the quota increase foreseen and suggests that the Commission should examine the implications of an elimination of the quota regime. In its concluding remarks, the Court suggests, among others, that the Commission should work towards finding approaches that enable policy objectives to be achieved in a more cost-effective manner.

This rather interesting Court Opinion, which contains many more critical comments on other aspects of the Commission's proposals for CAP reform under Agenda 2000, adopts a fairly broad perspective and makes a number of salient comments on fundamental and highly sensitive political issues, just as the Council began to prepare for the last round of negotiations on the Commission's proposals. The impression conveyed by the document is more one of a general political assessment than that of a detailed analytical exercise. For a full-scale *ex ante* evaluation in a technical sense a much more detailed analysis of the links between measures and objectives would be desirable. However, it is quite likely that such an exercise might generate results some of which could come close to some of the comments made by the Court.

In summary, most of the ECA's work on the CAP falls in the category of auditing financial management and technical implementation. However, occasionally the Court also presents assessments that go in the direction of policy evaluation, though in a rather limited way. In its most recent Opinion on Agenda 2000 the Court has adopted a fairly general political perspective.

If the Court's assessments go beyond pure audits, they are particularly strong where they build on the information regarding CAP implementation collected by the Court in Member States' administrations and companies, in a way that few other evaluators could ever match. On the other hand, the Court's economic arguments are not always convincing, and its work on the CAP does not, and is not expected to, fulfil the need for full-scale policy evaluation.

4. Current Practice of CAP Evaluation in the Member States

4.1. Germany

A formal process of policy evaluation like that now existing in the European Commission under the SEM 2000 initiative, with specialised evaluation units, evaluation plans, and evaluation reports, has not been established in German ministries, neither in the Ministry of Food, Agriculture and Forestry nor in any other Federal ministry. What potentially could come close to elements of a formalised process like that is a general regulation issued by the Ministry of Finance regarding cost-effectiveness ("Wirtschaftlichkeit") of public management in the context of the budgetary process.¹³ This regulation, valid for all public authorities at the Federal level (and hence also for the Ministry of

¹³ The most recent amendment of this regulation was circulated in Bundesministerium der Finanzen (1995).

Agriculture)¹⁴, aims at the best possible use of resources and requests that the principle of cost-effectiveness is pursued in all measures directly or indirectly relevant to the budget of Federal authorities. In particular, the regulation stipulates analyses of cost-effectiveness for all measures. It distinguishes between *ex ante* analyses during the planning process, intermediate analyses aimed at monitoring the effectiveness of current measures, and *ex post* analyses after completion of a measure.

For *ex ante* studies of measures being in their planning stage, the regulation requires that they should make statements on, among others, the objectives of the measures and any conflicts among them, benefits and costs, including non-monetary costs, budgetary implications, and criteria for assessing effectiveness. Intermediate and *ex post* studies are expected to assess whether and to what extent the objectives aimed at have been achieved, whether this was due to the measure adopted, and whether the measure was cost-effective. For measures that extend over more than two years, and in other appropriate cases, intermediate assessments are called for, at points in time to be scheduled as appropriate for the measure concerned. These interim assessments are expected to provide information necessary to decide whether and in which form the measure concerned should be continued, given the changes that might have occurred in the economic, social and technological environment. All measures that have been completed are expected to be studied *ex post*.

As far as methodology is concerned, the regulation makes a distinction between micro-economic and macro-economic analyses. Micro-economic analyses are requested for measures whose implications are constrained to the ministry or administration concerned (e.g. procurement of equipment or changes in administrative procedures). In such cases, the methods to be used are those also used in companies when they plan their investments. Macro-economic analyses, on the other hand, are requested in all cases where broader economic effects are to be expected. Obviously it is the latter category of analyses that are of particular relevance in the context of evaluation. For measures with such more wide-ranging effects, the methods to be used are those that consider overall economic effects, like cost-benefit analyses. The regulation stipulates that the studies have to be conducted by the services responsible for the measures concerned.

When rules like these were first introduced for the Federal ministries and other authorities in the early 1970s, a considerable number of cost-benefit analyses were indeed done, mostly commissioned from external experts. Agricultural policy was no exception, and several studies of the cost-effectiveness of concrete agricultural policy measures as well as analyses of more fundamental policy issues were conducted.¹⁵ However, more recently the formal requirement to engage in studies of cost-effectiveness, as laid down in the regulation cited above, no longer plays much of a role in the day-to-day practice of the Ministry of Agriculture, and cost-benefit-analyses or similar forms of more formal evaluations are rarely conducted. The reason given is twofold. On the one hand, the tight budget situation has meant that all measures not absolutely necessary have already been weeded out, and for the same budgetary reasons there is anyhow no scope for introducing new measures where *ex ante* analyses might be useful. On the other hand, those measures that are still in place are politically so

¹⁴ Equivalent provisions apply at the level of the German Länder, and are therefore relevant for the Länder ministries of agriculture.

¹⁵ Two examples of such studies looking into more strategic issues of agricultural policy are Koester and Tangermann (1976) and Sohn (1984).

sensitive that there is not much point in assessing their cost-effectiveness. All this applies to measures falling in the category of national agricultural policies. Union measures under the CAP are rarely in the domain of policies covered by the cost-effectiveness regulation cited above.

Though formal evaluation in the strict sense of the term is not pursued in the German Ministry of Agriculture, important policy measures are regularly reviewed, in a somewhat less formal process, in what is called the Planning Group of the Ministry. Each department in the Ministry has one unit responsible for general affairs and planning issues relating to matters under the responsibility of that department. The heads of these units, and in addition the head of the unit responsible for budgetary affairs, are members of the Planning Group, which is chaired by the deputy head of the department for general matters of agricultural policy. Fundamentally, the task of the Planning Group is to do the conceptual groundwork for policy decisions to be taken in the Ministry. For this purpose, the Planning Group can deal with all matters of fundamental importance, and cutting across the responsibilities of individual departments, whenever necessary. For example, the Planning Group would look at the Commission proposals for the MacSharry reform of the CAP and submit, after consideration in the group of department heads, a document commenting on the Commission proposals to the Minister.

On a regular basis, the Planning Group has to deal with what is called the integrated task and budget planning ("integrierte Aufgaben- und Finanzplanung"). Of particular importance in this context is the annual process of preparing the budget proposal for the coming year, and the related financial planning for the following three years. In going through the individual budget items, the Planning Group reviews all measures under the responsibility of the Ministry and considers changes that might be necessary from either a budgetary or a substantive perspective. On that basis, the Planning Group comes up with a draft budget proposal, which is forwarded to the department heads, and later the Minister, for decision. The budget proposal then forms the basis of the budget negotiations between the Ministry of Agriculture and the Ministry of Finance.

In a related process, also on an annual basis, the Planning Group considers which measures may require, in the coming year, specific analysis or political decision. The result of this process is a list of measures, existing and new, to which particular attention should be paid in the coming year ("Vorhabentableau"). The measures concerned could be both national policies (e.g. elimination of legal and tax provisions that inhibit structural adjustment in German agriculture) and Union policies under the CAP (e.g. the future of the quota regime for milk production). In a way, this process of selecting measures requiring review can be said to contain, implicitly, some elements of informal evaluation, as the measures listed are those where action is considered necessary because the operation is not satisfactory or the circumstances have changed. However, in establishing the list, a fair amount of political judgement appears to be involved. Measures where changes are considered infeasible for political reasons hardly find their way onto the list. Hence this process does not result in a regular review of all measures for which the Ministry bears responsibility. For each of the measures on this list, a document will then be prepared analysing the situation in detail, and assessing the operation of the measure concerned and the need for political action. Even though also not falling in the category of formal evaluation in a strict sense, these documents probably come closest to what could be called evaluation of individual measures. On the basis of these documents, the Planning Group then makes proposals for action, for consideration by the heads of departments and eventually the Minister.

In summary, a formal process of regular evaluation, comparable to that under the SEM 2000 initiative in the Commission, is not pursued in the German Ministry of Agriculture. The general regulation for the whole Federal government requiring periodic analyses of cost-effectiveness (through, for example, cost-benefit analysis) no longer appears to play much of a role in day-to-day practice (not only in the Ministry of Agriculture, but also in other ministries, it appears). The activities of the Planning Group come closest to what could be called evaluation, and result, among others, in documents that assess the merits and drawbacks of individual measures. However, there is no provision which would make sure that all measures are reviewed on a regular basis, in a more or less standardised format. Broadly speaking, it appears that considerations regarding political imperatives dominate over any felt need for routine processes of regular policy evaluation.

Only loosely related to evaluation, but worth mentioning in this context, are the activities of the Scientific Advisory Council of the Ministry of Agriculture ("Wissenschaftlicher Beirat"). This institution, also existing in similar forms for a number of other ministries in the Federal government, consists of up to 15 academics, usually professors of agricultural economics and general economics from German universities. Members are appointed by the Minister, on proposal by the Scientific Advisory Council. The Council is expected to provide advice to the Minister, in complete independence. The Council, meeting around three times a year, decides independently on its agenda, but consults with the Ministry on topics that might be of interest. The major activity of the Council is to draft reports on strategic issues in agricultural policy. Recent reports of the Council have dealt with issues such as integrating European agriculture into world agriculture; allocation of competencies to the Union, the Member States and regional authorities; decoupling of support; agricultural implications of Eastern enlargement; the future of land use in Germany; and fundamental reform of the EU sugar market regime.¹⁶ More recently the Council has also begun to draft short statements on topical issues, the first addressing the Commission proposals for reform of the CAP under Agenda 2000.¹⁷ The reports are submitted to the Minister, and discussed with him and his staff. The Minister then decides on publication of the reports. So far, all reports of the Council have been published.

Though also not policy evaluation in a formal sense, and far from being comprehensive in the coverage of all agricultural policy measures, the reports of the Scientific Advisory Council contribute to assessing the functioning of existing agricultural policies, with a particular emphasis on their economic effects. Generally the Council tends to argue for more market orientation in agricultural policies, and less government intervention. However, the Council also has tended to accept the need for gradual adjustment in policies, and for income compensation where price support is reduced. The reports are occasionally hotly debated in the general public, not the least because they often deviate fundamentally from the policy orientation of the Minister. The reports of the Scientific Advisory Council do not directly influence agricultural policy making in the Federal government, but it is occasionally said that they have some impact on thinking among the staff of the Ministry, and thereby potentially have some bearing on the preparation of longer run adjustments in policy positions.

¹⁶ Wissenschaftlicher Beirat beim Bundesministerium für Ernährung, Landwirtschaft und Forsten (1997a, b, c, 1996, 1994). The report on integrating European agriculture into world agriculture has not yet been submitted to the Minister and is not yet published.

¹⁷ This short statement has not yet been submitted to the Minister and is not yet published.

4.2. The Netherlands

The Netherlands have a long tradition in planning and evaluating economic policies. Moreover, in agriculture the Netherlands is among the countries where top research of excellent international standing is done, including research in agricultural economics. As far as academic research in the area of agricultural economics is concerned, Wageningen Agricultural University is an outstanding centre, though economists from other universities in the country also work in this area. In addition, agricultural economics research is also done at a large scale at the Agricultural Economics Research Institute (Landbouw-Economisch Instituut, LEI-DLO), established under the Ministry of Agriculture, Nature Management and Fisheries (LNV). The internationally well-known agricultural economics profession in the Netherlands generates a large amount of studies on the CAP and its various individual measures, improving the general understanding of the functioning of these policies and thereby creating the basis for successful evaluation. It also conducts, and contributes to, studies commissioned by the European Commission.¹⁸

The Dutch Ministry of Agriculture also commissions studies of strategic issues in agricultural policy, and uses the results in the preparation of policy decisions. For example, in order to analyse proposals for CAP reform, researchers may be asked to engage in studies based on quantitative modelling.¹⁹ An area that is being studied with particular emphasis in recent years is the relationship between the CAP and the environment, where a number of studies have been, and are being, done by LEI-DLO.²⁰ These studies have made significant contributions to evaluating the environmental effects of the CAP. Work is also done in LEI-DLO on the establishment of robust environmental indicators that can be used to evaluate policies. A study recently done by LEI-DLO, on the request of the Ministry of Agriculture, provided an economic and administrative evaluation of the innovation policy of the Ministry.²¹

As far as formal evaluation of agricultural policies is concerned, a routine process through which all relevant measures would be evaluated on a regular basis did not in the past exist in the Dutch Ministry of Agriculture. This is not to say that evaluation of policies was not done, but it occurred more on an *ad hoc* basis. Major policy measures and politically important issues were assessed from time to time, but the process was not comprehensive, leaving aside many individual measures. Moreover, there was no standard format nor a standard method for policy evaluation. In some cases, a one page document was produced, summarising the results of an evaluation and providing basic information. In other cases, research institutes were commissioned to engage in more in-depth evaluation studies. Such an evaluation could contain, but did not have to include, an analysis of the cost-effectiveness of the measure concerned. It could, though, also be rather descriptive, reporting, for example, on the number of people who had applied or the amount of paperwork involved. Evaluation studies of this type related mainly to structural policies in Dutch agriculture. It appears that an overall evaluation of price support policies under the CAP has not emerged from this process.

¹⁸ An example is Folmer, Keyzer, Merbis, Stolwijk and Veenendaal (1994).

¹⁹ A model occasionally used for such purposes is ECAM as developed and used by Folmer *et al.* (1994).

²⁰ See, for example, Brouwer and van Berkum (1996).

²¹ Diederren, van den Eeden and Kuper (1998).

However, in the context of the budget preparation 1996, an interdepartmental policy study of European expenditure was commissioned. Part I of this study dealt with agricultural expenditure.²² The study looked in particular into the net contribution of the Netherlands to agricultural expenditure of the European Union and how this might be affected by alternative scenarios for future developments of the Common Agricultural Policy, considering the implications of future developments in WTO negotiations on agriculture and of Eastern enlargement of the EU. The scenarios considered included the increasing use of direct payments, under a complete elimination of price support. As one option under such a scenario, the complete financing of the direct income payments by the Member States concerned was analysed, with the implication that there would no longer be any expenditure on agricultural policies by the European Union. It was found that a complete shift from price support to direct income payments would reduce the net contribution to the Union budget that the Netherlands make under current policies, and that national financing of all direct payments under this scenario would eliminate the Dutch net contribution to the EU budget altogether.

One type of use of the results of evaluation studies is related to the budgetary process in the Netherlands. When the Minister of Agriculture presents his draft budget to the Parliament, the document submitted also has one chapter devoted to the evaluation of existing policies. This chapter typically draws on the results of *ad hoc* evaluations. Occasionally, however, evaluation studies also had a decisive and direct impact on policy pursuit. Two cases are reported in which negative results of an evaluation triggered the decision to discontinue the measure concerned. One such case was the installation premium for young farmers, which was eliminated in 1992 after a study done in Wageningen University had shown that this measure was very cost-inefficient. The other case was the early retirement scheme for farmers, which was abolished after a negative evaluation it received in an internal report produced in the Ministry.

More recently, though, the Dutch government has embarked on the process of establishing routine policy evaluation. In 1991 the Cabinet took the decision that policies should be evaluated on a regular basis. Each ministry is responsible for evaluating the policies in its own domain. The overall process is overseen by the Ministry of Finance. In this context, the Ministry of Agriculture produced three papers in 1998, on (i) the state of the art in policy evaluation in the Ministry, (ii) the strategic goals of policy evaluation in the years to come, and (iii) a plan of action for how to reach these goals.²³ Starting in the fall of 1998, all departments in the Ministry are expected to plan their evaluation activities for the future.

In the paper on the state of the art²⁴, the current situation regarding policy evaluation in the Ministry of Agriculture is assessed. The purpose is to create a vision of the policy evaluation process and to discuss issues regarding the quality, scope and costs of this process and how to structure it. Findings are that in most LNV directorates no vision on policy evaluation was developed in the past, and a process of policy evaluation was not yet programmed. Furthermore, the point is made that little

²² Government of the Netherlands (1996).

²³ Directie Financieel-Economische zaken (1998 a, b and c).

²⁴ Directie Financieel-Economische zaken (1998a).

awareness of the quality aspects, precise aims and costs of evaluation exists. It is reported that in the period of 1994-1997 a total of 41 evaluation studies were carried out. One of them was an *ex ante* evaluation, while the remaining 40 were implicit *ex post* analyses. It is found that the number of monitoring systems is increasing, especially in the agriculture directorate.

In order to improve the current situation, aims of policy evaluation are formulated in the document on the strategic goals of policy evaluation in the years to come.²⁵ The goals are described in a fairly general way. Emphasis is placed on support of decision making; evaluation of means and performance, which is described as important in particular in the context of increasingly decentralised organisation of policy making; and the need for growing awareness of the responsibility for making sensible use of scarce resources.

Finally, in the third document a plan of action is outlined.²⁶ Implementation of the new process of policy evaluation is to be monitored by the directorate on financial and economic affairs of the LNV. It is envisaged to programme the policy evaluation process on a yearly basis starting in 1999. The intention is that all policy areas are evaluated at least once every five years. Moreover, an *ex ante* evaluation is foreseen for all newly designed policies. General criteria for the quality of policies are stressed, with an emphasis on the sufficient targeting of policies and the appropriateness of the policy with regard to the aims pursued. The directorate on financial and economic affairs is to provide a yearly overview of the policy evaluation process in the various directorates. This overview is to be presented to the *bestuursraad* of the LNV, a high ranking steering committee including the permanent secretary and the director generals. The *bestuursraad* has the possibility to initiate additional policy evaluation at the expense of the directorate concerned, if existing evaluation is found to be unsatisfactory.

As far as methods of policy evaluation are concerned, it is stressed that results must be generated in a way that they can be reproduced. Conclusions must follow logically from the evaluation, and recommendations on future action must follow logically from the conclusions. With regard to the issue of who should engage in policy evaluation, there is a preference for in-house evaluation, because of cost reasons, in order to improve the positive effects of evaluation in terms of learning for future policy decisions, and to secure a better quality of evaluation results. There is no indication of the nature of consequences that might be drawn from the results of evaluation activities. With regard to the domain of policies that might be evaluated, there is no explicit definition, but it would appear that both domestic policies pursued in the Netherlands and agricultural policies administered at the level of the Union are potential candidates for evaluation.

As a lack of knowledge about, and experience with, policy evaluation was found in the document on the state of the art, a "help-desk" is to be set up. Assistance in programming the process of policy evaluation is to be provided by the directorate on financial and economic affairs. It is also envisaged that a manual on policy evaluation will be produced and courses on policy evaluation will be organised.

²⁵ Directie Financieel-Economische zaken (1998b).

²⁶ Directie Financieel-Economische zaken (1998c).

Furthermore, an information system on policy evaluation is to be established by 1999. This information system is designed to provide information on the quality, quantity and costs of policy evaluations as well as a database of all evaluation reports within LNV. The system is also thought to make information available on external consultants who have contributed to the evaluation process.

As this explicit, regular and more formally structured process of policy evaluation in agriculture will not become fully operative before 1999 it is too early to say anything about its outcome. However, it will be interesting to see how this process develops and how both the procedures adopted and the results achieved, including implications for future policy design, compare to the similarly new process of policy evaluation established in the European Commission under the SEM 2000 initiative. In a number of regards it appears that there are similarities between the Commission initiative and the Dutch plans, though obviously both the audiences concerned and the policies to be covered are different. There may be a point in comparing these two processes in the future once the first solid experiences have been made. It should also be interesting to compare the results of these two exercises in cases where the same policy measures are evaluated at the two different levels, if that should ever be the case (e.g. changes to the individual market regimes under the CAP as resulting from Agenda 2000).

4.3. The United Kingdom

The current systematic and formal process of policy evaluation in the UK was described in a manual prepared by the Treasury in 1988, 'Policy Evaluation: a guide for managers'. UK usage refers to 'policy evaluation as the process of examining a policy while it is in operation or after it has come to an end' (which have been referred to in this report as ongoing and *ex post* evaluation). Policy appraisal is 'analysis done before a policy is launched', (called *ex ante* evaluation in this report). The procedures outlined in this guide were to be built into all new UK policy initiatives and all proposals arising from policy reviews. Thus there is a ten-year period during which the evaluation culture has been instilled into the British civil service. Appraisal and Evaluation are seen as important stages in the policy cycle pictured below.

The Treasury's Evaluation Guide is a simple and clear discussion of all the main issues of evaluation. It covers: the purpose of evaluation, the distinction between evaluation and monitoring, defining the scope of each evaluation, definition of objectives, the need for a base case (the UK term for the counterfactual policy), choosing measures and indicators, collecting information, the merits of external and internal evaluation, defining costs of policy and cost effectiveness and issues concerning publication. When asked if there was a manual for evaluation in UK government departments, this is the document which was produced by the Ministry of Agriculture, Fisheries and Food (MAFF). It should be noted that in the UK evaluation mainly refers to economic evaluation. There are parallel procedures, and guides, for environmental assessment of policies. This does not preclude economic evaluation of environmental policies, which as reported below have been the subject of many evaluations by MAFF.

4.3.1. CAP Evaluation by the UK National Audit Office

The UK National Audit Office (NAO) came into existence in 1983, as a financially, and operationally, independent institution answerable through its head, the Comptroller and Auditor General (C&AG), to the Public Accounts Committee of Parliament. However, its operations are completely independent of Government. It is primarily concerned with the financial operations of government departments and other bodies receiving public funds. It conducts 'value for money' (VFM) examinations of these organisations and reports to Parliament and the public, its findings on the economy, efficiency and effectiveness with which public resources have been used. The C&AG has access to all information and explanations from departments which he may reasonably require.

NAO's role is defined in the National Audit Act 1983. Its task is to question the instruments used and the effectiveness of their use, to achieve policy objectives. Its role is restricted insofar as it is not entitled to question the merits of policy objectives. Nevertheless, NAO has otherwise complete discretion in whether to carry out a VFM examination and how to carry it out.

The NAO has a team currently consisting of five full time staff engaged in VFM investigations of agriculture expenditure, including that under CAP Schemes administered by the Ministry of Agriculture, Fisheries and Food. Examinations of the CAP expenditures present some interesting issues. First, NAO publishes the results of its VFM studies via reports to the Parliament. The accounting officer for the activity under examination is called before the Committee of Public Accounts to defend or account for the actions of his department. In the case of CAP measures the underlying policy, policy instruments and even the nature and level of administrative controls are determined by the European Union. As such there may be greater sensitivity or difficulty in asking Accounting Officers to defend positions beyond their direct control. Second, as is discussed at length in Chapter 5 on objectives of the CAP, most of the current measures in the agricultural policy are not in place to serve absolute need in specific circumstances, but they have arrived or have been modified from pre-existing measures, as the outcome of a negotiation process. This makes it difficult to determine precisely their objectives and thus to assess their value for money.

Notwithstanding these issues, the NAO seeks to carry out VFM studies of CAP schemes which may contain information or results of interest to those involved in review of the effectiveness of the schemes and to assess the VFM of UK implementation.

The National Audit Act 1983 restricts the C&AG's powers to examination of use departments have made of their resources. For this reason the nature of the evaluations is exclusively *ex post*. This does, however, not impede the VFM analysis to be forward looking and comment on some future implications of the results.

To select which policies to evaluate the NAO is continuously in the process of reviewing from a number of sources and through direct contact with MAFF the implications of agricultural policies.

The NAO generally chooses the areas to examine using one or more of the following criteria:

- the size of public expenditure involved,
- where it is believed that a VFM examination in a particular area can add value or identify efficiency savings,
- where there is a potential public or parliamentary interest,

where the policy has been covered previously or has been planned to be examined by NAO, the Agricultural Select Committee of the House of Commons, the European Court of Auditors (ECA) or other relevant parties, if there is prima facie evidence of poor value for money.

To perform the VFM evaluations the NAO has a number of approaches and techniques. For each examination a team is selected which may involve outside consultants. Generally for agriculture most of the work has been performed by in-house staff. However, NAO nearly always employs advisory panels whose members are drawn from relevant areas. These panels normally consult trade and other representative associations.

The external input is in any case very important in all studies. It is the practice for all reports to discuss with MAFF the proposals, the preliminary findings, the proposed coverage and the methodology of all studies. Before publication the Accounting Officer must agree on the factual contents of the report and comment on its presentation and the conclusions. This is fundamental to the process whereby he is called to account by Committee of Public Accounts. Furthermore, currently all published reports are subject to quality reviews by the London School of Economics (LSE). Given the LSE's academic excellence and independent status, this not only reinforces the quality and independence of NAO's work, but also allows the NAO to revise and refine its methodology.

Since 1988 the NAO has published 15 reports on agriculture on subjects as different as fraud prevention to the protection of environmentally sensitive areas. The latest report is on the Arable Area Payments which is expected to be published in February 1999. The techniques used for this audit have been developed in collaboration with the Dutch and Swedish Audit offices.

The NAO has shown a strong interest in collaborating with opposite numbers in other Member States and with the European Court of Auditors. This interest is reflected by its publication of a book entitled "State Audit in the European Union" in 1996. In this book the NAO has examined the role of each state audit office, referred to as Supreme Audit Institutions (SAIs) in the 15 Member States and at the European Union as a whole. Auditing methods are discussed annually in a Contact Committee which is attended by the Auditor Generals of the SAIs from all Member States and the President of the ECA. Those of the CAP always feature because of the size and importance of the budget. Despite this rapprochement, collaboration is very uneven as accounting practices can diverge substantially between Member States.

4.3.2. CAP Evaluation in the Ministry of Agriculture

MAFF follows the UK convention of distinguishing between appraisal (*ex ante* evaluation) and evaluation (*ex post* evaluation) and undertakes or commissions a wide range of both types of evaluation of policies for which it is responsible. The administration of *ex post* economic evaluations of policy is co-ordinated by Economics (Resource Use) Division, where policy evaluation represents the largest work area for one branch of three officials. *Ex post* evaluation is advocated as an essential process which ensures that policies remain relevant and that they meet their objectives in a cost-effective way.

The Division has produced a Policy Evaluation Guidance booklet for use by all programme managers in the Ministry in drawing up evaluation plans (MAFF, 1998a).²⁷ *Ex post* evaluations are usually carried out by external researchers.

The MAFF have had a systematic rolling programme of *ex post* evaluation of specific programmes under the CAP (and UK policy instruments too) for the last ten years. Policies are selected for evaluation taking into account the importance of the policy, the state of knowledge about its effects, the need for evaluation to feed into a review, commitments to the European Commission or Treasury, age of the policy and time since last evaluation.

These evaluations are commissioned by MAFF from independent research institutions following a system of competitive tenders. The use of external organisations for this evaluation is justified on the grounds that it provides an independent and objective analysis of policies. The terms of reference of invitations to tender are not prescriptive. Part of the assessment of bids for such contracts is based on the experience and expertise of the organisation and their proposed methodology for conducting the evaluation. Contracting out the evaluation process is thought to provide best value for public expenditure and the results are given credence through the researchers' independence. However, this procedure does mean that there is no standard methodology for evaluation. This means that the policy is thoroughly assessed for its own effectiveness but is not compared in a measured way against other policies in the UK or Europe.

The purpose of MAFF policy evaluation is to provide a *critical and detached* (their emphasis) assessment of the following factors (taken from Policy Evaluation Guidance Booklet 1998, paragraph 6)

”At the very least, it must assess the extent to which the policy objectives have been or are being met. Additionally, it may be appropriate to assess one or more of the following, depending on the policy and the circumstances at the time of the evaluation,

Is the rationale for the policy (e.g. market failure) still valid and appropriate?

Are the policy objectives still valid and relevant?

What are the mechanisms and instruments used to pursue the objectives? Are they the most appropriate ones?”

The coverage of this process is comprehensive. The policy to be evaluated is broadly defined as ”The objectives of MAFF and the Government and the preferred means of trying to achieve them”.²⁸ MAFF distinguish the rationale of policy from its aims and objectives. All three help explain why the policy was adopted, its purpose and what it is hoped the policy will achieve. Rationale is usually defined in terms of perceived problems or difficulties that the policy is intended to correct or compensate for. Objectives are yardsticks against which to judge the success of a policy. Aims are simply more

²⁷ The programme of evaluations is drawn up by MAFF for the whole of the UK. The actual evaluations are then sometimes done for the whole UK (as was the case for the recent evaluation of milk quotas). On other occasions there are separate evaluations arranged by the departments of Agriculture of Wales, Scotland and N. Ireland (this was the case for the evaluation of Hill Livestock Compensatory Allowances).

²⁸ These objectives have been explicitly articulated (in one side of A4), and the Guidance document reproduces them as an annex.

specific objectives. MAFF urges that evaluation should always try to assess additionally - the extent to which the policy has induced a reaction which would not otherwise have occurred; dead-weight effects – the extent to which the effect observed would have occurred in the absence of the policy; and side effects.

The MAFF quite explicitly include EU policies which come under their competence in their evaluation programme. In doing so, they only select aspects of the CAP which have significance for UK interests, but they do so with respect to all significant UK objectives (it may certainly be assumed that this includes general macro-economic objectives including EU budget contributions). There is clear reference to possible divergence between EU and UK objectives for policies. Thus for UK evaluation of CAP measures the MAFF guide (paragraph 23) advises that the following questions are posed:

What are the EU policy objectives?

To what extent do these complement UK objectives?

Does the UK have additional objectives?

Are there fundamental differences between the EU and UK objectives?

There is an intrinsic link between the objectives of a policy and how the policy is evaluated. It is recognised that Member States will have their own hierarchy of objectives. Evaluation will be conducted in terms of these objectives. The broad nature of EU objectives facilitates appropriate interpretation by Member States. This has a cascade effect on evaluation. For example, MAFF will assess EU policies in terms of the UK objectives. An assessment of the same policy by a different Member State will be conducted from a different set of objectives and hence the evaluation results could potentially be very different. Any form of EU wide evaluation would have to assess the balance of interests of the EU and its Member States.

The MAFF Policy Evaluation Guidance booklet identifies a particular sensitivity of evaluation of supra-national policy such as the CAP. Policy decisions in this area are made in a process of negotiation in which member state interests are an important, if not the most important, factor. It is clear that no member state will wish to reveal, in advance, its negotiating position. The results of its own evaluations of policy effects may well do this. Therefore, if MAFF consider that an evaluation of EU policy contains information which is sensitive for UK interests, then the usual presumption of contracting out the evaluation and the publication of the findings may be set aside. The evaluation will then be conducted in-house and not published immediately if at all. In such cases, the Policy Evaluation Guide offers the assurance that such evaluations will be subject to independent scrutiny.²⁹

When evaluations are completed, the reports are published (usually nowadays on the internet) and frequently they feed into a formal policy review procedure. The Head of Economics (Resource Use) Division reports the findings to MAFF's Management Board and the relevant policy division appoints an action manager who has to produce an action plan which is to be agreed within 3 months of the receipt of the evaluation report and implemented as far as possible within one year.

²⁹

In practice, so far, the only recent evaluation done in-house was on MAFF Staff Welfare Policy, and this was done internally for reasons of sensitivity to staff interests.

Information on future, current and past evaluations is publicised by MAFF on the internet and in hard-copy booklets, (MAFF 1998b). The booklet explains the process of evaluation, how MAFF selects contractors for this type of work, it lists evaluations expected to be done in the coming year, and lists current and recent past evaluations. It also contains the standard contract for such work and a specimen 'Expression of interest' form.

The documentation of 'completed research' lists 38 evaluations; 1 each in 1988, 1989 and 1990, 5 in 1991, 2 in 1992, 3 in 1993, 1 in 1994, 8 in 1995, 5 in 1996, and 10 in 1997. There are 6 evaluations listed as currently underway in 1998. These have been conducted by 22 different organisations, 10 University departments, and 12 consultancy organisations.³⁰

A rough classification of the policies evaluated is summarised below, with the number of evaluations done, or in progress.

CAP commodity regimes, or aspects of those regimes	7
CAP agri-environmental or forestry measures	16
CAP rural development measures	6
UK policy measures ³¹	13
Fisheries policy	2
Total evaluations 1988 – 1998	44

It can be seen that the predominant targets of evaluation have been the discrete programmes with relatively confined objectives. The main examples are in the area of agri-environment, rural development and, in the case of the UK policy measures, such matters as: pollution, animal health, disease control, provision of advice, and a new Agricultural Tenancy Act. There have been seven evaluations of what might be termed the mainstream commodity programmes (which have accounted for approximately 90% of CAP budget expenditure). One was on the 5-year set aside policy, two were on livestock premia, two were concerned with the arable area payments, one with the hops regime, and one with the operation of milk quotas.

MAFF *ex ante* evaluation is conducted on a less formal basis. It is done mostly in-house within the Economics Division, although some appraisals have been commissioned externally, for example options for the development of the Beef Special Premium Scheme and the Suckler Cow Premium Scheme, and for the EU hops regime. This work is seen as an integral part of the on-going debate on the CAP. The kinds of questions addressed include the following. How well is the CAP working from the UK perspective? What are the pressures for further changes in the CAP? What options for policy change are under discussion? How would those options affect UK interests? Which does the UK prefer? Which Member States are likely to take the same view as the UK and which will prefer other options?

³⁰ These numbers are intended to be roughly indicative of the spread of organisations involved. Some evaluations involve more than one organisation, and in two cases there is a close link between the consultancy company and the educational establishment.

³¹ Including MAFF internal management policy.

The appraisal (i.e. *ex ante* evaluation) techniques used range from extremely simple impact calculations at sectoral or farm level, through use of partial equilibrium economic models of specific sectors, to the (contracted-out) use of general equilibrium models.³² It would be a reasonable summary to say that most of this work involves cost benefit analysis or investment appraisals, which in the case of CAP reform, use quite complex partial equilibrium models. Most of such work is done under the time pressure of on-going negotiations. More thorough and longer-term analyses are conducted in the lulls between major policy reform debates. Most of this appraisal work is judged to involve calculations and information which it would not be in the UK interests to have in the public domain. Such work is subjected to (sometimes intensive) scrutiny by other UK departments, the European Commission and other Governments, but because it is not published, it is not subject to wider public or academic scrutiny.

In conclusion, the UK is constructing a solid foundation of procedures and experience in public policy evaluation. There is strong support for strengthening complementary procedures in EU institutions and for more co-operation in evaluation between Member States and EU institutions. MAFF economists were encouraged by the recent collaborative policy evaluation work published by the Commission (CAP reform proposals – impact analyses, European Commission 1998b). Although there was no MAFF or UK input into that study, it was recognised as a helpful coming-together of academic and official researchers to do *ex ante* evaluation. MAFF officials speak of the need for closer co-operation between Member States in evaluation work and can see the potential for developing an evaluation framework across the EU. It is recognised that any form of EU wide evaluation must be based on technical expertise in evaluation systems and would have to be insulated against becoming subsumed by the political and budgetary agendas.

4.4. Sweden

Sweden is one of the Member States with an advanced auditing system at many levels. This is a consequence of the nature of the Government. Public Administration is characterised by considerable decentralisation. Government ministries are small with only a few hundred employees. This is because ministries have little or no involvement with administrative matters. These are dealt with by agencies and state boards which benefit from a considerable independence of action from the state, even though their directors are appointed by the Government and their performance is monitored by the state. This agency system with performance monitoring has made evaluation a key element in the functioning of the state and its policies, and agricultural policy has not been exempt from scrutiny.

Sweden has gone through an extraordinary series of switches in agricultural policy in the last decade, and evaluation has played an important role at each stage of this process. Ten years ago, Sweden, along with Norway, Switzerland and Japan, had amongst the highest recorded rates of support to its agriculture in the world.³³ After an extensive period of national debate it was decided to liberalise agricultural policy radically, reducing domestic support and border price protection and replacing them with more direct support to farmers. Maintaining farming in the north of the country remained

³² Chapter 7 discusses the merits and problems of these alternative approaches.

³³ As evidenced by PSE and CSE calculations made by the OECD.

an objective of policy. However, before the transition from the old to the new policy could be fully implemented, the decision was made to join the European Union, which of course meant that Sweden had to adopt the CAP. This, in turn, necessitated a close study of the expected impacts on both Swedish agriculture and the Swedish economy generally. With this background, since joining the Union, Sweden has been at the liberalising end of the spectrum of opinion about the CAP. At each of these three stages, before liberalising, before joining the EU, and since joining the EU there has been extensive, close evaluation of both the existing policy and of possible alternatives. A particular Swedish tradition in conducting these reviews, or evaluations, is that the process is quite widely drawn. It has involved, of course, both the Ministry of Agriculture and the Swedish National Audit Office (RRV), but has also involved agricultural interest groups and independent researchers.

After the accession to the EU, Sweden made a strong effort to evaluate the future effects of the CAP at various levels, as well as a possible reform of the Common Agricultural Policy. The set-up of the reports is discussed under section 4.4.2 on the work of the Ministry of Agriculture.

4.4.1. CAP Evaluation by the Swedish National Audit Office

The UK's National Audit Office has provided a detailed description of the workings of the Supreme Audit Institutions of the European Union (NAO, 1996). The Riksrevisionsverket (RRV) is the central agency for auditing and accounting in Sweden and is subordinate to the Ministry of Finance. It is headed by the Auditor General, who is appointed by the Government for a six-year renewable term. The RRV is concerned with efficiency in the work of central government, and aims at high standards of accounting and financial management. The Swedish government often turns to the RRV for comments on measures proposed by public commissions and ministry-led inquiries. The independence of the RRV is enshrined in legislation. The RRV chooses its own subjects for investigation, and the type of investigation to be made. The general approach of evaluation of public policy measures stresses the familiar three Es, efficiency, economy and effectiveness.

The RRV does not have a department specifically working on agriculture, but has units which for a period of time concentrate on specific government departments and policies. The studies published are used as indicators for the Government for the performance of the institutions and may be used for decisions on the institutions' future budget and operation. The reports are not used systematically as in the UK to bring the officers of the institutions to account for their actions. Therefore, it is not an obligation for the RRV to collaborate with the audited departments. Nevertheless the RRV often has a representative in the Committees that evaluate the CAP and discusses issues with the Ministry.

Without documenting in detail the history of involvement in agricultural policy of the RRV, it is of most interest to note their current activities in agriculture. In 1997 the Swedish National Audit Office joined a parallel audit with the UK NAO and the Dutch auditing body (Algemene Rekenkamer) of the Arable Area Payments scheme which was introduced by the 1992 reform of the CAP.

The RRV decided in 1998 to follow up its earlier work on the Arable Area Payments (AAP) scheme. A 'Pre-study' was first conducted to "judge whether an exhaustive performance audit of the objectives of the EU's common agricultural policy (CAP) should or could be undertaken, which the study concludes to be feasible." This pre-study, which was presented as an example of how the RRV

conducts complex audits, was published in English.³⁴ The purpose of the pre-study was to decide whether it would be feasible to try and answer three questions. "Has the arable support – in combination with other forms of support to agriculture – resulted in a 'fair standard of living' for arable producers? How has the scheme affected the development of productivity in the Swedish arable sector? How has the introduction of the support affected consumer costs, either directly through prices, or indirectly through taxation to finance the support?"³⁵ Having posed these questions, the pre-study then discussed a number of possible indicators of each of these three objectives and assessed the data required for these indicators and its likely availability. The study concluded that all three questions could be addressed and suggested methodologies for doing it. It was then decided to go ahead with three studies, one on each of the three questions concerned with the effects of the AAP scheme on farm incomes, productivity and consumers. The deadlines for completion of these three studies were set to be November 1998, January 1999 and February 1999 respectively. The work was to be conducted by staff of the RRV, with help from external consultants, especially on some of the modelling required. The first study on income effects was indeed published (also in English) on time.³⁶ It is interesting to note that this study has been done in close cooperation with the Dutch and UK National Audit Offices, which in turn have analysed this subject. The cooperation sought the goal to improve and discuss the best methodologies and standards to adopt so as to make the studies comparable (see section 4.3.1).

The very fact that the RRV has undertaken these studies is most interesting. The work is a refreshing attempt to ask simple, but fundamental questions about the purpose of what is now a central element of the CAP, and whether it meets the declared objectives of the CAP, as stated in Article 39 of the Rome Treaty. Like in all evaluations there are issues that can be criticised. In particular there are problems in defining the counterfactual (how the situation would have been in the absence of the policy) and in taking into account side effects of the policy. These are, however, frequent problems in evaluations of this kind and should not deflect from the fact that this is in any case an encouraging effort by a national audit office to scrutinise the CAP's performance and effectiveness. It is highly commendable that RRV has taken this approach and puts Sweden clearly among the leaders in agricultural policy auditing.

4.4.2. CAP Evaluation in the Swedish Ministry of Agriculture

The consequences of introducing the CAP in Sweden raised the political importance of agricultural policy. This is reflected in the considerable increase in the staff numbers of the Ministry of Agriculture, even though the size of the ministry staff is still small compared with comparable ministries in other EU countries. The amount of internal resources available for evaluations is still very small, but this is a consequence of the structure of the Swedish public sector and reliance on agencies and committees. During the past years the ambitions to perform long-term CAP analysis are enhanced and co-ordinated by the Ministry, which organises and participates in evaluation exercises of the CAP.

³⁴ RRV (1998a).

³⁵ RRV (1998a), p. 7.

³⁶ RRV (1998b).

It is typical that for large evaluations a committee is set up, consisting of quite a wide selection of representatives from different relevant organisations. Usually the committee consists of members from the Swedish Board of Agriculture, from the Parliament, researchers, representatives of trade and farmers unions, as well as members from other governmental and private organisations.

The Swedish Ministry of Agriculture has been deeply involved in a series of studies on the Common Agricultural Policy, an exercise that began with the accession negotiations and has been reinforced since accession. In particular, the Government decided in 1995 to perform a number of studies concerning the future development of agriculture and the Swedish position on agricultural policies.

During the period of 1995 to 1997, the Ministry organised four Committees to perform the following tasks:

"Common Agricultural Policy and World Food Supply", dealing with the consequences of the WTO commitments and enlargement to the East.

"Effects of the Common Agricultural Policy", giving an *ex post* analysis of the effects of the CAP in Europe.

"Future Options for the Common Agricultural Policy", presenting alternative strategies for reform.

"Common Agricultural Policy, the Environment and Regional Policy", discussing the environmental consequences of the CAP and possible reforms.

The reports are in Swedish, but the Ministry of Agriculture produced a report on the studies which has been translated into English (Committee for Reforming the CAP, 1997). The studies' emphasis and objective is to quantify the effects of the CAP in Sweden, in a critical and objective manner. The research is performed with the aim of clarifying the Swedish position on the CAP, what it is and how it should be. The strong international orientation is particularly interesting. The studies do not concentrate exclusively on Sweden, but also on the rest of the EU and on international trade. In fact the first study analyses the CAP in the context of international trade, spelling out Sweden's standpoint on the WTO and the role Europe should play in helping the developing world by improving fairness in trade and reducing the incentive to overproduce and subsidise exports. Given the time frame and the importance of the report, its organisation is a good example of the practices of evaluation in Sweden as done by the Ministry.

At this moment the Ministry does not have a system set up to perform systematic evaluations of the CAP, and the scope of evaluation work is rather different from that of the RRV. Studies are still more directed towards the traditional impact analysis. This has much to do with the problems originating in the inconsistency inherent in the objectives of the CAP. However, there is an increasing awareness of the future importance of value-for-money, and there is a concerted effort to find effective and efficient environmental policies.

In 1998 the Government also produced a Bill presented to the Parliament setting up its views on the development of agricultural policy. The Swedish government is concerned not only with the agricultural sector when setting up its agenda, but has great concerns on the effects on the wider economy, e.g. consumers, environment and the world market.

What can be concluded is that Sweden is one of the EU members at the forefront of CAP evaluation. In particular it is performing important studies on applying as far as possible the Three Es criteria. The importance of evaluating agricultural policies is in any case increasing steadily. This is reflected in the planned establishment in 1999 of an independent institution in the area of agricultural economics to perform academic studies.

5. Defining the Objectives and Costs of the CAP

All public policy can be represented as collective action in pursuit of an agreed goal or objective. As was explained in Chapter 2 and exemplified in Chapters 3 (for the EU) and 4 (for several Member States), a critical part of evaluation is concerned with the specification of the objectives of the policy and looking at the costs of achieving those objectives. This Chapter considers objectives and costs for the Common Agricultural Policy. The Chapter has turned out considerably longer than the authors initially intended. Whilst superficially straightforward, the objectives of the CAP as implemented are a complex matter, the original objectives spelled out in the Rome Treaty have become overlain with successive adjustments and reforms with their own objectives. This complexity may provide one of the principal explanations for the endurance of a policy which has been widely criticised for decades (by all except its beneficiaries and some of those most directly involved in its implementation). Without a clear understanding of objectives, evaluation can offer very little. The intention is that this Chapter provides the basis for defining more concrete and specific indicators of objectives, costs and cost-effectiveness which is the subject of Chapter 6.

5.1. Objectives of the CAP

At first glance, determining the objectives of the European Union's agricultural policy seems to be a very straightforward matter. Article 39 of the Treaty of Rome spells them out quite explicitly, viz.:

- ◆ *To increase agricultural productivity by promoting technical progress and by ensuring the rational development of agricultural production and the optimum utilisation of the factors of production, in particular labour;*
- ◆ *Thus to ensure a fair standard of living for the agricultural community, in particular by increasing the individual earnings of persons engaged in agriculture;*
 - To stabilise markets;
 - To assure the availability of supplies;
 - To ensure that supplies reach consumers at reasonable prices.

Four sets of considerations conspire to make these seemingly simple objectives difficult to work with and in particular make the task of evaluation of the CAP a far from straightforward exercise. These four sets of considerations are:

- 1) other EU policies,
- 2) the interpretations and weights of the objectives in the Member States,
- 3) the changing weights given to these objectives, and new objectives, over time, and given the longevity of the CAP,
- 4) second-order objectives which have emerged as apparently important drivers of policy change. Each of these will be considered in turn, but it is useful to make one more general point before discussing these issues.

Policy objectives are a vital part of evaluation, and also these objectives are an essential part of the European Commission's initiation of policy. The Commission may only propose policy actions on the basis of the Treaties of the EU. Therefore, in all proposed regulations and directives, the Commission cites the Treaty basis for its proposals. This is invariably a statement of objectives. These are spelled out in greater or lesser precision in the whereas clauses of the proposed regulations. After the general Treaty objectives have been examined we will turn to the 'Whereas' clauses of the 1992 reform proposals and Agenda 2000 to see the current expression of objectives.

However the actual policy decisions are made by the Council (after due consideration of the opinions of the European Parliament and the Economic and Social Committee). We can only guess the objectives in the minds of the Ministers who constitute the Council when they make their decisions. They may be the same as those stated by the Commission – and spelled out in the regulations - but they may also be different. This is a subject in its own right. The decisions about the precise settings of policy variables (e.g. institutional prices, instruments of supply management, levels of payments) are the outcome of a bargaining process. Analysts can, and do, discuss what the objectives of the participants in such negotiations may be. This is the subject of the political economy of agricultural policy which has produced a rich literature on possible explanations of the behaviour of the participants to such policy decision processes (see for example De Goerter and Tsur, 1991; Fennel, 1997; Moyer and Josling, 1990; Petit, 1985; Petit *et al.*, 1987; Swinnen, 1994). The point here is to issue a general warning. Whilst formal evaluation of the CAP can only proceed on the basis of stated, agreed objectives of the policy, we should not be too surprised if the results of such evaluations do not appear to have a very strong connection with observed policy decisions. These are taken in a political process in which, apparently, the participants have their own political motivations and objectives.

5.1.1. The CAP and Other EU Policies

Europe's agricultural policy should not conflict with its other policies. The objectives of the European Union, stated in the Treaty of Maastricht Article B are, *inter alia*, "to promote economic and social progress which is balanced and sustainable...through the strengthening of economic and social cohesion...to assert its identity on the international scene..." These have been long term objectives of the European Economic Community spelled out in the Rome Treaty and whose application has been widened in the Maastricht Treaty for the European Union to embrace monetary and price stability and high levels of employment. Clearly there is scope for agricultural policy to be in conflict with these objectives. The public expenditure costs of agricultural policy may mean higher than otherwise taxation levels. The macro-economic effects of such taxation and the macro-economic effects of high food prices, and agriculture's labour and capital use, may have a dampening effect on economic growth. Thus the pursuit of agricultural policy objectives may have impacts on the achievement of these broader economic objectives.

Similar points may be made about Europe's trade policy objectives and other foreign policy including its policy towards developing countries. These are well known areas in which European agricultural policy (the CAP) can, and often has, come into conflict with broader objectives of the Union. For example, until the Uruguay Round of the GATT, the CAP was widely seen as a major obstacle to harmonious relations in the field of international trade.

In particular, restrictions on access to EU food markets and the subsidisation of agricultural exports can be seen to have detrimental (or in some cases, advantageous) economic effects on some developing countries. Such wider, if not higher, objectives should not be overlooked in any evaluation of the CAP.

A third category of other European Policies merits a special mention. This is the Article 130 (r, s, t) objectives of the Maastricht Treaty on Environment. This mandates that environmental protection requirements must be integrated into the definition and implementation of other Community policies. The objectives of Community environment policy shall be: "preserving, protecting and improving the quality of the environment; protecting human health; prudent and rational utilisation of natural resources; and promoting measures at the international level to deal with regional or worldwide environmental problems." Article 130r goes on to say that a high level of environmental protection must take into account "the diversity of situations in the various regions of the Community." Also that environmental policy shall be based on "the precautionary principle and on the principles that preventative action should be taken, that environmental damage should as a priority be rectified at source and that the polluter should pay."

This is an area of growing importance in the hierarchy of public policy objectives. With economic growth, man's activities impose more obviously on resource use and the environment. As populations have their basic needs satisfied they find they have the time, the information and the communications systems to become aware and concerned about the deleterious impact of their activities on the environment. Agricultural policy is seen as having two kinds of relevance to environmental policy. First, to the extent that agricultural policy encourages expansion and intensification of agriculture and to the extent that this has negative environmental consequences, then the CAP can be seen to be conflicting with Europe's environmental policy. Second, to the extent that protecting agriculture has, indirectly, helped preserve environmentally benign, or even environmentally helpful, farm practices and landscapes, then such measures may enhance the achievement of environmental objectives. Clearly both impacts should properly be part of any evaluation of the CAP. It is argued by many that the positive environmental role of agriculture, which is viewed as a classic market failure, may ultimately be a major part of enduring public support to land managers (see Buckwell *et al.* 1997). Evaluations systems in Europe should therefore be designed to embrace the environmental objectives of agricultural (or rural) policy.

5.1.2. The Interpretation and Weights of the Five Objectives in Each Member State

The meaning of the five objectives of Article 39 has been debated for four decades. Their durability is partly a testament to the wide scope for different interpretation between different interest groups and changes in their interests over time. A more precisely defined set of objectives for the CAP which suited the six founder members of the EEC in 1957 might quickly have become obsolete for the wider Community as time passed and circumstances changed. We now review the five objectives of Article 39 to examine what they mean.

Improving agricultural productivity is uncontroversial. No one can be in favour of inefficient or wasteful resource use. For non-economists, productivity is invariably defined in terms of technical efficiency. This is expressed in terms of ratios of outputs to inputs: for example, yields of crops per hectare, or milk per cow, or live weight gain of pigs per kilo of feed, or hectares cultivated per man

day. These are indices of average physical productivity of, respectively, land, cows, pigs and labour. Such indicators are relatively uncontroversial – although we shall see in Chapter 6 that devising indicators of technical efficiency which embrace many outputs and many inputs where both are subject to quality changes is not without its conceptual and measurement problems.

Economists argue that technical efficiency is necessary, but is not the end of the story. They argue that a productive or economically efficient firm or sector is one which satisfies the conditions of allocative efficiency. The use of this concept is controversial. Defining economic efficiency can be approached in two ways. The first approach is familiar to all students of economics in the ritual torture inflicted on them by their professors. They are taught the theoretical conditions for optimally combining each input and output pair, all pairs of inputs and all pairs of outputs. These conditions are expressed as equating, at the margin, a technical productivity ratio (output : input), substitution rate (between inputs) or transformation rate (between products) to an appropriate price ratio (product to input, two inputs or two outputs, respectively). The second approach is to describe the economic structures and circumstances in which rational firms will meet these efficiency criteria. Again, abbreviating a lot of economic theory, economic efficiency is likely to come about in competitive markets in which there are no government induced distortions.³⁷ Departure from these conditions then indicates likely departure from economic efficiency. This concept of efficiency, of course, applies economy wide.

There is a clear potential for conflict here. It has been the case that the instruments of the CAP, though supposed, *inter alia*, to stimulate agricultural productivity improvement (defined in terms of technical efficiency), have in reality focused on stabilising and raising prices of farm products. This has been done using a mixture of domestic price support measures like intervention buying and border measures such as variable import levies. Such measures will, almost invariably, contradict the achievement of economic efficiency. Of course, there are many other ways to raise productivity, which have much less negative effects on economic efficiency, for example through publicly assisted research, development, extension and training, or measures to assist the restructuring of farms.

An immediate tension in evaluating the productivity objective is that most farmers' organisations, and often Ministries of Agriculture too, will interpret this objective in the narrow sense of technical efficiency. Finance Ministers and the wider public interest are more likely to be concerned with economic efficiency.

Ensuring a fair standard of living for the agricultural community turns out to be an even more difficult objective to pin down. It has multiple interpretations, and thus there are potentially many possible indicators of its achievement. Defining 'fair', 'standards of living' and 'persons engaged in agriculture' each poses difficulty.

³⁷

This result is subject to some extremely important qualifications. First, that producers are profit maximisers, and consumers utility maximisers. Second, that there are no significant imperfections to markets and information. Third, that there are no externalities. Furthermore the achievement of economic efficiency tells us nothing about the underlying equity of the outcome.

There has been a long tradition in agricultural policy in considering fair to mean parity between earnings levels in agriculture and in other sectors of the economy. Even this is not so helpful, as it begs the question of the acceptable degree of parity.

Measuring living standards is far from straightforward. Farm households have three important distinctions from most non-farm households. On farms, the business assets and income is inextricably integrated with the household assets and income. Second, most farms in Europe have income sources in addition to their income from farming. Third, the farm family accommodation and some food are derived within the business. These features and the unincorporated legal form of most farm businesses pose severe problems both conceptually and in measuring the 'living standards' of farm families. There is an extensive literature on these issues, see Hill (1996) for a recent review of the state of the art in measuring the 'Total Income of Agricultural Households'.

There are also practical problems in defining the target group of the policy, 'persons engaged in agriculture'. One such problem relates to the definition of the lower and upper thresholds of the target group. Consider the lower limit first. Should this be defined in terms of some minimum land area, or value of output, or whether the products are marketed through formal channels, or the proportion of income derived from such produce? In other words, how do we distinguish between someone who has a large garden or a spare time hobby activity of keeping sheep or a few vines, from persons engaged in agriculture? In practice most countries use a mixture of these criteria. However, the Member States each have their own definition of what constitutes a farm business or an agricultural holding or a person engaged in agriculture.³⁸ There are also wide differences in the relative importance of waged labour on farms. Are farm workers the targets of agricultural policy? Are they treated identically to self-employed farmers? The prospect of Eastern enlargement of the EU, bringing in countries with significantly different farm structures, widens this problem. The applicant countries have both very small 'farms' and 'household plots' which often generate a significant proportion of national output of some commodities, and also very large farming companies or production co-operatives with large numbers of hired workers or co-op members. The political and economic status of peasants, farm workers, co-op members and shareholders in farming companies or absentee landowners are different to each other and different to typical family farmers. This is particularly an area where each member state will have different, and usually unwritten, objectives for these various groups of 'persons engaged in agriculture'. This poses a problem for evaluation of the achievement of the income objective.

Thus far, it has been argued that there is considerable room for different interpretation of what is meant by improving productivity, rational development of production, optimum utilisation of resources and improving the living standards of those engaged in agriculture. To make matters more complicated there are clearly great differences in the way the EU Member States seem to interpret the interaction of these first two objectives of Article 39. One interpretation, which could be characterised as neo-classical or Anglo-Saxon, puts great stress on the "thus" of the second indent. Their interpretation is that the intention is to assist farmers to increase their productivity and rationalise farm structures and *thus* to achieve better living standards. The main ways in which their

³⁸ These three concepts are all distinct, and there is no agreement in the EU Member States concerning their definition for official statistics.

productivity would be raised was expected to be through a combination of: better training and education, helping less productive farmers to leave agriculture, farm enlargement and consolidation so that each (remaining) farmer had a larger resource base, equipping each farm worker with more capital, and through the use of improved technology. All this would necessarily mean that the resource structure of agriculture would *and should* change. This approach to raising productivity and thus living standards, requires fewer people engaged in agriculture, each operating larger, more capital-intensive farms. In this way they would achieve a 'fair' standard of living, comparable to those outside agriculture. On this view, the market failure is the initial inappropriate farm structure and tenure arrangements, the low levels of education and training of farmers and the poor functioning of rural capital markets and perhaps the marketing of farm produce. The policy actions to deal with this specification of the problem are legislative and institutional actions to encourage farm restructuring, education, research, extension and training of farmers and farm workers, and institutional development to encourage the emergence of rural credit and produce marketing. Notice that there is no mention in this interpretation of supporting markets or farmers incomes.

The alternative interpretation, which could be described in short hand as continental, takes a more conservative perspective on the likelihood and desirability of resource adjustment in agriculture. In this view, farms start with such a low resource base that even with necessary and desirable improvements in technical efficiency, and with politically acceptable rates of outflow of people from agriculture, those remaining will not, in the foreseeable future, achieve acceptable living standards. Therefore to deal with the remaining low income problem of farmers, some form of support is deemed necessary. It is this prescription which leads to the policy instruments of supporting prices of products or giving direct income supports to these farmers. It is plain that this interpretation of Article 39 is the predominant view, as the principal chosen instruments of the CAP have been, and remain, price and income supports, even though it is reasonably clear that price and income support does relatively little, if anything, to increase productivity, in both a technical and economic sense.³⁹

Another way of characterising this dichotomy of interpretation is to focus directly on the number of people engaged in agriculture, and to ask, is the intention of the policy to reverse the outflow of labour from agriculture, to halt it, to reduce it or to accelerate it? Article 39 offers little help on this point. However, it is a mistake to assume from this absence of clear reference to whether the policy is supposed to speed the restructuring of agriculture or retard it, that society is indifferent to the matter. The problem is that society wants conflicting ends. Simultaneously, we want an efficient agriculture which will integrate with the sophisticated food industry to produce the dependable supply of a wide range of quality products in our food stores. Yet we are also sympathetic to the hardships of farming, the existence of rural poverty, the preservation of rural traditions, and we are concerned about the 'desertification' of the rural areas (which most of us are glad our predecessors left!). The resulting lack of clarity is a major problem in evaluating the CAP. Politicians apparently believe that an explicit pronouncement that the intention of policy is to reduce the number of farmers so that those who remain may enjoy satisfactory living standards is too brutal to be spelled out. We all conspire in this obfuscation. We prefer to speak of 'rationalisation, restructuring and adjustment assistance'

³⁹ In a way, one could well argue that the phrase *by promoting technical progress and by ensuring the rational development of agricultural production and the optimum utilisation of the factors of production, in particular labour* has a lot more to do with farm restructuring than with providing price and income support to a farming sector in a given traditional structure.

rather than 'farm business enlargement, the outflow of labour from agriculture and redundancy payments'. All the while society is unprepared to be clear about its intentions on this issue it is inevitable that policy will be sub-optimal.

An important purpose of evaluation in this situation is therefore to spell out the competing interpretations of objectives, to make explicit the implicit or hidden objectives and to provide indicators of the achievement of policy for each of these targets. Because the targets are in conflict, the policy will succeed for some but not others. Ideally, the results of evaluation will make these conflicts explicit and measure the social losses resulting from the non-resolution of the conflicts. In this way the greater transparency of policy and its effects may lead to more informed debate about the priorities for policy and thus to more rational policy.

The third objective of the CAP is to *stabilise markets*. This too may be interpreted in several ways. Does it mean to stabilise market prices, or receipts from the market?⁴⁰ Does it mean to take steps to avoid spikes of very high prices - to protect poorest consumers, or to avoid market price collapse - to protect producers? Does this refer to short run within-season instability or longer run instability? Are all markets to be stabilised or are some more sensitive than others? Because market stability is not self evidently a universally desirable goal this objective tempts consideration of the more fundamental objective which lies behind it. Again, there will be different weights given to these interpretations in the Member States. In countries where food expenditure is still a large share of total consumer expenditure, the inflationary effects of periods of high food prices are of more concern than in countries where agricultural prices have much smaller weight in consumer expenditure. Instability of producer prices will also have different importance depending on: the structure and size of producers; the extent to which they are dependent on a few products; and the extent to which there are regional concentrations of badly affected groups of farmers. There is no simple way to resolve these, often conflicting, aspects of stabilisation. Furthermore, because agricultural product markets are closely linked (consumers and producers will substitute between commodities depending on their prices and their relative stability), this makes it even more complex to achieve a desired stabilisation goal. Finally, it is also the case that achieving more price stability may have an impact on the average price level.⁴¹ There will be quite different perceptions about this trade-off as seen by consumers and producers, within and between Member States.

Assuring the availability of supplies is often referred to as the food security objective. At the most general level it seems uncontroversial. All governments regard it as their duty to pursue policies which offer security of the basic food supplies for the population. However the statement of this desire does not translate easily into an unambiguous policy. First, it is often, and mistakenly, assumed that food security means one hundred per cent self-sufficiency. However if producing all food at

⁴⁰ There are predictable relationships between price instability and revenue (or expenditure) instability depending on the source of the instability (production or consumption) and on the responsiveness (elasticity) of supply and demand to price.

⁴¹ *To stabilise markets* is often interpreted as implying the objective to raise the average level of prices. This is clearly not the case, as stability only means the absence of fluctuations. Price stability is, thus, achieved if the price is kept at the same average level around which it would otherwise fluctuate. However, measures that achieve such stability may (inadvertently) also affect the average level.

home and importing none means that animal feeds, fertilisers and fuel, for example, have to be imported in large quantities it is not very clear that such domestic food supplies are any more secure. Second, once countries have achieved economic growth which ensures that food consumption patterns and levels have long passed subsistence levels, it is far from clear that dependence on imports constitutes a meaningful threat to national nutrition standards or political stability. This is quite apart from the economic rationality of using domestic resources to produce food at high cost which could be imported at lower cost from abroad. Third, once countries have achieved the income levels and expenditure patterns of the Member States of the EU, it is plain that there is no real food security problem for the population at large. However, the problem may apply to specific, socially excluded groups within the population. This redefines the food security problem as an issue of malnutrition or poverty within Member States, and, as such, removes it from the competence of the EU and its agricultural policy.

The fifth objective of the CAP is *to ensure that supplies reach consumers at reasonable prices*. Judging what is reasonable may be interpreted in many ways. Economists point to opportunity costs. If food can be obtained more cheaply from a foreign source, then some justification is required to explain why consumers should be denied this opportunity. In the EU, the broader political judgement has been that the achievement of the other objectives of the CAP – productivity enhancement, living standards of farmers and stability and security of supplies – justify some sacrifice from consumers in paying more for their food. This is plainly an area of conflict between goals. Given that society subscribes to both goals, then the purpose of evaluation is to quantify as far as possible the impacts of chosen policies on the achievement of each goal to make explicit the extent to which one is sacrificed in order to favour the other.

It should also be pointed out that the consumer price goal of the CAP seems never to have been taken very literally. Agricultural policy has generally restricted itself to markets at the farm level. Price policy has intervened at wholesale level, but rarely further down the food chain. Yet as food processing, preservation and distribution technology has developed and as consumer tastes for purchasing increasing services (processing, preparation and convenience) with their food, the raw material, or wholesale price plays a decreasing role in consumer (i.e. retail) food prices. If this fifth objective of the CAP really was to mean reasonable *retail* food prices, then food policy would have to embrace much wider concerns than agriculture. For example, these considerations would take us into issues of food quality and competition policy which go far beyond the concerns customarily discussed under agricultural policy.

In short, the apparently clear goals of the CAP spelled out in the Treaty do not provide a clear basis for policy action. There are multiple goals for the CAP; there are several interpretations which can be made of each of these goals; and the relative weights given to them are likely to differ between the Member States. There are no available unambiguous statements of the interpretations and weights considered by the Member States. This makes it impossible to contemplate constructing a single index of the achievement of the goals (let alone the cost-effectiveness of their achievement). The purpose for evaluation (both *ex post* and *ex ante*) of the CAP should therefore be to attempt to express each of the possible interpretations of the objectives in measurable or recordable criteria. This information can provide the public and policy makers with the basis for judging the acceptability or otherwise of the policy.

5.1.3. The Changing Weights of Objectives Over Time, and New Objectives

Although there has been no change in the stated objectives of the CAP for forty years, the relative weights attached to them and their interpretation have certainly changed over time.

Agricultural productivity has improved and farming structures have changed. There has been a massive switch in the resource base of farming, a large outflow of labour, a smaller loss of agricultural land and an equally large inflow of capital. Whilst there will always be scope for further structural and technical change, the number one objective of the CAP can no longer be the drive to increase productivity. Indeed it is much more frequently the case in the 1990s that the public questions the need for any further technical progress in agriculture at all. This has shown up first in the concerns expressed about the impacts of modern technology on the environment, and second in debates over hormone use in livestock production and genetically modified seeds in crop production.

The concern about farmers' incomes and living standards persists but becomes no clearer. It might be expected that the trends to pluri-activity amongst farmers and thus multiple income sources of farming households would have diminished, or at least reshaped the debate on farming incomes. There is little agreement on this matter in the Union. However, the change in policy instruments pursued since 1992 have significantly changed the terms of this debate. Similarly, notwithstanding the confusion which surrounds the question of whether the CAP was supposed to stem or encourage the reduction of farm labour, the fact that the agricultural labour force has, in the meantime, massively shrunk, has switched emphasis away from agriculture, and farmers' incomes, to broader goals of rural development. These issues are discussed further below.

The emphasis given to market stability, security of supplies and consumer prices have also changed over the years. Implicitly, through the chosen principal instruments of the CAP from the late 1960s until 1992, great weight was given to stabilising the market prices of the major field crops, dairy products, beef, sheep meat and, to a lesser extent, pigs and poultry. These products were protected by systems of variable import levies, intervention prices and, where necessary, variable export refunds. There is little doubt that the result was both to stabilise EU prices for these commodities as well as to support them above international levels. Since the MacSharry reforms of 1992 and in anticipation of the Uruguay Round agreement of 1994, the simultaneous lowering of support prices, switch in border protection towards tariffs and the limits on the use of export subsidies, have the effect of exposing EU farmers to greater price volatility than before. Qualitatively, these policy switches could also be interpreted to imply reduced emphasis for maintaining a high degree of self supply, because of the agreement, in principle, to improve market access, and a greater concern for ensuring reasonable prices for consumers. However it would be a mistake to read too much into this suggested change of emphasis. The extent to which import access has been increased as a result of the Uruguay Round agreement on agriculture is slight (Tangermann *et al.*, 1997). Also there are few objective signs that the European Commission or Council have switched priority from farmer to consumer protection.

What undoubtedly has changed is the weight given to environmental protection within agricultural policy in Europe. Throughout the 1960s to 1980s the clear emphasis in agricultural policy was the production objectives. The focus on technical productivity (as defined above) and farmers' incomes, using price supports as the principal instrument, encouraged a sustained increase in capital intensity

of production and increase in output itself. It should be said that these structural and technical tendencies were common in all countries of the world during this time period, irrespective of the very different agricultural policies pursued.⁴² These changes, together with the expansion of EU membership, caused a switch in the Community's market balance for all major commodities from a position of net imports to net exports. The net impact of these developments on the environment was perceived as mostly harmful. As time progressed the side effects of modern farm technology on the environment were, first, alleged, then scientifically documented and then politically registered. The effects of concern were: pollution of soil, water, and atmosphere; reductions in bio-diversity both on farmed land and in the habitats surrounding agriculture; and changes in landscapes as field boundaries, water features, hedges and woodland on farms were removed.

The precise degree of causation between CAP instruments, farming systems and these environmental changes is complex and not well understood. However, in the context of the production incentives given to farmers and the resulting oversupply of the markets, it was clear that whatever the responsibility of the CAP in causing the environmental problems, it should surely bear some responsibility for helping to deal with them. The need to ensure that all EC policies paid due regard to their environmental impacts was first enshrined in the Single European Act in 1986. However it was not really until the 1992 CAP reforms, following the Treaty amendments of 1991 (Maastricht) that agri-environmental objectives for the CAP were first spelled out (see section 5.1.1 above). The so-called agri-environment accompanying measure spelled out the following general objectives in Article 1:

- accompany the changes to be introduced under the market organisation rules,*
- contribute to the achievement of the Community's policy objectives regarding agriculture and the environment,*
- contribute to providing an appropriate income for farmers.*

The same article then specifies what this Community aid scheme should promote to achieve this wider goals:

- a) the use of farming practices which reduce the polluting effects of agriculture (...);
- b) an environmentally favourable extensification of crop farming, and sheep and cattle farming, including the conversion of arable land into extensive grassland;
- c) ways of using agricultural land which are compatible with protection and improvement of the environment, the countryside, the landscape, natural resources, the soil and genetic diversity;
- d) the unkeep of abandoned farmland and woodlands where this is necessary for environmental reasons or because of natural hazards and fire risks, and thereby avert the dangers associated with the depopulation of agricultural areas;
- e) long-term set-aside of agricultural land for reasons connected with the environment;
- f) land management for public access and leisure activities;
- g) education and training for farmers in types of farming compatible with the requirements of environmental protection and unkeep of the countryside.

⁴² This illustrates a very important lesson for policy evaluation which is explained in detail in Chapter 7, namely that it is a mistake to assign simple cause and effect to changes in agriculture and agricultural policy pursued. Many changes in technology and structure have more fundamental causes than the agricultural policy measures employed.

This enunciation of environmental objectives for the CAP has further complicated the problem of the relative weights given to the objectives. It has made more explicit an important new dimension of incoherence in the CAP. One set of instruments, the price and structural policies encourage rationalisation and intensification of production and thereby, in part, induce environmental damage, while another set of instruments under agri-environmental policy offers assistance to mitigate or avoid this damage. If relative public expenditure on these two parts of the policy reveals the preferences of the EU's agricultural policy makers, it suggests that the productivity-farmer income objectives still far outweigh the environmental objectives in importance. However, these weights are slowly changing.

The other area in which new objectives for the CAP are emerging is in regard to rural development. The importance of agriculture in the rural regions diminishes over time as measured by its contribution to Gross Regional Product and to employment. At the same time, rural regions still display problems of general concern: low growth, lower income levels, poorer educational levels, and poorer access to health and social services. Rectifying such inequalities of opportunity is what is meant by striving for social cohesion. Whilst a few decades ago, supporting agriculture could be justified as a proxy for supporting the rural economy, this is much less obviously the case now. The component of the CAP which has dealt with these issues is the structural or so-called guidance section. The components of structural policy affecting agriculture and rural development are the Objective 1⁴³ (lagging regions) and Objective 5a and 5b⁴⁴ measures. It has been policy to give greater assistance to these objectives since the redesigns of structural measures in 1988 (Council Regulations (EEC) 2052/88 and 4253/88). This is indicated by the steady increase of the structural funds in the two periods from 1989-93 and 1994-1999. The revision of the Regulations in July 1993 for the period 1994-1999 (Council Regulations (EEC) 2080-2085/93) increased considerably the funding and also paid particular attention to the protection of rural areas.

The full regulations and commentary for the Structural and Cohesion Funds for the Period 1994-99 can be found in European Commission (1996b). This changing emphasis has, by no means, reached the end of the road. The discussion surrounding Agenda 2000 quite explicitly spells out that the importance of rural development should continue to grow within the umbrella of what is increasingly being referred to as the Common Agricultural and Rural Policy for Europe. Indeed, the proposed rural development regulation (replacing nine existing structural regulations – including the three MacSharry accompanying measures) is destined to become the second pillar of the CAP.

⁴³ Promoting the development and structural adjustment of regions whose development is lagging behind. The aim is to help in the development of regions at NUTS level II whose per capita GDP is less than 75% of the Community average. It also makes provisions for other regions which for their particular circumstances are eligible.

⁴⁴ Promoting rural development by:
(a) speeding up the adjustment of agricultural structures in the framework of the reform of the Common Agricultural Policy.
(b) facilitating the development of structural adjustment of rural areas.
5(b) areas are those not covered by Objective 1, but meet one or more of the following criteria: a high share of agricultural employment; a low level of agricultural income; a low population density and/or significant depopulation trend. This Objective has a wide range of aims from the diversification of activities in order to create non-agricultural jobs (e.g. Tourism and SMEs) to fostering telecommunications.

However, the lack of clarity and indeed tension between market policy and environmental policy referred to above, is mirrored here between market policy and rural development policy. What is Rural Development Policy supposed to do? Is its task to make agriculture more productive, to encourage farm value added and thus to assist the survival of farms? Or is the prime purpose to help stimulate alternative activities in rural areas to provide employment possibilities for farmers and their families. To date, the balance between these has clearly favoured the former interpretation – farm restructuring. Two developments are leading to a reassessment of this. First, the ‘over success’ of past price and structural policy, which created embarrassing surpluses, led to the situation that actions to stimulate farm restructuring were only eligible for structural support provided it could be demonstrated that it did not contribute to further over-production. Second, it is increasingly recognised that the (relative) output and employment contributions of agriculture to the rural economy are now very low, even in the most agricultural of rural regions. In this situation, it is much less clear that structural policy focusing on farming is the most effective way to achieve agricultural policy objectives. Low standards of living of resource-poor farmers are better tackled by stimulating other rural employment activities. This may enable them to remain on their land and become part time farmers, but simultaneously to enjoy living standards comparable to the rest of society. There are strong political statements made to indicate that the European Commission favours this evolution of policy objectives (see for example the speeches of the Commissioner for Agriculture and Rural Development, and the Cork Declaration, 1996). However, the balance of resources between market and rural development policy, and the balance of objectives spelled out in the proposed Rural Development regulation (European Commission 1997) indicate that market policy is still the favoured tool and farm development is still favoured over wider rural development. It is hard to dispute the notion that resource deployment decisions probably give a truer reflection of true objectives than political speeches and papers.

5.1.4. Second-Order Objectives of the CAP

The CAP has endured for so long, and has survived so many political and economic changes in Europe that it has acquired its own momentum. This has happened to such an extent that it has stimulated the extraordinary and convoluted situation that new, quite explicit, objectives have been formulated with the goal to reduce problems caused in pursuit of the original objectives! This creation of second order objectives emerged in the 1980s. Following the switch in the Union’s agricultural net trade position for major commodities from being a net importer to a net exporter, it became apparent that the growth of surplus production and of the associated budgetary cost were not sustainable. Thus a number of new regulations, and instruments, were introduced with the objective of directly acting against the effects of existing instruments. Three examples are offered. First, co-responsibility levies were introduced for cereals (Council Regulation (EEC) 2040/86) and milk (Council Regulations (EEC)1079/77 and 1822/77). These had the purpose of lowering production incentives for producers (which were simultaneously raised by the intervention prices) and to raise resources for the budget (which were depleted by the costs of intervention). Second, milk quotas were introduced in 1984 (Council Regulation (EEC) 856/84) with the prime objective of constraining the growth of milk production (despite the continued incentive to over-produce for the market provided by the supported prices). Third, maximum guaranteed quantities with linked price cutting mechanisms (Council Regulation (EEC) 2221/88) were introduced in the cereal sector to control growth in budgetary expenditure. These measures were not introduced because they were expected to have desired effects on the primary objectives of the CAP. The objective of these, second order, measures was to mitigate

the undesired side effects of the primary measures. A charitable interpretation would be to say that they were introduced in order to allow the continued application of the primary measures which *were* justified in terms of the original Treaty objectives.

In such situations, the real objectives of policy become even less clear. If instruments are in place which simultaneously encourage and discourage production and productivity growth, and which increase and decrease farmers incomes⁴⁵, it is difficult to determine what the real aim of the policy is with regard to productivity and farmers incomes.

To put it another way, as time progressed and as the effects of deep-seated and enduring policy themselves became an accepted part of the policy context, the real objectives subtly changed, and new instruments were introduced. The confusion stems from the fact that the old objectives and instruments were not removed, they remained.

This creates a difficult problem for evaluation. To this point it has been argued that there are five levels of complexity in identifying the objectives for the CAP:

- it has multiple objectives,
- the interpretation of these objectives differs between Member States,
- the weights given to the objectives also differs between the Member States,
- the weight given to these objectives has changed over time, and
- new, secondary, objectives have been introduced which countermand pre-existing objectives.

It was argued in Chapter 2 that policy objectives are central in the process of policy evaluation. If the objectives are unclear, incoherent or conflicting, then no amount of evaluation is going to be very helpful in guiding us to a conclusion on whether the policy has achieved its objectives and whether it has done so in a cost-effective way. This might, in part, explain why formal systematic evaluation has not been done in the Commission or the Member States for most of the period of the CAP.

The most startling reformulation of the objectives of the CAP came about in the 1992 reform and this process is being consolidated in the agricultural reform proposals which form part of the Agenda 2000 package.

The Commission's summary of the "primary aims" of the 1992 reforms were directed:

- "at improving balance on agricultural markets through more effective control of production,
- at increasing the competitiveness of European farm products through a three year programme of substantial price reduction but with compensation paid directly to farmers provided production was cut back via set aside,
- at encouraging increased respect for the environment;
- at facilitating some redistribution of support towards more vulnerable holdings;
- at maintaining a sufficiently high number of farmers on the land thus helping to preserve rural society" (European Commission, 1994).

⁴⁵ The instruments which do this are, respectively, on the one hand, price supports and on the other, production quotas and co-responsibility levies.

The first two, and arguably also the third and fourth, of these objectives are further examples of secondary objectives introduced in order to correct the undesirable effects of the instruments chosen in pursuit of the primary objectives. Markets which are not subject to government interference have no difficulty in finding their own balance. The imbalance referred to is of course the overproduction induced by the high price regime. The uncompetitiveness of EU exports, necessitating export refunds, is also a direct consequence of the price supports.⁴⁶ Regarding the third objective of the 1992 reform, responsibility for some aspects of the negative environmental effects of modern agriculture can be attributed to the CAP of the 1970s and 1980s. To this extent, this objective can be seen as correcting the effects of previous policy. However, a broader and more charitable interpretation is that it represents the formal embodiment of the environmental objectives of the Maastricht Treaty into the CAP, and a recognition of the positive environmental services supplied by farmers as joint products of their food outputs and which are not paid for by the market. The fourth objective introduces explicitly an important new dimension into the CAP – redistribution. This will be explored below. The fifth objective is an explicit statement that the purpose of support is to maintain the number of farmers. This is at least a welcome clarification of what hitherto was unstated. However it still leaves unresolved whether the target is to stop the labour outflow or slow it.

The most important aspect of the 1992 reform is undoubtedly the switch in the basis of much of the support from the old instruments of domestic price intervention and border controls to raise producer prices, to the new instrument of direct payments. This approach is being taken further by the Agenda 2000 reform proposals. By 1997, three-fifths of the budget costs of the CAP was devoted to the direct payments to cereal, oilseed, sheep, beef, olive oil and tobacco farmers. If Agenda 2000 is implemented as proposed, by 2006 this figure will rise to approximately 73% of FEOGA guarantee expenditure. This raises the question, what is the objective of these payments?

Their legal basis is quite clear, it is stated in the regulation introducing the payments and repeated in the above list of objectives. These payments are compensation for the cuts in intervention prices.⁴⁷ The compensation principle is common in twentieth century affairs. If some citizens are harmed by the effects of a change in public policy made for the wider social benefit, then it is reasonable to compensate the losers. It is also, of course, an important principle in welfare economics that to achieve an overall welfare gain it may be necessary to make some worse off. By offering compensation to this group, they may accede to the change in policy, which, understandably, otherwise they would oppose. If the basis for the payments is compensation, then of course the distribution of the payments amongst farmers should be to those who are hurt by the price cuts. This is mostly what has been done.

⁴⁶ There is a further conflict embedded in these new objectives. If the desire is to achieve market balance and international competitiveness – i.e. to avoid EU prices being held artificially above international price levels, then what is the purpose of ‘more effective control of production’? Supply management measures like set-aside and quotas are part of the apparatus which keeps prices above international levels. Paradoxically, they are short-term instruments, yet because of the notorious difficulty of removing them once in place, they tend to delay long-term adjustment.

⁴⁷ Regulation 1765/92 defines the purpose of the support system for certain arable crops, as follows: “to ensure better market balances by approximating the Community prices of certain arable crops to the prices of the world market and to compensate the loss of income caused by the reduction in institutional prices by compensatory payments to producers who sow such products...”

As the payments are paid in proportion to areas of Cereals, Oilseeds and Proteins (COPs) planted, and to numbers of beef cattle and sheep, it is to be expected that this will have the effect of distributing most of the benefit to the largest producers.

However, as part of the change in the basis of support, the 1992 reform also tried explicitly to redirect the distribution of benefits amongst farmers, this is referred to as modulation of the payments. This was done in two ways. First, compensation payments per hectare are calibrated at (regional) average yields. This has the effect of giving relatively more compensation to farms operating at below-average yields than those with above-average yields. Second, larger farmers (measured in terms of COPs output) were required to put some of their land out of production.⁴⁸ Two more elements of redistribution of direct payments are proposed in the Agenda 2000 measures, namely labour unit modulation and a system of degressive payment ceilings. If accepted by the Council and implemented these will mean that the most capital intensive and larger producers will receive only a proportion of the payments calculated on the basis of their crop areas and animal numbers. These elements of policy reflect the desire to correct the redistributive effect of the old CAP. It was a prominent part of the debate on the 1992 reforms that supporting farmers through instruments which raise product prices has the effect of providing most benefits to the largest producers. This was often referred to as the 80:20 problem: the Commission estimated that 80% of the benefits of price support accrue to the biggest 20% of farmers (in terms of volume of supported products). The redistribution measures are thus expressions of Objective iv listed above.

Introducing such an objective is helpful in that it makes explicit what previously was understood but not stated. At the same time it introduces further definitional difficulties. A prime purpose of the CAP has always been to support the income of poor people in rural areas – most of whom in the 1950s and 1960s were engaged in agriculture. The new objective makes it slightly clearer that the aim should be to help most those who are "vulnerable" (objective iv) and who presumably might quit agriculture and thus be lost to "rural society" (Objective v). Defining vulnerable will not be straightforward. Vulnerability depends on many factors beyond the level of income derived from farming (which is the only statistic assembled regularly by Ministries of Agriculture and the Commission). It embraces assets, the structure of the household, degree of dependence on agriculture, extent of indebtedness, and local opportunities inside and outside agriculture. It will not be easy to evaluate if the new payments are well directed to the vulnerable.

At a more fundamental level, and this is well understood by European politicians, there is an obvious trap which lies down the road of defining too explicitly the 'income' objective of the CAP. If the policy becomes essentially a system of social payments from taxpayers to farmers because they have low incomes, it begs two questions. What test of income is applied to such payments? Why are they paid only to the, occupationally defined, group 'those engaged in agriculture'? It is well known that the socially excluded and poor in rural areas include considerably more people than those engaged in agriculture. The trap is that if the CAP becomes the social security administration for rural areas, it loses its legitimacy. The competence for social policy is with the Member States, not the Union.

⁴⁸ The Commission actually proposed a stronger version of modulation of payments which would have reduced the payments to the larger farms. The Council was unwilling to accept this.

The Commission, and some, but a minority, of the Member States, are well aware of the force of these arguments. In response they are attempting to move the justification of the direct payments away from 'productivity... and thus... the individual earnings of those engaged in agriculture' towards adjustment assistance and payment for the delivery of cultural and environmental landscape services. The first of these is implicit, the second explicit.

It is clear that no group in society can be compensated forever for a once-for-all change in policy. Yet, few clear signs have been given to farmers that the compensation payments will decrease over time or may even cease eventually.⁴⁹ The notion of declining and finite payments was an explicit part of the new 'transition to market' payments introduced in the 1996 US Farm Bill.⁵⁰ There are only rather subtle, understated, signs of the same principles for the CAP payments. These are, first that the Commission proposed (during 1996 and 1997) to cut the cereal compensation payments. The arguments advanced were strong evidence of overcompensation in the early years of the cereal payments and the need to channel funds to the beef sector to deal with the crisis caused by the announcement of the possible BSE-CJD link. The Council rejected these arguments. Second, the payments are defined in nominal Euro and thus their real value will decay at the rate of inflation.⁵¹ Third, there is some signal in the Agenda 2000 proposals that full compensation forever for price cuts is inappropriate. In those proposals farmers are offered 50% compensation for the next round of cereal price cuts, and less than full compensation for beef and milk price reductions.⁵²

The issue is important because without clarity about the planned endurance of the payments it is difficult to evaluate their effectiveness. If the real purpose of the payments is to provide adjustment assistance to help farmers live in the world of unsupported prices, then the payments should be evaluated in terms of the extent to which the payments stimulate farm families to make adjustments and the success of such adjustments. In the meantime, if the political justification for direct payments is compensation, then evaluation must consider the extent to which former beneficiaries of price supports are being compensated for their income losses. In turn, recognising this as a second-order objective, it suggests evaluating the payments in relation to the primary objectives of the CAP.

These new directions in objectives and instruments introduced in the 1992 reforms were the start of a process which is likely to take many years to complete. Part of the problem is that the Commission

⁴⁹ Tangermann was one of the first to point out that converting price support to direct payments means that the (enduring) stream of support must be converted to a finite asset. His proposed mechanism was to offer farmers bonds, which could be used to provide an income stream or they could be capitalised and sold (Tangermann, 1991).

⁵⁰ There is nonetheless considerable scepticism that these payments will cease altogether after the expiry of the present Bill in 2002.

⁵¹ With inflation at the upper limit of the ECB range of 2%, it will take 35 years to halve their real value.

⁵² It might be inferred that a fourth signal given by the Commission that the direct payments will have to be further modified if not ended is that they will not be offered to farmers in the countries of Central and Eastern Europe. The grounds for this are that farmers who did not enjoy the high price protection do not justify compensation. It cannot be a permanent state of affairs in a Common policy that poor farmers in the East do not get the payments received by the richer ones in the West. However, it would be dangerous to read too much into this signal, not least because the Commission has further muddied the waters by referring to the compensation payments as direct *income* supports.

and Council judge that it is not politically helpful to spell out the destination of this process in very clear terms. It has become apparent in the debate leading up to the Agenda 2000 proposals that the process will extend well beyond 2006. There is a gap between the rhetoric of the reforms and the concrete proposals made. The rhetoric refers to a two-pillar model for the CAP. The first pillar is the redefined market policy based increasingly on direct payments to producers and less on domestic and border measures to support prices. The second pillar is the new rural development regulation – which embraces the agri-environmental measures. However, even if fully implemented as proposed, the bulk of the expenditure under the CAP will take place under the first pillar, the second pillar is allocated only 15% of projected FEOGA expenditure in 2006.

In terms of objectives, rather than instruments, the new approach stresses international competitiveness, market balance, environmental protection and preservation of rural society. This suggests quite different goals and thus criteria for evaluating the CAP than the Article 39 objectives. It suggests that as commodity support is phased out (to achieve international competitiveness and market balance) the residual, enduring, role for commodity programmes under the CAP will be concerned with stabilisation and perhaps, security of supplies. The main market failures which then remain to be the focus of agricultural or rural policy are supporting agri-environment and rural development. In time, as the direct payments justified either as compensation or as adjustment assistance are phased out, payments for these two activities will absorb the bulk of public support to rural areas.⁵³ When (and if) this point is reached, evaluation will have to focus on the cost-effectiveness of the delivery of these objectives. But this transformation of the CAP is a slow process. Before we arrive at this projected 'mature' integrated rural policy, there will be many years of operation of the Agenda 2000 model in which the dominating feature is the direct payment to farmers. The difficulty is determining how to evaluate the effectiveness of these instruments which are transitional in all but name. Should they be evaluated against the past (Article 39) objectives, against the present (compensation) objective or against the futuristic (adjustment assistance and agri-environmental) objectives? Another possibility is to evaluate the extent to which they facilitate (or obstruct) the evolution of policy through these stages.

Before leaving the discussion of objectives and turning to costs, a little more should be said about the objectives of the second pillar, the rural development and agri-environmental policies within the CAP. These are subjects in their own right and each supports a massive literature. Because most post-industrial societies are moving increasingly in this direction, the OECD (1988, 1996, 1997a and 1998a,b,c) has compiled much useful discussion and information on these areas of policy, including criteria or indicators of their goals. Considerably more work is required to bring these issues within the framework of international agreements.

It has first to be said that Rural Development (RD) and Agri-Environment (AE) goals can only be specified in detail at the regional level. These activities have widely different meaning and expression depending on the geographical context. Natural conditions – soils, climate, topography, and social and economic conditions vary so much that appropriate goals and measures will differ greatly amongst the regions. Thus objectives at the European level are necessarily defined in rather general terms. Implementation of these aspects of policy is the responsibility of the Member States based on

⁵³ The full arguments for this logic are contained in Buckwell *et al.* (1997).

partnership and a regional programming approach. This means that evaluation of the operational programmes has to be conducted for each region and with respect to the specific objectives set within the region.

The goals of rural development are spelled out in Commission documents such as the Future of the Rural World (European Commission 1988), the Cork Declaration (European Commission 1996c) and the Rural Development Regulation of Agenda 2000 (European Commission 1998). The objectives are broad, referring to income levels, employment opportunities, access to services and assistance to overcome the twin rural penalties of dispersion and remoteness. At the conceptual level, there are three main levels of difficulty for evaluation of rural development measures:

The interpretation and weighting given to the stated rural development objectives; there are many objectives which are not spelled out with precision.

The relative importance given to *agricultural* development as opposed to (wider) *rural* development; in other words, who are the intended beneficiaries?

The intended relative contribution of rural policy *per se* as compared to the rural impacts of general policy instruments (i.e. general economic, health, education, transport and social services policy).

EU environmental policy is based on Title XVI of the Treaty establishing the Communities as amended at the Treaty of Maastricht. The three articles 130r, 130s and 130t (see section 5.1.1) spell out the general principles. The direct goals of agri-environment policy stated in the regulations (see section 5.1.3) and in other EU statements can be summarised in four main points:

The protection of the natural resource base of agriculture (soil, water and atmosphere),

The conservation and enhancement of natural and semi-natural habitats in the rural areas,

Biodiversity conservation,

The protection and enhancement of rural cultural landscapes.

The first three are elements of what is usually defined as sustainability. The fourth has aesthetic, cultural and social elements which make up what we mean by the countryside and rural society. Each of these objectives is multi-dimensional. There are many difficulties of definition and interpretation, and all the familiar problems of determining weights given to each. This suggests problems for evaluation of agri-environment programmes. However, there is a strong offsetting advantage in this area which is evident in the considerable attention already given to evaluating agri-environment schemes. This is the existence of large, well resourced, well-informed and powerfully motivated environmental pressure groups in Europe. Environmentalists may not all agree on the desired elements of this part of policy and may also disagree about the weights of the components, but they are intensely interested in the outcomes of environmental policy and are willing to be part of the monitoring process.

Evaluation of the individual projects and programmes within the RD and AE policy poses one kind of challenge, another is to assess the relative balance between them. To illustrate, both the market measures (the direct compensation payments) and the rural development measures may be designed, partly, to help improve the earnings of those engaged in agriculture. Market policy has a direct effect

on the earnings from farming. Rural development policy may assist non-agricultural earnings of members of the farming household by catalysing the development of some other economic activities in the region. Thorough evaluation should be capable of measuring both contributions to incomes of farming households, and to offer information on the cost-effectiveness of each.

There are, unfortunately many areas where market policy, rural development and agri-environment policy are in direct conflict. To the extent that market policy encourages livestock densities in excess of the carrying capacity of the land, this can act directly against the aims of agri-environment measures. The European taxpayer can find he is paying farmers both to damage the environment and to rectify the damage. To pursue this clash one step further, consider the cross compliance proposal in the horizontal regulation in Agenda 2000 (European Commission 1998). The objectives of this measure are not very clearly specified.⁵⁴ The essence is that Member States are invited to lay down certain basic environmental standards, or codes of good farming practice, which must be respected as a condition for receipt of the arable direct payments. This seems very much like farmers receiving payment not to pollute, i.e. in direct conflict with the polluter pays principle. A different interpretation of the direct payments in this situation is that they are offered as 'temporary, adjustment assistance' precisely to provide farmers with the resources to adjust their technology to modern environmental standards.⁵⁵ Likewise, rural development policy may assist structural changes in farming which increase productivity and output, counteracting certain supply management measures which are designed to restrict output. Also some structural programmes, for example, to provide irrigation infrastructure can result in environmental damage which is the subject of agri-environmental schemes.

These new objectives express the same problems discussed at length above for the old Article 39 objectives. Where there are multiple objectives, and interpretations, the task of evaluation must be to state clearly and explicitly the analysts interpretation of the objectives and provide information on how far they are achieved.

Finally on objectives, two other considerations have appeared from time to time in discussions of the purpose of the CAP or reforms to the CAP, these are reducing the budget cost and simplification. In the context of evaluation, it makes no sense to consider these matters as objectives. Indeed it is absurd to have the policy of reducing public spending on a policy! The budget considerations about success of achieving the objectives of agricultural policy fall into the discussion of costs not objectives. Likewise, the current fashion for simplification is a clumsy way of affirming that the achievement of policy objectives should have an eye on the administrative costs of implementing the policy. The important thing is that policy implementation should be efficient and the associated costs proportionate to the benefits achieved. If a desired public objective is complex (as objectives of most agri-environment policy will inevitably be), then the measures and administration (the transactions

⁵⁴ Member States shall take the environmental measures they consider to be appropriate in view of the specific situation of the agricultural land used and the production concerned. These measures may include support in return for agri-environmental undertakings, general mandatory environmental requirements and specific environmental requirements constituting a condition for direct payments. Member States shall decide the sanctions... of not observing the mandatory environmental requirements. (...)

⁵⁵ For example meeting the nitrate, habitat or bio-diversity directives.

costs of the policy) may have to be complex too. Yet if conducted effectively and efficiently they can in principle represent sound policy. Simplicity *per se* seems an unhelpful consideration, and in some circumstances could prevent the realisation of objectives.

5.2. The Nature of Agricultural Policy Costs

The title and terms of reference of this study refer to the evaluation of public expenditure under the CAP. It is proper that a Parliamentary budget control committee should wish to scrutinise the way public funds are expended. However, there is a danger that focusing on the budget costs of the CAP might imply that these costs are the only costs which should be considered. The thrust of this section is to make the case that a full evaluation of the CAP should embrace a much wider concept of costs. Indeed, fundamentally the concept of costs is as wide as that of objectives, because costs are essentially nothing else than negative contributions to objectives, and hence it is to an extent arbitrary whether given policy effects are called objectives or costs – it is only the sign that matters. Yet, it has become customary to categorize certain policy effects as costs, and such costs are discussed in this section.

5.2.1. Economic Costs of Agricultural Policy

The fundamental cost of any economic activity is determined by considering the cost of the best alternative. This is the concept of opportunity cost. If there truly is no alternative to the present action or policy (which will be rare) then the opportunity cost is zero: the present action or policy is optimal.⁵⁶ This idea is the basis of economics. Resources are scarce, therefore any allocation of resources or any policy which influences the allocation of resources, causes these resources to be used in a particular way. The cost of this allocation or policy is the value of what could otherwise be produced.

These general ideas can be applied to consider the cost of the Common Agricultural Policy.⁵⁷ The commodity regimes of the CAP comprise a set of border taxes and subsidies, domestic price support arrangements, various rules on output levels, constraints on resource use for producing certain products, plus schemes of direct payments to the producers of a range of products. This creates a particular set of signals and inducements for using land, labour and capital and produces a given output – the gross agricultural product. The resources have alternative uses. If they were not retained in agriculture, they could be used by other sectors of the economy. Conceptually, the cost of this

⁵⁶ Using the phrase opportunity cost requires a certain amount of mental agility. First, because the words costs and benefits are sometimes interchanged; as in, 'the opportunity *cost* of being an MEP is the earning or *benefit* one could enjoy in the best alternative'. Second, the phrase is sometimes used to denote *total* benefit, and sometimes the *incremental, or marginal, benefit*, from the best alternative compared to the current action. Take, for example, a farmer who earns a 2% rate of return on the assets he has invested in his farm business. If the best alternative opportunity for this capital would earn a return of 8%, then the opportunity cost would usually be defined as the extra earnings foregone, (=6%), but it is not uncommon to find the whole 8% referred to as the opportunity cost.

⁵⁷ See Buckwell *et al.* (1982) for an exploration of these costs. This study was motivated precisely to widen the debate about the costs of the CAP beyond consideration of the budget costs and to point out that the costs could only usefully be defined with respect to some specified feasible, alternative policy. This section concentrates on the cost concepts, Chapter 7 examines the issue of defining the counterfactual policy.

policy is the output forgone by not allowing the best alternative use. It should be clear from this definition first that this concept of costs is general and it embraces much more than the budgetary costs of the CAP, and second, that defining and measuring costs involves defining the best alternative policy.

The full resource costs of agricultural policy on the economy at large show up in many ways. Most economic analyses categorise them into three broad types. First, policy may encourage the retention of resources in agriculture which would otherwise be used in other productive sectors of the economy. This applies in principle to capital, land and labour. Second, the policy influences wholesale prices of agricultural products directly (through border taxes and intervention price support systems) and indirectly (by output restricting measures such as marketing quotas and set-aside). These price effects influence the costs of production in the food processing and distribution sectors, and real incomes of final consumers. Third, there can be macroeconomic impacts of agricultural policy created partly through the sum of the above two microeconomic effects, and partly because of the direct effect of the large public expenditure on fiscal balance, and hence interest rates and even the exchange rate. A little more will be said about each of these three types of cost.

Resource use effects: capital. The concept that a supported agriculture will attract more capital investment than would occur otherwise is not very controversial.⁵⁸ Putting a value on this effect is more complicated. Conceptually, the effect can be measured as the output this capital could have produced in its best alternative use. However, it can be argued that once capital is invested in agriculture, its salvage value is low, hence its opportunity cost is low, and this is what should be measured. In the short run this logic is reasonable. Hence a thorough analysis will usually differentiate costs depending on the time horizon. In the very short run when no resource adjustment can take place these resource effects will be ignored. This is usually called an impact analysis of a policy. The short run in economics is the time period in which some factors can adjust, but asset fixity and low salvage values will result in low resource costs. When no time period is referred to it is usually understood that the analysis is being made for the long run, defined as the period in which all factors can be adjusted. This is sometimes referred to as the planning period when society can contemplate the best allocation of resources.

Resource use effects: labour. Similar considerations apply to labour used in agriculture. Here we speak of the immobility of labour. The extent to which farm labour, retained in agriculture as a result of agricultural support, could move into other productive sectors depends on many factors. If farm workers have highly specialist skills, of little use in other occupations, their opportunity costs may be low. If they have transferable skills – ability to work with sophisticated machinery, or skills in accountancy, management or marketing skills, then job mobility would be less of a problem. But the labour market has determinants far beyond these direct economic factors. Age, housing, remoteness, and information are all vitally important variables affecting the mobility of farm labour. In addition there will be less tangible, social and psychological, factors such as life style, proximity to family, and loyalty to place and community. These are the practical considerations which determine the mobility of labour. From the analytical point of view of calibrating the (economic) cost of agricultural policies,

⁵⁸ Farmers' organisations are often reluctant to accept this point, yet they readily argue the opposite case, that investment in the industry will suffer if the support is curtailed.

the important thing is their net effect. Just as with capital, the opportunity costs depend on the time horizon concerned. In the short run a different policy might have no detectable effect on labour use. Given more time to adjust, in the long run, existing farmers and farm workers will retire, and a less supported agriculture would attract fewer replacements than the outflow. This is the basis for economic calculations of the opportunity cost of labour retained in farming as a result of agricultural support.

Resource use effects: land. The land market raises quite different considerations based on the special characteristics of land: its supply is fixed, and that has implications on the attitudes to land occupation and ownership and thus legal arrangements for its use and transfer. For two reasons the opportunity cost of land, retained in agriculture as a result of agricultural policies, does not figure in most economic analyses. First, conceptually, as the supply is (more or less) fixed the return to land is a residual, called economic rent. Second, the rate of transfer of land to other uses (some of which have a rate of return that is massively higher than farming) is itself highly regulated by the land planning regulations of each Member State and is largely independent of the economic situation in agriculture.

The food price effects of agricultural policy show up in two ways. First, if farm policy has the effect of raising raw material prices for the food and feed processing industries, then this can be expected to have an impact on the international competitiveness, output and employment of these industries. Second, higher farm-gate prices will be transmitted through the food chain and must have an effect on real incomes in the economy. Given that agricultural produce is internationally traded, that there is relatively easy substitution between many foods and feeds and that policy has different impacts on different commodity prices, these price effects will influence the utilisation or consumption levels of EU farm output. A full analysis of the CAP would seek to embrace these effects. It would follow the downstream effect of agricultural policy on the food industry and it would also measure the consumer welfare impact of the high price regime.

Macroeconomic effects. Agricultural economists and politicians often overlooked the macroeconomic effect of agricultural policy. However since agricultural support has proved to be so resistant to change and has grown to such magnitudes – occupying the major share of the EU budget and providing a significant share of the value of agricultural output, it became clear that these impacts should not be ignored. Even though agriculture accounts for a small and steadily declining share of EU GDP and total employment, the magnitude of the distortions to resource use and prices caused by the CAP are noticeable. They have a measurable effect on macro-economic variables such as GDP, its growth rate, employment, the trade balance, price levels, the fiscal deficit and hence monetary variables.

It is one thing to cite these resource use, price and macroeconomic effects of agricultural policy, and another thing to measure or evaluate them. How this is done is the subject of Chapter 7.

5.2.2. Budget Costs of Agricultural Policy

Budget expenditure refers to the deployment of resources to achieve social aims through a collective (Member State or EU) decision process rather than leaving citizens to make their own individual resource allocation decisions. These expenditures are made from funds gathered from citizens through

taxation.⁵⁹ Because one of the prime tasks of the political process is to decide on the use of these collective resources, their deployment is naturally the focus of attention when examining public policy. But this focus on the budget costs of achieving policy objectives should not lead us into the trap of assuming that it is always desirable to try and minimise total budgetary expenditure. A corollary of considering wider economic costs as the appropriate concept of cost is that it allows the possibility that an improved policy may be one which requires higher (not lower) budget expenditure. If the only category of cost which could be considered in evaluating the CAP is the budget cost, then this solution could easily be overlooked. This was precisely the argument which lay behind the logic of the 1992 CAP reforms. Reduced intervention prices combined with full compensation of farmers for their gross revenue losses, were expected to reduce overall economic costs of the CAP despite the fact that the budget costs is higher. The same expectation of further reduction in overall economic costs, yet increases in budget expenditure, is the Commission's projected outcome of the Agenda 2000 proposals.

The definition of budget cost itself, has to be carefully defined. Public expenditure on a particular policy is not confined only to that part of the budget labelled specifically for that purpose. There are often indirect expenditure effects, or sometimes savings, in other parts of the public budget. Thus the budgetary effect of EU agricultural policy is not limited to FEOGA expenditure. It also comes in the form of, for example, tariff revenue (which reduces net expenditure on agricultural) or (EU or Member State) public expenditures on some environmental measures (cleaning up water polluted as a result of agricultural activity encouraged by the CAP), and so on. In addition, individual Member States would wish to include their own budgetary contributions to policy measures which are co-financed and their own public administration costs of the CAP.⁶⁰

5.2.3. Wider International Costs of Agricultural Policy

Because the European economies are relatively open and are fully integrated into the international economy, there are bound to be significant spill-over effects from Europe's agricultural policy into international markets. These are of concern because the EU is a significant economic, political and moral player in international affairs. They are also of concern because of the feed-back effects for the EU itself.

Broadly, the protectionist stance of European agricultural policy means that European agriculture is larger than it would otherwise be, and its non-agricultural sector smaller than otherwise. Its agricultural exports are therefore greater than otherwise and imports lower, and vice-versa for non-agricultural trade. Because of the relative size of the European economy, these distortions can be expected to influence world market prices for goods and services and thus to have economic effects

⁵⁹ In the case of the EU budget the taxation is based on the EU common external tariff, a share of member state value-added tax, agricultural levies and a fourth resource is based on member state GDP shares i.e. raised from general taxation in the Member States.

⁶⁰ Whilst it is not relevant to the evaluation of the *common* agricultural policy, evaluation of agricultural policy in each member state should of course include national policy measures and the corresponding national public expenditure costs – which, importantly for agriculture, include many tax benefits, e.g. favourable fuel and value-added tax treatment of farmers.

abroad. To be more specific, the subsidised export of farm products from Europe depresses world prices for these products. It means that consumers in countries which import cereals and dairy products from the world market, many of which are developing countries, are helped by this, and third country producers who export to the world market are hurt. It would not generally be expected that these third country benefits and costs of the CAP should be calculated and aggregated with the domestic costs and benefits in an evaluation of the CAP. Nonetheless, because Europe has a concern, and a policy, for international development, qualitatively at least, these impacts of the CAP are relevant in its evaluation. Europe's agricultural policy should be in harmony with its policy towards developing countries.

Maybe a more powerful, and certainly a more direct, impact of these 3rd Country effects are the economic feedback effects for the CAP itself. These effects are manifest through international prices. It was finally accepted in the early 1990s that the high price policy of the CAP was self-defeating. The more Europe produced, the more it exported with subsidy, hence the lower it drove international prices and the more expensive those subsidies became. Quantifying the size of this feedback effect is clearly an important part of assessing the costs of the policy. This means that analysis of the costs of domestic EU policy must embrace the economic forces in relevant international markets – that is the export demand and import supply facing Europe's farm product exporters and importers.

Apart from the economic feedback on the CAP over several decades there has grown a considerable political feedback effect. This culminated in the frontal assault on the CAP by Europe's trade 'partners' in the Uruguay Round negotiations of the GATT from 1986 to 1994. Although European politicians maintained the story that there was no formal link between these negotiations and the 1992 CAP reforms, it is plain that the international impact of the CAP was a very important consideration in the design of the new CAP. Likewise it is explicit in the Commission proposals for Agenda 2000 that the international effects of the new regimes are important ingredients in evaluating feasible policy change.

5.2.4. Environmental Costs of Agricultural Policy

The importance of the impacts of farming systems on the environment was discussed in sections 5.1.3 and 5.1.4 above. Environmental impacts of farming systems can show up as negative effects, or costs - pollution, habitat destruction and reduction in bio-diversity – or as positive effects or benefits – the creation or preservation of landscape features or semi-natural habitats. The classification of a particular effect as a cost or a benefit is usually clear and is generally accepted. However, it should be noted that this process is also somewhat arbitrary, depending partly on social valuation of what is an undesirable effect and partly on the definition of property rights.⁶¹ From the point of view of

⁶¹ The arbitrariness comes in as a result of the non-precision of definitions and environmental science. Defining an acceptable cut-off for a pollutant, e.g. nitrates in drinking water, involves difficult judgements about the concept of sustainability, about the science of the effects of nitrates on habitats and human health and about social evaluation of risks. The regulatory definition of the acceptable pollution threshold has to embrace these uncertainties. The definition of property rights also determines our view of these external effects and what is regarded as a cost (positive) and benefit (negative). In the case of aquifers and watercourses the property rights are assigned to society (i.e. not the farmer or land owner) thus we speak of the negative effect of farming on water. In the case of landscape, we say the property right, on how to manage the land, is privately owned by the farmer, hence his delivery of nice landscape is deemed a benefit and he must be paid to do it.

evaluation the important thing is that these effects are considered. It makes no real difference whether they are considered under objectives or costs. As far as this report is concerned, because environmental objectives have now been elevated to the status of a pillar of the CAP, then it makes sense to consider them in this way. There is no difficulty in principle in devising indicators which specify the reduction in the negative environmental effects as well as the increase in cost-effective delivery of the positive environmental effects.

6. Towards a System of Indicators of Effectiveness of the CAP

Conceptually, from each of the objectives of the CAP a number of indicators of success of that objective could be defined and measured. Likewise indicators could be derived for each concept of cost. Cost-effectiveness is then established by constructing and examining ratios of objectives and costs. Some objectives lend themselves to being expressed as achievement per Euro of cost, generally, the bigger the better. For others the effectiveness ratio will be expressed as the costs in Euro per unit of achievement, generally, the less the better. Conceptually, these effectiveness ratios are averages. For some purposes it is useful to have measures of the extra or marginal product of achievement of an extra Euro expenditure, or the marginal costs for an additional unit of achievement of the objective.⁶² Such calculations will be suitable for some objectives, but not others. In particular, these calculations can only be done where indicators of objectives and costs can be quantified. If this is not possible, then qualitative assessments must be made of the degree of satisfaction of the objective. Moreover, establishing ratios between individual indicators of achievement and individual indicators of costs makes sense only where the policy measure concerned involves only one (major) type of cost and is supposed to achieve only one (major) type of objective. In most cases, though, policies have several types of costs and achieve several types of objectives at the same time, and then there is the issue of whether and how to aggregate the different individual indicators of costs and achievements.

It is clear from the discussion of objectives in Chapter 5 above, that there are potentially a very large number of indicators of the large range of broad objectives for the CAP. It is neither desirable nor sensible to try and calculate them all for every evaluation. Also, the indicator choice will be highly conditioned by the purpose of the evaluation, and its audience. Because thorough evaluation is a costly process, an important part of the process should be spent in clarifying the most important questions sought and thus the most appropriate indicators. *Ex post* evaluation must pay attention to the objectives set in the past, whereas *ex ante* evaluation will put more emphasis on new objectives. In order to determine useful indicators for policy evaluation, it is necessary to have a good understanding of the mechanics of the policy instruments used, and of the structure and behaviour of the economic sector where they are expected to yield benefits. This is because the overall set of indicators to be used in evaluating a given policy should be both comprehensive, in terms of gauging all the important policy effects, and economical, in the sense of not trying to measure effects that are unlikely to be related to the policy concerned. Clearly the 'art' of good evaluation lies in judging what is important in such matters.

⁶² Calculation of marginal products or marginal costs require the estimation of the relationship between achievement (Q) and costs (C) and are then derived as the appropriate partial derivatives dQ/dC or dC/dQ .

The complexity of the CAP, embodied in hundreds of regulations and directives, covering tens of agricultural and food products and the rural territory and societies of 15 nations, is such that there is, in principle, an unmanageably large number of indicators usable for CAP evaluation. It is not possible, nor would it be useful, to attempt to list and discuss all of them. The approach adopted here is to contribute towards developing a framework for defining indicators of all the main areas in which the CAP tries to achieve some objective. If this framework is found useful, it could be the task of a follow-up to this study to assemble the complete CAP evaluator's toolkit by pulling together a more systematic and comprehensive check-list of indicators and their measurement methods. Section 6.1 offers a start in this process for the main categories of objectives and costs discussed in Chapter 5. The approach here is to list relevant indicators of these objectives which would be used in a strategic evaluation of the policy as a whole. In section 6.2 this broad list of indicators is sharpened to show how evaluation of *changes* in the CAP might be tackled. To help focus the discussion, this is done in relation to the 1992, MacSharry, reform of the CAP. The list of indicators suggested is not intended to be comprehensive, and other indicators can also be analysed. It is emphasised that this Chapter will not actually conduct the evaluation of the CAP or the MacSharry reform, but suggest suitable indicators for doing it.

6.1. Indicators for a Strategic Evaluation of the CAP

These will be listed with brief discussions offered about the sources of information, methods of measurement and some of the difficulties and controversies involved in their definition and measurement. They are presented in the same order as in Chapter 5.

Increasing productivity: rational development of agricultural production and optimum utilisation of factors of production, in particular labour

Technical productivity measures include average and marginal productivity for individual factors and products (yields per hectare of crops, yields and growth rates of animals). To deal with the situation of multiple products and factors it is customary to compile average indices such as output per Euro 100 of inputs or a more comprehensive indicator such as an index of total factor productivity. Constructing these indices requires data on outputs and inputs and their prices. The data required to compute such ratios and indices is reasonably widely available. The average ratios are routinely calculated by statistical authorities from sectoral level data. The broader indices such as total factor productivity require more detailed, farm-level, data on inputs and outputs. This is obtainable from farm surveys such as FADN. There is a lot of experience of this kind of analysis, though there are also difficult issues to resolve, for example, the appropriate definition of the index used, the weighting of the components, dealing with quality change, and the numéraire or comparator chosen. More general measurement of the rate of technical progress and its causal factors is much more difficult. There is a large literature on the measurement of productivity growth, but there is very little understood about how to disentangle technical and structural change and the relative contribution of the level and stability of farm support to this process.

Measuring the rational development of production and optimal use of resources can be interpreted to mean economic efficiency of resource use. This too can be tested through the classic optimality conditions of production economics. This in turn necessitates econometric estimation of production, cost or profit functions. Data for such studies can be derived from farm survey information.

Another angle to develop indicators on the rational development of agriculture is to trace the changing structure and resource use of agriculture. To the extent that an important, but not very explicitly stated, objective is to moderate, if not slow or reverse, the rate of outflow of labour from agriculture, indicators of the structure of factor use of the industry should be compiled. Ideally these should be regionalised to enable some analysis of rates of change in structure in relation to different levels and types of support.

It will become apparent in each of the categories of indicators discussed, that measuring, in a static sense, the 'achievement' level of a particular indicator of a policy objective e.g. the productivity of labour in agriculture does not constitute evaluation of the effect of the policy. Chapter 7 discusses in depth the necessity for some kind of counterfactual reference situation in order to be able to ascertain the extent to which any observed change in the indicator has resulted from the policy, as opposed to autonomous changes in the variable or the effects of other policies. This point applies to all the indicators discussed though it will not be stated for each.

Fair standard of living for the agricultural community, increasing the individual earnings of persons engaged in agriculture

There are extensive efforts made to measure the income from farming of persons engaged in agriculture. This is done at the aggregate sectoral level by individual Member States as well as by Eurostat. There are a number of aggregate indicators in common use, the most pervasive one being net value added per annual work unit. Other indicators are computed which make deductions for (actual or estimated) payments to capital, the labour of the farmer and his family and for land. The narrowest such indicator measures the management and investment return to the farmer. The sectoral level data are supplemented by information on farming incomes derived on a (more or less) common basis across the Union from the Farm Accountancy Data Network (FADN). However, the income concepts at farm level do not match those at sectoral level.

There are a number of well known deficiencies of these data as reasonable measures of the living standards of those engaged in agriculture, (see Hill, 1996, for a discussion of these issues). There are problems in consistently defining 'those engaged in agriculture', especially deciding the lower size limits. There are difficulties in determining the real labour input into farming of many part-time farmers. There are severe difficulties in knowing the other income sources of these part time farmers. This latter is plainly necessary if living standards of such people or families are to be assessed. There is also a debate on the correct treatment of appreciation of assets in assessing real incomes.

Judging fairness of living standards and individual earnings requires indicators of the distribution of the effects of policy induced income support. Thus data are needed not just on aggregate and average income levels, but also the distribution of income between farms of different sizes, types and regions. It is similarly necessary to have indicators of the distribution of benefits of policy – especially the direct payments which increasingly form the majority of farm support. Are these benefits distributed widely across all types and sizes of farms or concentrated on particular groups? Assessing fairness also requires comparator groups outside farming. Ideally these should be based on small, self-employed businesses with comparable turnover and risk.

Market stability

Interpreting this broadly as discussed in section 5.1.2, this requires indicators of price variability and income variability. Such data are relatively easily available from sectoral and farm survey data respectively. There are several statistical measures of variability from range to standard deviation. It is of interest also to know any skewness in the variability of prices and incomes, because downward movements are of more concern to producers than upward fluctuations, and vice versa for consumers. Drawing conclusions about the connection between policy and market stability will require comparative analysis of market stability inside and outside Europe.

At least one of the purposes of the stabilisation objective is to ensure the survival of 'normally' viable farms under extraordinary conditions. To evaluate this, other indicators of the degree of achievement of this objective would focus on instances where farms fail, i.e. insolvency, bankruptcy or other involuntary changes in the status of farmers. Arriving at useful conclusions about the role of policy will, almost certainly, require more data than simply the national or regional statistics of various forms of farm failure. Individual farm data and even case studies may be necessary to examine this dimension of policy.

Assuring availability of supplies

Availability of supplies is commonly, and usefully, measured through the construction of supply balance sheets for all the important commodities. These bring together information on opening stocks, domestic production, domestic utilisation for various purposes, exports, imports, additions to stocks and thus closing stocks. These sheets are constructed for the past and in prospect, based on expected production and projected utilisation. Supply balances are continually updated as the market situation develops and as information becomes available. Information of that type is clearly of importance for the detailed week by week management of markets particularly from the point of view of achieving market stability.

Assessing whether the CAP has contributed to longer-term food security raises quite different questions, and would require a completely different analytical approach and data to answer. It raises issues of the instability and reliability of the international food system. If providing greater stability and assurance of food supplies in Europe is at the expense of instability and less assurance outside Europe, then this may well have important feedback effects in trade and foreign relations which European policy makers would wish to take into account. This short discussion serves to indicate that supply balances and indices of self-supply alone are not very adequate indicators in the field of food security.

This information also gives only a crude, short term, indication about the 'security' of supplies. Stocks-to-use ratios are a normal first warning indicator of impending shortage. Such data tell little *per se* about the security of supplies. If there really is concern about the ability of the EU to feed itself in the medium to longer run under various scenarios, then this would require a quite different form of analysis and different indicators and data. Contingency analysis of anticipated threats would require information on, or assumptions about, imported raw material prices (fuel, fertiliser ingredients) as well as imported food prices. This information would have to be combined with calculations on the response of domestic production to these circumstances and food industry and consumer reactions to the different prices.

Ensuring supplies reach consumers at reasonable prices

To discover how well this objective is reached would require indicators of relative prices at all stages of the food chain from the farm gate through to the consumer. These could be available for international comparisons or for analysis of the actual policy compared to alternative price policies under the CAP. Such analysis can be done, although the further through the chain the more difficult it becomes. As the raw material is processed and mixed with other ingredients and services added (preparation, cooking, packaging, storage and distribution), meaningful comparisons become more difficult. Different margins, or farm-gate to retail price spreads, between commodities or between countries can reflect differences in the services provided, in the quality of produce, in the asset and cost structures of the food industry, in the risks and uncertainties involved, in the efficiency of the food industry, in the competitive structure of that industry, and in the tastes and relative values of consumers. It is notoriously difficult to draw firm conclusions from empirical analysis in this area, though this should not deter efforts to improve the data and techniques.

Competitiveness

The context in which this objective has emerged is that Europe found itself to be a major exporter of agricultural products with prices politically set at above international market levels. This clearly put European exporters in an internationally uncompetitive position. The obvious first indicator of competitiveness is therefore to compare prices of exportables in the EU with comparable products on the world market. This is the familiar 'price gap' analysis. Variants on this kind of indicator are nominal protection coefficients (these are measured as the ratio of domestic to international price, rather than the difference), and effective protection coefficients (ratios of value added per unit of the product which take account of price differences in inputs as well as outputs).⁶³ Moving up the scale to more all-embracing indicators of relative prices and protection or support offered to the domestic producers are Producer Subsidy Equivalents, PSEs. These attempt to measure the equivalent subsidy which would have to be offered to producers to make them no worse off if all supports were to be withdrawn. PSEs can be expressed in aggregate, per unit of the product and as a percentage of the value of the product. The OECD has put considerable effort into measuring PSEs for all major products and all OECD members. Strictly, none of these indicators measures competitiveness. To do this the concept of domestic resource cost should be employed. This measures the cost per unit of a product taking into account all taxes and subsidies and the real social opportunity cost of resources involved in its production, and compares that cost to the value added contained in that product, measured at world market prices of the output and the tradable inputs.

There are numerous practical issues in calculating all of these indicators. The fundamental data are prices. These should correctly reflect the time, place, quality and volumes transacted, and should be comparable – across products and countries. The latter comparison requires an appropriate exchange rate and should take account of transport, insurance and handling charges. These practical details are, in principle, tractable. More difficult are the conceptual assumptions embodied in the measures listed above. The most important of these is the implicit assumption that the international price provides an invariant yardstick against which the EU's competitiveness can be judged. Unfortunately, it is not reasonable to assume that international prices are unaffected by EU policy and the EU's net trade

⁶³ For a discussion of these indicators and the other indicators of competitiveness mentioned below, see Josling and Tangermann (1989).

position.⁶⁴ If the EU prices were to be cut and the volume of EU exports fell, this would almost certainly have a noticeable (positive) impact on international prices (all other things being equal). Expressing EU competitiveness in terms of the price gap is therefore tantamount to using a stretchy measuring rod.

It is tempting to respond to this by suggesting that competitiveness could therefore be better measured by examining the costs of production in Europe compared to abroad. This is a useful thing to do in its own right, but it does not circumvent the above problem. Measuring competitiveness by attempting to compare costs has its own hazards. Two will be mentioned. First is defining and measuring costs. Differences in farming structures and technologies, in accountancy concepts, particularly the treatment of capital depreciation, allocation of overheads and the treatment of farmers' own labour make cost comparisons difficult. Second, there is an interaction between policy pursued and costs. In a highly protected area such as the EU the benefits of support are capitalised into fixed asset values, especially land, and reflected in the rental cost of land. Moreover, high-cost farms are kept in business which might have left the farming industry already had price support not been provided. Therefore a static comparison of costs between heavily protected and unprotected producers will not say much about the long run competitiveness after input prices and asset values have had time to adjust.

Environment

At the EU level it is only possible to specify broad objectives for agri-environmental policy. There is such diversity in the physical nature of local rural environments from the cold north of Finland and Sweden to the Mediterranean Islands of Greece, and from the Atlantic West of Ireland to the Continental East of Austria, that it would be meaningless to try and define environmental indicators in common. The broad EU objectives have to be translated into more specific objectives at the regional level. Thus EU-level environmental indicators will therefore, necessarily, be rather broad. Reference has been made above in section 5.1.4 to the extensive literature on the environmental impacts of agriculture and the range of concerns and thus indicators of both environmental goods and bads. In the fields of environmental science and applied ecology there has been an enormous volume of research into indicators of environmental change. However, because this area of policy is so recent, there is a long way to go to develop practicably operational environmental indicators which can be built into the policy implementation, monitoring and evaluation process.

The most recent statement of these objectives is to be found in the Rural Development regulation of the Agenda 2000 proposals. In particular Article 20 (on agri-environment) in the Agenda 2000 proposals specifies that "agri-environmental support shall promote:

- ways of using agricultural land which are compatible with protection and improvement of the environment, the landscape and its features, natural resources, the soil and genetic diversity, an environmentally favourable extensification of farming and management of low intensity pasture systems,
- the conservation of high nature value farmed environments which are under threat,

⁶⁴ Economists refer to this underlying assumption as the small country assumption. A small country is assumed to be able to export any volume it likes and have no impact on international prices, a big country is defined as a country whose size relative to the volume of international trade is such that its actions will affect observed international prices.

the upkeep of the landscape and historical features on farmland, the use of environmental planning in farming practice.”

To deal with all these matters will require indicators of: resource protection for soil, water and atmosphere; of the extent and health of natural and semi-natural habitats; of bio-diversity presumably on both farmed land and in field boundaries and land surrounding agriculture; landscape features and landscape change; historical and heritage features.

In Article 13 referring to the proposed redefinition of support to farming in the less favoured areas as an environmental measure, there are references to:

ensuring continued agricultural land use and thereby contributing to the maintenance of a viable community,
preserving the countryside,
maintaining and promoting sustainable farming systems,
assuring environmental requirements.

Taking the first of these at face value requires indicators of land use and numbers and living standards of rural communities. For the other three objectives, the necessary indicators are the same as those developed to address the Article 20 objectives.

Rural development

Similar considerations apply as for agri-environmental indicators. The objectives at the EU level are broad, and thus the indicators will also be broad. The specific objectives for rural development will differ between regions and thus the more detailed indicators have to be developed by the Member States at that level.

The broad objective for rural development policy is to ”restore and enhance competitiveness of rural areas and, therefore, contribute to the maintenance and creation of employment in these areas” (sixth whereas clause of the proposed Rural development regulation, Agenda 2000). Indicators of rural development should therefore include the normal indicators of development, but measured specifically for the rural areas. These would include demographic features, the employment structure, income levels, access to transport, service provision, and social, education and health attainment indicators. One of the many difficulties of assembling and interpreting such data is whether the defined rural regions in such analyses include or exclude the towns within those regions. The greater and more extensive the connections (for employment, and service provision) between such towns and their rural hinterland the stronger the case for including the towns.

Within this broad objective and apart from the environmental objectives, there are many specific rural development actions each with their own objectives: investment in agricultural holdings (A4), setting up young farmers (A8), training (A9), early retirement (A10), improving processing and marketing of agricultural products (A23), forestry (A27), and promoting the adaptation and development of rural areas (A31). Indicators of each of these, with the exception of the last, are reasonably straightforwardly suggested by the specific objectives. The title of Article 31 sounds the broadest, but the precise objectives are unclear. Is the purpose mostly directed to *agricultural* development, as suggested by the reference to measures ”related to farming activities” or is this the intended main instrument for rural policy to extend beyond farming to the wider rural community as suggested by

the fifth (village renovation), sixth (diversification of activities) and eighth indents (development of rural infrastructure)? Without clear objectives it is difficult to devise clear indicators. The solution to this problem is presumably to devise as many indicators as possible embracing all these interpretations to enable each to be tested.

Economic and budgetary costs

As discussed in Chapter 5, the concept of policy costs is much wider than budgetary expenditure and includes in particular resource costs, but also elements such as international implications and environmental effects. It is therefore important that the cost indicators used in policy evaluation also go far beyond budgetary expenditure. Most of what needs to be said about indicators of non-economic costs, including environmental effects, has already been discussed in the preceding sections. However, a few comments should be added on how to measure and compare, through appropriate indicators, the resource costs and budgetary expenditure resulting from the CAP.

The resource costs of policies are closely related to the opportunity cost, i.e. the best alternative use, of all factors of production that are somehow affected by the policy measures concerned, as discussed in Chapter 5. More specifically, if the CAP draws resources (labour, capital, land, intermediary inputs) into agricultural production that could have been used more profitably in other sectors of the economy, then the value of these resources used in agriculture should be considered in measuring the cost of the CAP. Two conceptual issues have to be considered in this context, i.e. the interpretation of "profitable", and the appropriate measurement of policy costs. The concept of "profitable" in the context considered here has to be defined from the perspective of society on aggregate, rather than based on individual business (i.e. private) profit. For society, the "profit" of producing one extra ton of agricultural output is the difference between what society earns by selling that produce to the rest of the world⁶⁵, i.e. the world market price⁶⁶, and the (opportunity) value of the resources used to produce that output. The amount of resources that can be profitably employed in agriculture, from this perspective, is reached when the last ton of output produced requires resources whose value equals the world market price of that produce. Any production beyond this point uses resources that could have been employed more profitably in other sectors of the economy. The loss resulting for society, and hence the resource cost of a policy that draws resources into agriculture beyond this point, is the difference between the (opportunity) value of these resources and the value of agricultural output produced by them, valued at world market prices.

In graphical representation, the (opportunity) cost of producing one extra unit of agricultural output is measured along the supply curve (based on the assumption that producers equate marginal cost to price received). The cost of drawing, through policy measures such as price support, extra resources into agriculture, is then represented by the triangle formed by the supply curve, the world market price and the quantity produced at the domestic price. In economic analysis, the area of this triangle (appropriately calculated and expressed in monetary units) is used as the indicator of the policy costs in the form of resource misallocation.

⁶⁵ If the country concerned (i.e. the EU) is a net exporter. Alternatively the relevant product value is the price that would have had to be paid to the rest of the world for importing that ton of output.

⁶⁶ For a large country like the EU, the appropriate valuation is marginal export revenue (below world market price) respectively marginal import expenditure (above world market price).

On the consumer (or more generally user) side, misallocation can be measured in an equivalent way. Domestic food prices above world market prices make the consumer spend money that has an opportunity value higher than the real value (i.e. world market price) of the food bought, while buying less food than would have been the case had food been made available at the world market price. The monetary value of that policy cost can equally be measured by the equivalent triangle along the demand curve.

While these triangles measure true costs to society in the sense of a deadweight loss of aggregate welfare, there are closely related indicators which measure redistribution of welfare among different groups in society. In particular, changes in so-called producer and consumer "rent" indicate policy-induced effects on the well-being of these two groups. In graphical representation, the relevant measurements are the trapezium-shaped areas between the vertical axis, the domestic and world market prices and, respectively, the supply and demand curves. These areas measure the benefit to agricultural producers, and the burden on food consumers resulting from agricultural price support. Similar measures can be used to indicate the producer benefit resulting from other forms of agricultural support. In economic welfare analysis, it is in this context that budget expenditure (or revenue) also becomes part of the calculation, because it brings the third relevant group of economic agents into the picture, i.e. taxpayers. Roughly speaking, producer benefits (measured as change in producer rent) result from transfers originating either from consumers (measured as consumer burden in the form of change in consumer rent) or from taxpayers, measured as government expenditure. While these three flows (producer benefit, consumer burden, taxpayer burden) generally indicate a redistribution of welfare among these three groups, their balance at the same time is equivalent to the deadweight loss of aggregate welfare for society. This is because the sum of consumer and taxpayer burden is generally larger than the producer benefit. The balance is the loss in overall welfare resulting from the misallocation of resources caused by agricultural policy. In effect this means that some of the income taken away, through agricultural policy measures, from consumers and taxpayers never ends up in the hands of farmers as it finances the unprofitable use of extra resources in agriculture.

An implication of such considerations is that public expenditure *per se* is not a cost to society on aggregate, but a redistribution from taxpayers towards the beneficiaries of the policy concerned. This is particularly obvious for pure lump sum payments (e.g. completely decoupled compensation payments to farmers). Society on aggregate does not lose from such payments, because they do nothing else than transfer income from one group of society to another. There may, or may not, be good reasons for such redistribution, and policy evaluation should look into these reasons, and also analyse whether the payments are sufficiently targeted. However, for economists concerned with overall welfare of society, the nature of such pure transfers is something rather different from overall welfare costs, i.e. losses for society on aggregate, resulting from a misallocation of resources. These latter "true" costs can, of course, also result from public expenditure, but they are then an effect of the signals that some of the policies financed by public expenditure send to producers and consumers, and the resulting distortions of resource allocation.⁶⁷ In most cases these "true" resource costs are

⁶⁷ Moreover, taxation as required to finance budget expenditure can also affect (and distort) resource allocation. At the same time, public money may psychologically be considered more scarce than private money, and this can also add an element of "true" cost to public expenditure.

only a fraction of the budget expenditure involved in the policy.⁶⁸ In any case, as the economic meaning of budget expenditure is very much different from that of overall welfare costs, the two must not be added up.

As far as empirical indicators of all these elements are concerned, public expenditure under existing policies is easily gauged from the appropriate budget accounts. Public expenditure under alternative policies is not always easily estimated, in particular where complex agricultural policy instruments are involved. However, administrators have quite some experience in making budget estimates for future policies, not the least since they usually have to present them along with proposals for policy changes. Estimates of consumer burden and producer benefits, and hence of income redistribution, are also relatively easily done as long as quantities consumed and produced are assumed to remain unaffected by the policy measures concerned. However, this is precisely where the analysis can go wrong, as it is exactly because quantities respond to policies that resource costs arise. Hence for an evaluation of overall welfare implications it is unavoidable to make assumptions on (or econometric estimates of) supply and demand elasticities. Some comments on the analytical frameworks in which all these economic effects can be assessed will be made in Chapter 7 below.

6.2. Indicators for Evaluating Changes to the CAP: The Example of the MacSharry Reform

As said in the introduction to this Chapter, there is a potentially very large number of empirical indicators that can and should be used in practical evaluations of the CAP. It would go far beyond the scope of this study to establish a comprehensive list of such indicators. However, in order to provide some impression of what is meant by indicators, and how they are different, though derived, from policy objectives and costs, one particular example of a hypothetical policy evaluation will be used here to suggest an illustrative list of indicators for that specific case. The example used will be one of specific policy change under the CAP, rather than introduction or elimination of the whole CAP. Because of its political significance, and given its close relationship with the imminent Agenda 2000 reforms, the 1992 MacSharry reform of the CMO for cereals, including the introduction of direct payments and set-aside, will be used as this example. The idea is to suggest a list of indicators that could be applied in an evaluation of the extent to which this particular change in the CAP impinged on the objectives and costs of the CAP in general, and of the specific policy measures concerned. As far as the reference situation is concerned (see below, Chapter 7), it is assumed that the evaluation aims at a comparison of the post-MacSharry situation, say in the second half of the 1990s, with the situation that would have occurred had the CAP not been changed in 1992. In order to limit the scope of the analysis, only the changes originally introduced into the COM for cereals (Council regulation 1766/92) and the related compensation payments and set-aside in their original form (Council regulation 1765/92) will be considered. It is emphasised again that the purpose is not to conduct an evaluation of the MacSharry reform, but to illustrate the type of indicators that might be used in such an evaluation.

⁶⁸

There can, though, also be cases where the resource costs resulting from policies are much larger than the budget expenditure involved. An example are regulations that determine who can and who must not produce (a given quantity of) a given product. Such a policy involves no budget expenditure (apart from the administrative cost), but can distort resource use significantly.

As argued above, the MacSharry reform to a large extent was concerned with correcting some of the major problems that had resulted from past policies of relatively high price support, in particular for cereals. There was an obvious, though not explicit, relationship with the ongoing negotiations of the Uruguay Round. Broadly speaking, the MacSharry reform aimed at improving the market balance for cereals, i.e. to reduce the amount of surplus to be exported with subsidies. On the demand side the idea was that a price reduction should make cereals more competitive with cereal substitutes in feed rations. On the supply side, quasi-mandatory set-aside was introduced in order to reduce cereals output. Direct payments were added to the package in order to compensate the negative income effect of the price reduction. Quite obviously the MacSharry reform was a major departure from past strategies under the CAP, and further reforms along the new direction established by MacSharry are envisaged in the Agenda 2000 programme. The indicators suggested here should, therefore, also be of some use in evaluations of future policy adjustments which go in a similar direction.

Given the detailed discussion of CAP objectives and costs presented above, and what has already been said about indicators in the preceding section, it is not necessary here to explain how the individual indicators suggested are derived from CAP objectives and the whereas clauses in the Council regulations concerned. Instead, the allocation of the different groups of indicators to broad categories of objectives and costs will only be indicated by short-hand references in sub-headings. Only in a few cases brief comments on individual indicators are offered. For brevity, the indicators are suggested here only in their absolute form, though in an actual evaluation they would be used in the form of the difference between the actual outcome observed and the situation that might have occurred had the MacSharry reform not taken place (reference situation). Given that it is not the purpose of this study to serve as a technical manual, no references to statistical sources are suggested here.

Improved market balance. Major objective of Council regulation 1766/92. Derived from Article 39 objective of optimal use of resources.

- production and use (food and feed) of individual cereals
- gross imports of individual cereals
- gross exports and change in intervention stocks⁶⁹
- imports of cereal substitutes
- gap between EU and world market prices (reaching world market price level is explicitly stated aim in whereas clause)
- comparison of gap towards world market prices between individual types of cereals (whereas clause says that target price should be the same for all cereals in order to avoid bias of production towards individual types of cereals)⁷⁰

⁶⁹ The evaluator should, though, be aware that these variables are problematic indicators because they depend very much on market management, in particular the level of export subsidies set at any particular point in time.

⁷⁰ The economic rationale behind this "objective" is not really clear. It would make sense, as part of the evaluation, to provide an estimate of what the EU could have saved had it set domestic prices for all cereals to the same proportion of world market prices. With the same average level of EU prices for cereals, the EU price for wheat would then have been higher, and that for coarse grains lower. There would then have been more production, less use and more export of wheat, and vice versa for coarse grains. The result would have been a saving in expenditure on export restitutions, and higher overall economic welfare in the EU. See Koester 1982.

- monthly pattern of intervention purchases (whereas clause says that storage and interest cost of stockholders have to be covered)⁷¹
- monthly schedule of quantities exported and export restitutions paid over several years⁷²
- quantities of potatoes processed into starch
- export revenues forgone as a result of export tax/restrictions imposed in 1995/96
- area set aside in individual years and Member States
- area of voluntary set-aside (beyond required percentage)
- estimate of production effect of set-aside; comparison of that production effect with export chances/possibilities, given world market price developments and WTO constraints
- price developments on EU cereals markets, relative to intervention prices and considering world market prices⁷³
- levels of effective protection for first and second stage processed products (to see whether derived levies and restitutions for these products were set correctly)

Farm incomes. Major objective of Council regulation 1765/92. Derived from Article 39 objective relating to standard of living.

- actual prices received compared to intervention prices, in order to analyse over-/undercompensation, by member state (like analysis done by Commission)
- accuracy of compensation for selected farms with yields above/below regional average yields
- income effect of set-aside (based on sample farm budgets, considering higher payments for set-aside)
- income effect of special payments for durum wheat (based on sample farm budgets)
- income effect of price change and direct payments for oilseeds (based on sample farm budgets)
- income effect of price change plus direct payments on small producers, compared to what it would have been had they been treated like other producers
- income effects on farmers producing other agricultural products (beef, dairy, pork, poultry, potatoes etc.)
- number of farmers/farm workers who switched to alternative economic activities (to assess extent to which direct payments have effectively helped farmers to adjust)
- estimate of administrative cost for farmers and bureaucracy resulting from direct payments and set-aside

⁷¹ In this context, the evaluation could discuss why there is any need at all to make intervention purchases in more than one month towards the end of the crop year, and hence why stockholding activities during the rest of the crop year are not completely left to private agents.

⁷² This indicator provides information which can be used, in an extension of this evaluation, to analyse whether subsidised exports were made at the right moments. In this context the evaluation could discuss why there is intervention buying at all, rather than exporting surplus whenever it arises, at a domestic EU market price kept, through appropriately set export restitutions, at the level of what so far is the intervention price.

⁷³ In this context, the evaluation could analyse why it was that EU prices often remained above intervention prices even at times when export restitutions were paid.

FEOGA expenditure. Derived from macro-economic objectives.

- expenditure on intervention buying and export subsidies, minus import levy revenue, for cereals
- same for other agricultural products (resulting from cross-effects of cereal price cuts)
- expenditure on direct payments

Ove-rall economic welfare. Derived from macro-economic objectives.

- change in consumer rent, by product consumed
- change in producer rent, by product produced
- welfare triangles from partial equilibrium analysis, by product
- aggregation of these welfare effects across products (including budget expenditure)
- aggregate effect on overall EU welfare derived from general equilibrium analysis (see below, Chapter 7)

International trade relations. Derived from Article 110. To be stated only in qualitative terms.

- contribution to Uruguay Round negotiations
- effect on bilateral tensions over EU export subsidies with third countries
- implications for food importing developing countries, both in "normal" years and during the period when EU taxed cereal exports
- effect on stability/instability of world market prices for cereals

Effects on the environment and on rural development. Derived from objectives of Maastricht treaty, and the agri-environment regulation 2078/92.

Here a potentially large number of indicators could be used, as discussed above in section 6.1. The analysis would largely have to be done at the level of individual Member States and regions and would for the most part be limited to a qualitative nature.

Concluding comment

It is obvious that such a long list of indicators, which could easily be extended, provides a large amount of rather inhomogeneous information. Indeed, the multidimensional information generated may be rather contradictory. Some indicators may point to "success" of the policy, while others point to "failure". Policy makers and the general public may, on the other hand, have a preference for a clear-cut single message. Has the MacSharry reform been successful, yes or no? What was the ratio between achievements and costs of that reform? Was it cost-effective, yes or no? In other words, there may be a tendency to hope that the evaluator can aggregate the information provided by the various indicators into a small number of "supra-indicators" allowing an easy assessment of policy success or failure.

Is there any general way one can aggregate such indicators? The answer, unfortunately, is in the negative. The world has many dimensions, and policy makers, and the audience of policy evaluations, have to live with that, and effectively do so all the time. Essentially, aggregation into one single indicator of success or failure has to occur in a very informal, subjective way in the brain of the policy maker or his critic, and the unavoidable implication is that different people will assess success or failure differently. This does not say that there are not some subsets of indicators which can be aggregated, for example into measures such as cost-benefit ratios or cost-effectiveness (where costs of alternative ways of achieving a given objective, in itself not measurable in the same dimension as the cost, are compared). Such measures are very useful and should, wherever possible, be estimated

and provided. However, for a complex policy – and the CAP definitely is complex – such single-dimensional indicators cannot provide more than a very partial insight into the results achieved. A full picture requires a full set of individual indicators, and it will usually have brighter as well as darker spots.

If there is really a very strong interest in formal aggregation over all the various indicators of policy results, adoption of the following approach could be considered: All indicators are written into the head column of a table; in the first column of that table the outcomes for the individual indicators are ranked on a, say, -5 to +5 point scale, where -5 stands for massive failure and +5 for great success; the second column is used to insert weights for the individual criteria, say between 1 (only marginally important) and 10 (very important); in the third column the figures from columns one and two are multiplied, and in the bottom line the sum total is calculated. If that total is above zero, the policy is said to have been a success, and vice versa. This approach could also be used to compare evaluations of different policies. However, there should not be the slightest doubt that an approach like this cannot really solve the aggregation problem in a satisfactory way, because it essentially pretends to compare elements that are not really comparable in any formal sense. Moreover, there will rarely be two evaluators who agree on both the ranking and the weighting of the individual indicators, and hence an approach like this necessarily remains subjective. However, as a framework for discussing policy results in a structured fashion this approach might have limited merits.

7. Methodological Issues in Evaluating the CAP

The definition of policy objectives and costs, as discussed in Chapter 5, is of course fundamental to any policy evaluation because it says what the policy is supposed to achieve (objectives) and avoid (costs). For implementing policy evaluation in practice, one also needs yardsticks that can be used to determine empirically and in a rather concrete fashion the extent to which the policy objectives have been achieved and costs have been incurred. It is, therefore, useful to specify concrete indicators of objectives and costs that can be observed, as discussed in Chapter 6. When this is done, it may appear that the only remaining task then is to monitor the behaviour of these indicators over time, and on that basis to draw conclusions regarding the effectiveness of the policy to be evaluated. However, unfortunately one cannot short-circuit the process of policy evaluation in this way, because one also needs to establish the causal links between the policy concerned and the indicators observed. Was an improvement in a given indicator (say market balance) really caused by the policy to be evaluated (say a reduction of the support price)? Or was it coincidental? Has it possibly even occurred in spite of, rather than because of, the policy we want to evaluate? This is the type of issue to be discussed in the present Chapter.

Unfortunately this issue is rather complex, both conceptually and technically. Moreover, it is the type of issue where practitioners and theorists can easily disagree, both between them and amongst them. Theorists tend to demand an absolutely immaculate determination of the causal links between a policy measure and its effects, and suggest the use of potentially extremely complex tools of quantitative analysis in order to gauge policy effects empirically. Moreover, different scholars favour different approaches.

Practitioners of policy evaluation (and in particular policy makers themselves) tend to think that an intelligent look at reality tells us what policies have achieved, and that attempts at establishing cause-effect links are doomed to involve speculation that undermines the usefulness of their results. There is probably some truth in each of these positions, and it is necessary to steer clear of both oversimplification and unmanageable complexity.

In dealing with these issues it is unavoidable to start with a few conceptual issues involved in determining policy effects (Section 7.1). It is then necessary to consider where to rely on qualitative analysis and where quantitative analysis is necessary (Section 7.2). If quantitative tools of analysis have to be employed there is a wide choice of alternative approaches discussed in Section 7.3, and it is also necessary to consider the appropriate treatment of short term and external factors (Section 7.4).

7.1. Determining Policy Effects

Policies are introduced, or changed, at given points in time, and developments after that moment reflect the working of the new or changed policy.⁷⁴ In *ex post* policy evaluation, it may therefore appear reasonable to believe that changes over time of variables that are policy targets are a good basis to measure the effects of the policy concerned. For example, one of the major objectives of the MacSharry reform was to improve the market balance in the cereals sector. As far as the demand side was concerned, this was pursued mainly by correcting the price ratio between cereals and cereal substitutes, so as to make cereals again more price-competitive as a component of livestock feed. Indeed, after the MacSharry price cuts for cereals, starting in 1992, use of cereals in the EU, in particular feed use, increased noticeably, after a long period of stagnant if not declining cereal consumption in the EU. In the EU-15, total use of cereals increased from around 152 mio t in 1991/92 to nearly 170 mio t in 1996/97, i.e. by around 18 mio t. During the same period, feed use of cereals grew from around 91 mio t to 104 mio t, i.e. by 13 mio t.⁷⁵ At the same time, imports of cereal substitutes into the EU dropped by around 7 mio t. Based on these statistics, it appears there can be little doubt that the MacSharry reform was successful in achieving its objective of improving the market balance for cereals through making cereals more competitive in use.

There is no point here in questioning the "success" of the MacSharry reform in terms of changing the market balance for cereals. The 1992 reform has definitely made a major contribution to the significant growth in cereals consumption in recent years. However, it is also clear that such statistics cannot really tell what precisely that contribution was in terms of million tons. The simple fact is that statistics by themselves don't provide any information on causality, and hence cannot uncover effects caused by policies. In our example, statistics don't say what the development of cereals use in the EU would have been had the CAP not been reformed in 1992. Of course, if the MacSharry reform had

⁷⁴ Occasionally, policies can have effects even before having been introduced, resulting from expectations regarding their future nature ("announcement effects"). For example, the expectation that future compensation payments to farmers may be degressive with increasing farm size may lead some farmers to split their farms before the degressive payment scheme is introduced.

⁷⁵ Data are from European Commission (1997b and 1998a), with 1991 figures roughly adjusted to reflect the situation in the EU-15 (rather than EU-12 as provided in the Commission document).

been the only factor that changed since 1992, then all change that took place in reality could be completely attributed to the MacSharry reform. However, it is absolutely clear that many other things have happened since 1992 in addition to the MacSharry reform, and some of them might well have affected the volume of cereal use in the EU. Hence, even if the MacSharry reform had not taken place it is extremely unlikely that cereal use would not have changed.

Indeed, it might well be the case that consumption of cereals might have further declined in the absence of the MacSharry reform. EU beef production has fallen significantly after 1991, and this must have reduced feed demand. Feeding efficiency has continued to improve since 1992 and this too will have reduced feed use of cereals. More hypothetically, declining world market prices for cereal substitutes (in response, say, to lower import demand in the Former Soviet Union, or due to larger ethanol production and hence corn gluten supply in the USA) might have made cereal substitutes even more attractive for EU feed mills and could therefore have reduced their intake of cereals. In other words, had the MacSharry reform not taken place in 1992, then 1996/97 cereals use might have been even below that of 1991/92. Let us assume that in the absence of the MacSharry reform consumption of cereals in the EU would have declined by 5 mio t, rather than having increased by 18 mio t between 1991/92 and 1996/97. The actual effect of the MacSharry reform on cereals use would then have been 23 mio t, rather than the 18 mio t which are visible in the statistics. In other words, the "success" of the MacSharry reform might well be underestimated by looking at the development of this target variable ("indicator") over time.⁷⁶

Of course, there can equally well be cases where the "success" of a policy is overestimated by looking at developments of statistical indicators over time. For example, the long run growth of farm incomes per labour unit shown in the EU statistics cannot be taken to indicate the quantitative effect that the CAP has had on income development in EU agriculture. Even in countries that do not pursue CAP-like policies, farm incomes tend to grow in parallel with incomes in other sectors of the economy, as a result of labour market adjustments. Hence a large part of actual income growth in EU agriculture over the long run is due to factors that have nothing to do with the CAP.

Even more problematic, there can well be cases where the direction of a policy effect is opposite to that of the change over time of the target variable concerned. For example, the reduction of price support for beef under the MacSharry reform was hoped to reduce the surplus on the EU market for beef. However, a look at the statistics shows that the surplus on the EU beef market has increased dramatically in recent years. Of course everybody knows that this was due mainly to the BSE crisis. Had this drama not occurred, the surplus on the EU beef market might well have declined as a result of the price reduction implemented, and the MacSharry reform of the beef market was therefore probably successful in its own right. Indeed, it has probably helped to avoid an even larger accumulation of surplus. However, it is clearly impossible to read this directly from the statistics.

⁷⁶ Obviously, we do not really want to argue here that EU cereals use would have declined in the absence of the MacSharry reform, but simply use this hypothetical example to illustrate the general point that statistics don't tell policy effects.

More importantly, there can be many cases where the direction of change over time in a variable is opposite to the effect of the policy to be evaluated, because many other factors have counteracted the policy, and in such cases it is obvious that not even the nature of policy effects can be assessed by looking at time series.

Does all this say that it is impossible to determine the effects of policies? No, it simply says that policy effects cannot be determined by looking at the development of variables over time. In order to determine policy effects, there is no way around constructing, in some way, the counterfactual situation that would have prevailed in the absence of the policy concerned.⁷⁷ For the cereals example cited above, it is essential to establish what the post-MacSharry market balance for cereals in the EU would have been in the absence of the MacSharry reform if one wants to make a fair assessment of the "success" of the MacSharry reform for the EU cereals market. In other words, policy effects can only be determined through comparison with a benchmark, i.e. a reference situation that describes the world without the policy concerned. The crux, of course, is that the reference situation can only be hypothetical. This causes no major difficulty if the policy we want to evaluate is in the future. We are well used to imagining different sets of future events, one where the world moves on without a new policy that we are considering, and one where the new policy is introduced and has all sorts of effects on how the world will look like in future. Since both worlds are in the future, they are both hypothetical and there is no psychological barrier to comparing them.

However, if we look at policies introduced in the past, in *ex post*-evaluation, the problem appears to be different. One world is real, and that is the world in which the policy concerned has operated. We know this world, and we know that it has been shaped by the policy concerned. The alternative world in which the policy would not have been implemented is only hypothetical, i.e. counterfactual, and we find it difficult to accept that we should compare something real with something counterfactual. We therefore have the tendency to compare the real world in which the policy has operated with another real world in which it has not, i.e. the world as it looked before the policy was introduced. This is probably the reason why the before-after comparison is so popular in policy evaluation. However, it should be clear by now that this approach is based on a fundamentally flawed logic. As the policy we want to look at is not the only thing that has changed in the past, we cannot simply attribute all the changes over time in the variables we are interested in to this policy. The logical inconsistency of the before-after comparison is emphasised here so strongly because this is the single most frequently committed mistake in policy evaluation.

If it is accepted that there is no way to avoid constructing a counterfactual reference situation that would have occurred in the absence of the policy to be evaluated, then the next question is how to do this. Two approaches appear particularly attractive because they are strongly based on actual observations, rather than on speculation about how the world might look like under such and such conditions. First, two different parts of the world are compared at one point in time, where the policy exists in one of them, while it does not in the other (which is then taken as the reference situation). For example, if set-aside is not imposed on some farmers ("small producers") while it is required from

⁷⁷ The European Commission's guide on *ex post* and intermediate evaluation (European Commission, 1997a) makes ample reference to the problem of causal inference and the need for constructing a counterfactual situation, in section 4.1 on evaluation designs.

others, can we then not determine the yield effect of set-aside by comparing yields of small producers to those of all other farmers? Of course, we can do that only if there is reason to believe that set-aside is the only factor that distinguishes small producers from others. Unfortunately that is not the case (for example because farm size structure differs significantly from region to region in the EU, as does the quality of soil and a host of other natural conditions). In specific cases, where (nearly) everything is equal between two parts of the world, with the policy concerned being the only (noticeable) difference, this approach may work. However, such cases are rare.

Second, rather than making a simple before-after comparison, a more sophisticated approach is to look at trends of variables before the policy was introduced, and assume that these trends would have continued had the policy not been put in place. The trend extrapolation then serves as the reference case. For example, rather than comparing cereals use in 1991/92 (i.e. the situation immediately before the MacSharry reform) with that in 1996/97, the trend of cereals use between, say, 1981/82 and 1991/92 could be determined and extrapolated to 1996/97. Actual 1996/97 is then compared to this trend extrapolation, and the difference is attributed to the MacSharry reform. Of course this approach, though considerably more promising than a simple before-after comparison, still has to rely on the assumption that all other exogenous factors have continued to develop along their past trends, and that the policy was the only factor to change. This assumption may be much more justifiable than the assumption that nothing else changed except the policy concerned, as implicit in the before-after comparison. However, in many cases even this assumption is not convincing. The BSE crisis and its effect on the EU beef market is an obvious case in point.

In some cases it may be obvious why neither the two-parts-of-the-world approach nor trend extrapolation can provide appropriate insights. In other cases various elements of such approaches are mixed together in policy evaluation, and it is not always clear at first glance what (implicit) reference situation was used, nor whether the results are useful. Consider the following example. At some stage the need was felt to evaluate the agri-monetary regime as it stood at the time, where some EU countries had considerably higher prices than others. The background was that some Member States felt that the regime distorted agricultural production and hence trade within the Union. In particular, the question was asked whether positive monetary compensatory amounts (MCAs) and the consequent high-prices did not give a competitive advantage to producers in the countries concerned, allowing them to capture a larger share of the EU market. The analysis was done in the following way. The growth rate of agricultural production over a given period of several years in which the agri-monetary regime existed was calculated for each member state. Member states were then grouped into two categories, those with positive MCAs and those with negative MCAs. Finally, growth rates of production were compared between these two groups of countries. It turned out that both groups of Member States included countries with high and low growth rates of production. On that basis the conclusion was drawn that the agri-monetary regime did not distort agricultural production and trade of the EU Member States. Based on this result the Council of (agriculture) Ministers decided that the agri-monetary regime should be continued.

In this evaluation of the agri-monetary regime, the reference situation used implicitly in the analysis was an intricate mix of trend extrapolation (because growth rates were calculated) and the "two parts of the world" approach (countries with positive and countries with negative MCAs were compared). However, some thought should make it clear that the results of this type of analysis were bound to be misleading, for two reasons. First, there are so many differences between Member States in the

factors explaining the growth of agricultural production that it is simplistic to attribute any gaps in growth rates solely to the agri-monetary regime. Second, since the agri-monetary regime had been in existence throughout the whole period included in the analysis its major effect probably was to raise the level of production in the countries with positive MCAs, while growth rates may have been largely unaffected by the agri-monetary regime. The only analytical approach that could have yielded an appropriate answer to the question asked would have been to estimate what the levels of production in the individual Member States might have been in the absence of the agri-monetary regime, and then to compare actual production to that estimate.

In any case, in many instances neither the extrapolation of a trend (from a period where the policy concerned did not exist) nor the situation in a part of the world not affected by that policy can serve to establish the counterfactual reference situation with which current reality can be compared in order to determine the effects of the policy we want to evaluate. Where that is the case, some form of synthetic model, more or less sophisticated, may have to be used for constructing the counterfactual situation. Indeed, even pure guesswork (essentially based on a "synthetic" model in the head of the evaluator) on what the situation would have been in the absence of the policy concerned may be an option for establishing the counterfactual reference situation. In any case, it is important that the reference situation is made explicit by the evaluator.

The use of quantitative models in policy evaluation will be discussed in the following sections. The trouble of course is that use of models may not only be cumbersome and costly. It also is necessarily open to criticism because no model can ever claim to be able to describe unequivocally how the world would really have looked under the scenario assumed. Indeed, it will rarely be the case that two analysts agree completely on how the model should be constructed, and hence how reality would have looked in the absence of the policy to be analysed. For policy makers unhappy with the result of a given policy evaluation this unavoidably creates the possibility for them to argue that the model used in the analysis was wrong, and that their policy was much more successful than shown in the evaluation. There are certain methodological approaches that can be used to minimise this danger, and sensitivity analysis is the most important of them. However, for the simple logical reasons discussed above there is no proper way around using some form of anti-monde in policy evaluation, and hence there is always the danger of falling into that trap. But short of dispensing with policy evaluation altogether there is not really a way round this problem.

In addition to the fundamental logical issue of the need for constructing a reference situation, there is the more practical issue of how to define that situation with regard to the nature of the policy measures concerned. In particular, should the comparison be with a situation in which there would have been no policy at all, and market forces alone would have determined the situation ("no policy" case)? Or should it be with the old policies that existed before the policy to be evaluated was introduced ("no change in policy")? Or should a given alternative policy be used as the benchmark ("alternative policy")? Again, the benchmark is hypothetical in this regard, too, and its choice may result in credibility problems for the evaluator, in particular in *ex post* evaluations. Obviously the policy to be evaluated was introduced in reality, and one could argue that there was no alternative because that was the only policy on which decision makers could agree. However, if this argument were accepted, then there would be little point in any *ex post* policy evaluation.

Which reference policy is actually used in evaluation depends on the type of question asked. In some

cases, the "no policy" reference may be of interest. For example, if we want to learn about the effect of set-aside we can compare the actual situation with one in which set-aside would not have been imposed on EU farmers. Alternatively, if the interest is in evaluating the annual decisions on year-specific rates of set-aside, we can compare actual market developments with those that would have prevailed had the default rate of set-aside (17.5 per cent under current rules) been adopted all the time. Of course, for pragmatic reasons it often makes sense to define the reference policy such that it exhibits some (political and economic) realism. For example, in evaluating dairy quotas it would be pretty unrealistic to compare the actual situation to one in which the quota regime is completely abandoned without any reduction in the level of price support. Moreover, it is probably also realistic to assume that a large reduction in price support cannot be achieved without some degree of income compensation for farmers. In many cases of *ex post* evaluation, dealing with given policy measures introduced at a given point in time, a particularly interesting (and most often used) reference policy is that of "no change in policy", i.e. continuation of the policy as it stood at the time the new measure was introduced.

Though the choice of reference policy depends on the question asked in evaluation, fundamentally it can be argued that the "alternative policy" approach is the most appropriate. After all, there must have been a reason why the policy to be evaluated was adopted. Decision makers wanted to reach certain objectives that otherwise would not have been achieved, and hence felt the need to introduce new, or change old, policies. Had the measure to be evaluated not been introduced, an alternative policy would have had to be adopted. The trouble, though, is that the definition of the benchmark policy is open to debate. Indeed, in using the "alternative policy" approach, the evaluator can, deliberately or unwittingly, influence, and possibly distort, the result of evaluation through the choice of that alternative. Consider the following example. In 1991 the OECD produced a report aimed at providing "useful information on the possible implications of reductions in the level of agricultural assistance in OECD countries".⁷⁸ For dairy policies, (among others) two alternative ways of reducing the level of assistance (as measured by the producer subsidy equivalent, PSE) by 10 per cent in all OECD countries were compared, i.e. a 10 per cent reduction in market price support, and a 10 per cent reduction of milk quotas. It was found that the reduction in market price support would raise the world market price of dairy products by 8 per cent, while the reduction in milk quotas would raise the world market price by 28 per cent.⁷⁹ This result was widely publicised, and many commentators in the media concluded that a reduction in milk quotas was superior to a reduction in price support. There was nothing wrong with the analytics of this OECD study. However, the two policy options compared, though seemingly equivalent because both resulting in a 10 per cent reduction of assistance, simply were not realistic alternatives in most other regards, including their effects on producer incomes, consumer welfare, budget expenditure, overall economic welfare and longer run effects on the competitiveness of dairy producers. Hence there was not much point in comparing them.

Fundamentally, the most appropriate alternative policy to be used as a benchmark in policy evaluation is that policy which achieves the same objectives in the most cost-efficient way. In practice, though,

⁷⁸ OECD (1991). The citation is from p. 14.

⁷⁹ *ibid*, p. 69.

it is not, *a priori*, clear which policy does that, and hence the policy evaluator might have to analyse many alternative policies before being able to determine the most appropriate benchmark. Moreover, in many cases it will be difficult to find an alternative policy (or a set of measures) that achieves exactly the same objectives. Hence, choice of the benchmark policy to an extent necessarily remains an "art" where experienced evaluators need to take decisions which cannot mechanically be derived from a simple analytical framework. In any case, whatever the policy analyst defines as the reference policy, it is important that it is spelled out clearly so that the audience knows exactly what the scenario is against which the evaluation is made. It is also desirable that a justification be given for why this particular reference policy was chosen.

All these potentially controversial issues involved in *ex post* evaluations are, of course, also relevant for *ex ante* evaluation, though they tend to be less controversial in that case. As said above, the need for constructing a counterfactual reference situation is more easily accepted in *ex ante* evaluation, because the future effects of the policy to be evaluated are necessarily counterfactual as well. The need to define a (future) reference policy is also obvious in *ex ante* evaluation. Again in many cases the "no change in policy" approach is the natural candidate, as the alternative to the introduction of new measures often is the continuation of present policies. In some cases, though, existing policies have a built-in "sunset" clause by which they expire at a given date. In such cases the "no policy" case may be a more appropriate reference policy. For example, the current dairy quotas under the CAP expire in the year 2000. In an *ex ante* evaluation of the Commission's Agenda 2000 proposals for the dairy sector, it may therefore be appropriate to define the reference policy such that it does not involve dairy quotas. In any case, in order to determine policy effects appropriately, both *ex post* and *ex ante* policy evaluations require the construction of a well defined counterfactual reference situation, based on an explicitly assumed reference policy.

7.2. Quantitative Versus Qualitative Approaches

Apart from the conceptual issues related to a counterfactual reference case, there is also the more practical issue of whether some form of quantitative analysis has to be done in policy evaluation. In most instances that will be the case, and the methods available will be discussed in the following section. However, there are also cases where quantitative approaches are either not appropriate or not necessary. In such cases, qualitative reasoning may well be part of policy evaluation. Some cases of this nature will be discussed briefly in this section.

First, there is the issue of consistency between the nature of policy effects and the objectives pursued. In some cases, policies can have effects which are inconsistent with their objectives, and qualitative analysis is sufficient to show this because it is not so much a matter of the magnitude of the effects, but of their very existence. For example, the EU provides trade preferences, in the form of reduced tariffs, to various groups of third countries, in order to foster their economic development and strengthen political ties with the Union. In agriculture, such preferences are often limited to specific, usually small, quantities, in order to prevent them from undermining the sustainability of the CAP and to constrain their negative effects on farm income in the EU. Such limits on trade preferences imply the necessity to control the quantities that can be imported at reduced tariffs. In many cases this is done by issuing licences to traders. In the EU, such licenses for preferential imports are usually issued to importing companies registered in the EU. The reason given is that this makes it easier to monitor compliance. However, economic theory clearly tells, and empirical research of many such cases has

shown, that such licences confer extra profits on their owners (in economic jargon called "rents"), equivalent to the margin by which the tariff is reduced. In other words, even though the intention behind tariff preferences is to provide economic benefits to the exporting countries, their actual effect in many cases is that trading companies in the EU make windfall profits (which of course do not accrue to the exporting countries).⁸⁰ In cases such as this, qualitative analysis is sufficient to show that the actual effects of a measure are inconsistent with its objectives.

Second, in some rare cases of policy evaluation one may only be interested in the direction but not the size of policy effects. For example, a policy analyst might ask: if compensation payments for arable crops were to be decoupled from production, for example by being based on the area an individual farmer had in a past reference period, how might that affect land values? Economic theory predicts that this would let land values decline. If that is the only effect of interest in, then there is no need for a quantitative analysis. However, in many cases even the direction of a policy effect cannot be unambiguously predicted without quantitative analysis, because negative and positive effects occur at the same time, and the net effect is a matter of quantitative impacts. For example, how does withdrawing surplus fruit production by producer associations in the EU affect export chances of third countries shipping fruit to the EU? By taking produce off the EU market, export chances are improved. However, at the same time the prospect of higher average prices on EU fruit market may induce EU fruit growers to produce more fruit in the future, thereby making it more difficult for third countries to access the EU market. The net effect is not clear and can only be determined through quantitative analysis.

Third, there are many policies where neither the objectives nor the costs are easily quantified. Consider the case of co-financing CAP compensation payments by the Member States as recently analysed by the European Commission.⁸¹ It is clear that the objectives of that policy change, if it were adopted, would be to redistribute the budgetary burden among Member States. It can also be calculated, as done by the Commission, how much would be redistributed under any particular formula. Such implications have to be analysed in quantitative terms, as done by the Commission. However, the underlying objective of strengthening political cohesion among Member States is beyond quantification. Also, at face value no net cost would be involved in that policy change because the total amount of expenditure needed to finance compensation payments would not change. It is only the distribution among Member States that changes. However, there is little doubt that a policy change of this kind would have significant implications for the future of the CAP, and its long run costs. Yet, it is plainly impossible to determine such longer run effects in quantitative form. Hence a policy change of that nature has to be evaluated in qualitative analysis, even though some of its more immediate implications can be quantified.

On the other hand, there are many policies whose objectives are not easily quantified, but whose costs can be quantified. Many environmental policies are cases in point. Despite the efforts of environmental economists, which have generated estimates of the monetary benefits of some environmental improvements, it is still difficult or impossible to attribute quantitative values to many

⁸⁰ This is, for example, clearly the case for most of the agricultural preferences granted to the countries in Central Europe under the Europe Agreements. For a comprehensive empirical analysis see Overberg (1996).

⁸¹ Europäische Kommission (1998).

environmental objectives. Many other policy objectives will probably never be amenable to quantification, for example food safety. However, in many such cases the costs of achieving the objectives can well be quantified, and it is often possible to determine in a quantitative manner which policy measures can, or have to be, used in order to attain a given (qualitatively defined) objective. This is the domain of cost-effectiveness analysis, where the costs of alternative measures are compared that achieve the same objective (as opposed to cost-benefit analysis where benefits are expressed in the same monetary way as costs).

As a matter of fact, the indicator-approach outlined above in Chapter 6 also falls in this category. Such an approach is suggested because in most cases it is impossible, or not sensible, to aggregate all relevant policy effects into one aggregate benefit and one aggregate cost figure. However, this does not say that it is impossible or not sensible to use quantitative analysis for determining the effects that alternative policies have on the various indicators. Indeed, apart from the limited number of policy issues amenable only to qualitative analysis as discussed above, in most cases some form of quantitative analysis can, and should, be applied in policy evaluation. Some of the methods available will be discussed in the following section.

7.3. Methods of Quantitative Analysis

The agricultural economics literature is full of examples of quantitative policy analysis, including analyses of the CAP and its individual measures. It is impossible here to provide even a rough survey of the approaches available. Moreover, the most appropriate method to be used always depends on both the issue to be studied and the resources available. Thus only a few and necessarily highly selective observations can be offered here on how to decide on the most appropriate method of quantitative analysis for evaluating EU agricultural market policies. We shall start with simple techniques and progress to more sophisticated tools.

As discussed above, in *ex post* evaluations practitioners sometimes have a tendency to adopt the "before-after" approach, where changes over time in the relevant indicators after the policy was introduced are attributed to the policy concerned. For the reasons mentioned, this approach is rarely satisfactory. However, there is another still relatively simple technique, also mentioned above, that holds more promise, i.e. extrapolation of past trends before the policy concerned was introduced. The resources required for such analysis are relatively limited, though the difficulties of making appropriate trend extrapolations should also not be underestimated. What is the appropriate base period on which to base the trends? Are data easily available for that period? How should changes in data definition be dealt with, or changes in the composition of the variables observed (for example, enlargement of the EU by East Germany in 1990 and by three new Member States in 1995)? More fundamentally, quantitative analysis based on trend extrapolation is possible only if the benchmark policy (assumed for the reference situation) is the continuation of old policies, i.e. the "no change in policies" case. Evaluation of current policy against some alternative policy that was not pursued in the past cannot be performed on the basis of trend extrapolation.

Another simple technique of quantitative analysis is the use of accounting frameworks. For example, in order to assess the impact of the cereals price reduction under the MacSharry reform on farm income in given farm groups and regions, farm budgets (based, for example, on FADN data) for a pre-MacSharry period can be used, where actual prices recorded (or implicit) in the data are replaced

by the post-MacSharry prices. Similarly, the impact on budget expenditure for export restitutions can be assessed by multiplying the quantities exported by lower restitutions derived from the reduced prices. The effects of policy-driven price changes at the farm or wholesale level on consumer prices, for example, can be assessed in an accounting framework based on the assumption of constant processing and marketing margins. Such accounting frameworks are relatively easy to establish and have the great advantage of not requiring assumptions on the behaviour of market participants. They are versatile in the sense of being amenable to evaluating given policies against all sorts of reference situations. If the resources available for policy evaluation are limited, quantitative analysis, where necessary, will often have to be based on this type of handy tool.

However, it is clear that accounting frameworks can only yield impressions of the first-round impacts of the policies concerned, because they (implicitly) assume that everything not directly determined by the policy remains constant. For example, the use of farm budgets to calculate the income effects of price reductions assumes that quantities produced do not respond to the price changes. The assumption of no response in variables not directly determined by the policy concerned may be reasonable when assessing the immediate effects of a policy change⁸², but it becomes less reasonable the more time there is to have effects on other variables, including quantities produced and consumed. Such responses can either dampen or magnify the first-round impacts of policy changes. In the case of farm incomes, adjustments in the structure of both outputs and inputs are likely to dampen the first-round impacts of price reductions. In the case of export restitutions, savings will be magnified because the quantities exported decline as a result of price reductions.

It is because of such responses of variables that are not directly determined by policies that models play an important role in policy evaluation. Models essentially try to reflect the behaviour of reality in a much simplified and rather partial way. In economic analysis they are a weak but indispensable substitute for what experiments can do in natural science. If it were possible to "play", say, the year 1996 again, without the MacSharry reform, everything else being exactly equal, then we would not need a model to find out what the effects of the MacSharry reform were, and *ex post* policy evaluation would be much easier. If we could "play" the year 2005 twice in our laboratory, with and without Agenda 2000, then we could exactly say what the effects of that policy reform will be, without the need to resort to synthetic models in *ex ante* policy evaluation. However, such experiments are (fortunately) impossible in economic policy, and hence there is no way around using models in much of policy evaluation.

Models usable for policy evaluation come in a wide variety, from strikingly simple to unbelievably complex. In a way, trend extrapolation also is a "model" because it hypothesises that variables would have continued to behave over time as they did in the past, had it not been for the change in policies. The simple market diagrams with supply and demand curves that economists like to draw on the back of an envelop to visualise the effects of policies are models, and if drawn to scale they can even be

⁸² In many cases, market participants respond immediately to policy changes, and then such first-round impacts are never observed in reality. This is particularly the case when policy changes are announced before actually taking place, like often under the CAP (for example in the case of the MacSharry reform).

used for simple quantitative analysis. More typically, though, models consist of sets of mathematical relations with parameters thought to describe certain aspects of the behaviour of real-world people such as farmers and food consumers.

The models typically used in agricultural policy evaluation broadly speaking fall in three categories, i.e. farm models, market models and models of the overall economy. Farm models try to capture the behaviour of individual farmers when deciding which inputs to use and which outputs to produce. A typical example is a mathematical programming model that describes how a farmer allocates his labour, land, buildings and machinery to the various products he can produce, in order to obtain maximum profit. Many behavioural characteristics can be introduced into the model, and sometimes the link between the farm as a production unit and the farm household as a consuming and saving unit is explicitly included. Models are often filled with real-world data from farm samples and hence reflect "typical" farms in a given type of farming system, farm size and region. Levels of aggregation differ. For example, a whole region may be represented by one farm model based on the average characteristics of farms in that region, but there may also be several models for the major types of different farming systems and sizes operating in that region. Often several farm models are combined into a larger modelling framework so as to reflect a greater variety of farm types and regions. Some sets of farm models are combined with coefficients that allow projection of results to the level of national aggregates, so that they represent the whole farming industry in a country.

Market models typically come at the level of countries, if not the whole European Union, and reflect the aggregate supply of and demand for agricultural products (and sometimes inputs). They do this for either individual products or sets of products. In the latter case, market models typically reflect the cross-linkages among different products (e.g. when wheat price rises, wheat production expands and barley production contracts, and vice versa for consumption). Most market models operate at the farm gate or wholesale level, though some of them also include processing and marketing activities and extend to the level of consumer goods. Some market models include price-responsive world markets, in some cases with many countries being represented in the model (in which case one also speaks of trade models). Though market models typically represent the aggregate country level, they do not include the effects on sectors outside agriculture, nor the macro-economic feedback effects through factor prices and income formation. This is why they are called "partial equilibrium" models.

On the other hand, models of the overall economy, to the extent they are used in agricultural policy analysis, typically are "general equilibrium" models capturing the linkages that operate at the macro-economic level. For example, a rise in agricultural price support may result in higher budget expenditure, requiring larger tax revenues which reduce the spending power of consumers and hence their demand for food, in addition to the direct demand-depressing effect of rising food prices as a result of higher price support. General equilibrium models always consist of several economic sectors, one of which may be agriculture. However, there are also general equilibrium models that distinguish between several sub-sectors within agriculture, typically along the lines of major product categories. General equilibrium models always include factor markets (labour, capital, sometimes land) and can therefore reflect the effects of policies on factor prices. In some cases, general equilibrium models of several countries are combined into international trade models.

All these different categories of models can come in widely varying degrees of detail and complexity. Moreover, they can either be "static", representing the situation at a given point in time, or "dynamic", reflecting inter-temporal linkages between economic variables (such as investments in one year expanding production capacity in the following year). Not to be confused with truly dynamic models, sequential runs of a given static model for several years (possibly with changing parameters to reflect developments such as technical progress) can be used to generate time series of results.

Which particular type of model (if any) should be used in a given evaluation exercise depends on many factors, including the nature of the issue to be analysed, the time available, the resources that can be invested, the availability of a ready-made model of appropriate structure and, last but not least, the ease of communication with the model builder. In most cases of day-to-day policy evaluation it will not be possible, and by far too costly, to construct a new and comprehensive model just for that purpose. Policy evaluators will, rather, have to rely on existing models which may, or may not, have to be adapted to the particular issue to be analysed. Where the appropriate model is not available in-house, external expertise from model "owners" may have to be acquired. As far as the nature of issues to be analysed in evaluation is concerned, it is pretty self-evident which category of model makes sense for which type of issue. For example, policy effects on farm incomes are best studied in farm models, as are issues such as regional production structures and land use. Implications for market balance and foreign trade, as well as many budget and welfare effects of market policies are appropriately analysed in market models. General equilibrium models can provide insights regarding effects of agricultural policies on things such as factor markets, overall economic welfare and macro-economic variables like exchange rates. For some more comprehensive evaluation exercises, several different models may be used simultaneously, and in some cases they can even be formally linked to form a "super model".

There is a long history of the use of quantitative models in the evaluation of the CAP – by far too long and too complex to be reported here. There are also so many models around that can be and have been used in CAP evaluation that an attempt to survey them here in anything like a comprehensive fashion would not fit the scope of this study. Instead, a recent example can well illustrate the combined use of different categories of models in a particularly important *ex ante* evaluation of the CAP. In October 1998, DG VI of the European Commission issued a document that reports on quantitative analyses of the potential effects of the Agenda 2000 proposals for further reform of the CAP.⁸³ Four different major models were used, some (now) available in-house in the Commission, and others provided by academic researchers. In addition, effects on consumer welfare were assessed on the basis tools available in DG VI. In the context of this study, the results on Agenda 2000 are less interesting than the nature of the quantitative tools used and the way in which they were employed for *ex ante* policy evaluation.

The four models used can be briefly described as follows. The SPEL/EU-MFSS model, developed in the University of Bonn, is the Medium-term Forecast and Simulation System of the Sectoral Production and Income Model for Agriculture, frequently used by the Commission for *ex ante* analysis of changes in agricultural policy. It is based on an activity-based accounting approach complying with the definitions used in the Economic Accounts for Agriculture (EAA). Supply and

⁸³ European Commission (1998b)

demand of agricultural products are modelled to respond to economic incentives. Modelling is done at the level of individual Member States, and then aggregated to the EU level. An external trade component allows for net trade flows between the EU and the rest of the world and determines equilibrium world market prices. RAUMIS is a supply model for the agricultural sector in Germany based on linear programming. This model, also developed at the University of Bonn, divides Germany into 431 regions, each having a particular production model. It analyses the impact of policies on the structure of agricultural production, on farm income and land prices at the regional level in Germany. It is a comparative-static model for a given year, based on 77 crop activities and 16 livestock activities. Production intensity adjusts endogenously to the price situation. Some environmental impacts of agricultural production can also be assessed. These two models were used to evaluate the effects that Agenda 2000 reforms of the CAP might have at the sectoral and regional level in agriculture, in terms of variables such as supply, demand, trade, farm income and environmental impacts.

The CAP-Modelling and Accounting Tool (CAPMAT), developed by three Dutch institutes, consists of a dedicated database, an applied general equilibrium model to simulate overall medium term effects and a simulation and accounting tool that uses the outcomes of the previous parts to perform scenario calculations. The basic analytical engine is the ECAM model, an applied general equilibrium model which was already previously used for evaluating the CAP (Folmer *et al.*, 1994). ECAM works at the level of individual Member States, of which the nine countries that were members before Greece joined are represented in the model. The model structure is typical for applied general equilibrium models, with fully responsive elements for supply and demand in several sectors of the economy, including agriculture, and a complete macro-economic framework. The results of ECAM are used as data in the Simulation and Accounting Tool (SAT), an independent model that processes the comparative-static results of ECAM into dynamic simulation trends over time, for the individual policy scenarios analysed. SAT also introduces the Member States not included in ECAM, by adding a 'sister' country with results derived from those of ECAM. CAPMAT was used to evaluate expected effects on variables such as supply and demand for agricultural products, EU trade, farm income and consumer burden, and budget expenditure.

QUEST, finally, used in version II for this study, is a model developed by DG II of the European Commission, to analyse the economies of the EU Member States and their interaction with the rest of the world. It is a macro-economic growth model without sectoral breakdown that estimates the effects of changes in policies on the growth rate of the economy in the short and the long run. The model contains modules for the EU Member States, the US, Japan and for another ten countries/regions. The real rate of interest and the real rate of exchange are determined by private savings behaviour and demand and supply in international trade. The model allows the analysis of the effects of a change in prices of agricultural products on wage settings and investment costs of firms and checks if there is any lasting effect on the long-run growth rate of the Member States. It was used in this exercise to evaluate Agenda 2000 effects on the longer run development of macro-economic variables such as GDP, investment, wage costs and employment.

In addition to these major models, a number of partial equilibrium models of EU agricultural markets, internally developed in DG VI, were used to evaluate the impact of the proposed policy changes on consumer welfare in the EU.

In this brief reference to the recent model-based evaluations of the Agenda 2000 proposals for the future of the CAP, no comments can be offered on the substance and quality of the results achieved. The point to be made here is only that a multitude of different quantitative models are available, both in the Commission and in academic institutions, that can well be brought to bear on evaluations of the CAP. It is reassuring to note that the European Commission makes use of these opportunities in analysing, on an *ex ante* basis, important policy proposals such as Agenda 2000.

As a final point on quantitative methods for agricultural policy evaluation, an approach now pursued in OECD should be mentioned. Since the mid-1980s, the OECD has used the concept of producer subsidy equivalent (PSE) to measure the value of monetary transfers to agriculture resulting from agricultural policies. Equally, the consumer subsidy equivalent (CSE) measures the monetary value of the burden agricultural policies impose on consumers. The results of this analysis, done on an annual basis, have shown that in recent years the overall level of agricultural support in most OECD countries has slightly declined, but in particular that the composition of support has generally changed from market price support towards direct payments.⁸⁴ In an attempt to identify the implications of this development for consumer and taxpayer costs of support, trade distortions and the economic well-being of farmers, the OECD has recently embarked on an analytical project that is hoped to provide new insights by using the PSE concept in conjunction with quantitative modelling of policy effects. The analytical framework used is called a Policy Evaluation Matrix (PEM). This matrix, to be established for individual countries (and the EU) consists of columns denoting different agricultural support measures, such as price support, direct payments and input subsidies. The rows of the matrix represent indicators of the effects of these measures on variables of interest to policy makers. For the time being, the indicators analysed are taxpayer costs, domestic consumer costs, farm household income, input supplier profits, production volume and value, consumption volume and value, and net trade volume and value. All these effects are calculated for individual agricultural products at the aggregate level of the country concerned (for the EU aggregate in the case of the Union). The effects of the various policy measures on these indicators are estimated using a quantitative model of the agricultural sector, filled with country-specific parameters.⁸⁵ For the time being, results are reported in the form of impacts on the various indicators of one monetary unit (say, one ECU) extra support provided through alternative policy measures (price support, direct payments, input subsidies). Work on this analytical framework is still in progress in OECD, and results have not yet been published. DG VI of the European Commission is involved in this project, and works at applying this type of analysis to the CAP.

Fundamentally, PEM analysis as pursued by the OECD is not a completely new type of policy evaluation, and not a new alternative to quantitative policy evaluation as outlined above. It continues the tradition of using quantitative models for evaluating certain effects of given agricultural policy measures. In terms of the different categories of policy evaluation discussed above, PEM analysis would appear to fall in the class of *ex ante* evaluation, and may therefore help to take decisions on future policies, rather than evaluating concrete measures adopted in the past. What makes this work particularly interesting and worthwhile mentioning in the context of this study is the systematic way

⁸⁴ See OECD (1997b).

⁸⁵ In terms of the model categories described above this model is a market model.

in which a number of indicators of policy effects, very much in line with some of the indicators discussed above in Chapter 6 is established. Moreover, as the same analytical framework will be applied by OECD to various countries (very much like the PSE analysis is applied to all OECD countries), the PEM analysis will yield interesting insights of a comparative nature at the international level. If OECD should get to the point where PEMs are established at regular intervals like already is the case for PSEs, this type of analysis will add another useful element to the process of evaluating measures under the CAP on a periodic basis.

7.4. Treatment of Short Term and External Factors

In *ex post* policy evaluation, actual data on the indicators of interest are compared with the reference situation (say, actual 1996 farm income is compared to farm income that would have resulted from the reference policy). As actual data also reflect the impact of short term and external factors beyond the control of policy makers, this comparison implies the danger that systematic influences of policy measures cannot be distinguished from the random effects of such short term and external factors. For example, actual farm income in a given year may have been depressed because of adverse weather conditions and resulting unusually low yields in that year, or because prices for non-agricultural inputs were particularly high, say as a result of price hikes on international markets for crude oil. Of course these low incomes can then not be taken to reflect only the effects of agricultural policies. Similarly, FEOGA expenditure depends, among others, on the development of world market prices as affected for example by the Russian crisis, and not only on the way the CAP is pursued. This is a somewhat technical problem, but in the practice of policy evaluation it can well cause serious headaches.

There are two seemingly different ways to account for this problem, which though fundamentally involve the same analytical requirements. First, one can "clean" actual data by removing the "dirt" caused by short term and external factors. Second, one can augment the reference situation used in policy evaluation by the same "dirt" as contained in actual data, so as to make it comparable to actual reality. Of course, there is not really a difference between these two approaches, because in both cases we need to know the type and amount of "dirt" caused by short term and external factors, and the real problem is to determine that "dirt".

In quantitative policy evaluation based on modelling approaches, ideally this problem should not occur because the ideal model would contain all exogenous factors potentially affecting the indicators we are interested in. If we plug the policies actually pursued, and to be evaluated, into this model, the results should therefore reconstruct reality precisely, including the effects of all short term and external factors. We could then use that same constellation of all exogenous factors in our model, and change only the policies concerned so as to create the reference situation. This scenario would then also contain all "dirt", and should therefore be completely comparable with reality, without the "dirt" distorting our policy evaluation. However, models are never so complete and can therefore not solve the problem of short term and external factors in this way. To be sure, any other approach not based on quantitative models will not do better in this regard, because it offers even less of the possibility to account for short term and external factors in any systematic way.

There are two other ways of dealing with this problem, and both of them are used in the practice of policy evaluation. First, rather than comparing the model-based reference situation with reality, it can be compared with a quasi-reality also resulting from the same model, where the policy measures that

were implemented are plugged into the model instead of the reference policy. *Ex post* evaluation pursued in this way is then methodologically very similar to *ex ante* evaluation which also compares two synthetically constructed situations, because future reality cannot really be predicted in *ex ante* evaluation of policies. Second, actual data are filtered in some way so as to remove the "dirt" resulting from short term and external factors. This filtering can be done in various ways, and the appropriate approach to be adopted depends on the case concerned. If (some of) the disturbing short term or external factors are known and their influence can be determined in some way, then the "dirt" can be removed in a very direct way. For example, if it is known that in the year concerned cereal yields were six percent below their usual levels, then actually observed farm incomes can be corrected, through accounting procedures, for this factor. In many cases, though, this approach may not be feasible, not the least because too many short term and external factors were at play. In such cases, the pragmatic approach of averaging is usually adopted. Rather than using actual 1996 farm income, for example, the three-year average of 1995-97 is taken to represent "normal" farm income in 1996. In a geographical dimension, rather than taking farm incomes of individual villages, average farm incomes in larger regions are used. Of course this approach makes sense only if the policy measures to be evaluated did not differ between the observations (years, villages) used in averaging.

Though the problem caused by short term and external factors is, in a way, a purely technical issue in policy evaluation, much care should be applied in dealing with it. The reason is that the credibility of policy evaluation can otherwise be easily undermined. Policy makers unhappy with the result of a given policy evaluation will happily point out that the observation on which the evaluation was based was atypical because such and such short term or external factors were at work. It is therefore fundamental that this issue is thoroughly considered in the practice of policy evaluation.

8. Recommendations: Towards Implementing a Comprehensive System of CAP Evaluation

The preceding Chapters have explained why it is important to evaluate the CAP on a regular basis, how evaluation is currently practised at the Union level and in a number of selected Member States, how the objectives of the CAP may be interpreted and turned into empirically observable indicators that can be used in practical evaluation exercises, and some methodological requirements which should be considered when evaluating the CAP. Considering these observations, what can be said constructively about the central question posed in this study, i.e. the purpose and methodology of CAP evaluation? Are there any operational recommendations that can be offered, any improvements of the state of affairs that can be suggested? These are questions to which the present Chapter will respond.

As this study is not designed to generate a manual for CAP evaluation, the recommendations are of a conceptual nature, directed at some of the more fundamental issues rather than the technical implementation of evaluation practice. Moreover, there is no doubt that a number of points to be discussed here are political in nature, and potentially sensitive. Such issues transcend the realm of academic analysis, and decisions have to be taken at a political level. The proposals that will be made here should, therefore, be interpreted as suggested starting points for a discussion on these issues, rather than a blueprint for a holistic regime.

The Different Purposes of CAP Evaluation

Fundamentally, the two major purposes of evaluation are justifying the necessity of policy in the first place, and then managing justified policy efficiently. As these are different purposes of policy evaluation, they require different approaches. As far as practical evaluation of the CAP is concerned, the distinction between different purposes can even go a step beyond these two fundamental categories of evaluation. It is suggested that four different purposes of evaluation are distinguished by their focus and purpose, i.e. strategic evaluation of the whole policy, evaluation of changes in the CAP, evaluation of specific programmes, and evaluation of the operational implementation of policy instruments.

- A) At the highest, strategic level, evaluation should examine the objectives and instruments of the policy as a whole. Its purpose is: to examine the coherence of the objectives one to another and with respect to other objectives of the EU and its Member States; to test if the CAP objectives remain relevant; to judge if they are achievable; to assess if they are being achieved; and to judge whether the costs of achieving them is proportionate to the benefits gained. This sort of evaluation must cover the whole EU and all the dimension of the CAP (market policy, rural development and agri-environment).
- B) At a second level, significant changes in the pre-existing policy regime should be evaluated. In this type of evaluation, the purpose of evaluation is not so much to assess whether the policy is justified in the first place, but how it should be adjusted to the changing economic, social and political environment, and considering the effects the policy had in the past. In the case of the CAP, evaluation of such policy changes involves the business of correcting for the less desirable effects of previous policies and thus introducing the second order objectives discussed above in Chapter 5. Evaluation of policy change should assist the choice of optimal instruments, and discover the extent to which any changes made in the past met their objectives.
- C) At a third level, specific ongoing programmes should be evaluated on a regular basis, independently of whether significant changes are being considered or not. The purpose of this type of evaluation is to learn whether objectives have been met, to inform judgements about the continued necessity of the programmes or adjustments to them in the light of changing circumstances and past achievement of objectives.
- D) Finally, at a more technical level the operational implementation of individual policy measures should be evaluated. The purpose of this operational evaluation is to ensure that policy administration is cost-effective and safe from corruption.

As the purposes of evaluation differ between these four types of cases, the approaches to be adopted may have to differ as well, regarding the timing of evaluation exercises, the perspective in terms of *ex post* and *ex ante* evaluation, and the involvement of internal and external expertise.

The Timing and Perspective of CAP Evaluation

Evaluation of the CAP is an ongoing business, and it should be considered as a continuous process. However, for reasons of both cost-effectiveness of evaluation and continuity of policies, evaluation cannot occur too frequently. Hence there is reason to consider something like an "optimal" frequency of evaluation. Given the varying purposes of the different types of evaluations suggested above, their frequency may have to differ as well.

For obvious reasons, strategic evaluations of type A can and should be done only once in a while, say every five years. The perspective to be adopted in these evaluations will have to be mainly *ex post*, with an eye on whether the general policy orientation was justified in the first place. To an extent, though, some *ex ante* evaluation of fundamentally different policy regimes for the CAP may also be necessary in this context, in order to provide some insight into alternative strategies that could potentially replace the CAP as it was pursued in the past. In such evaluations, the question should, for example, be asked why it is considered necessary to provide market price support at all, rather than leaving market forces to govern supply of and demand for agricultural products. One fundamentally different policy alternative that could be assessed in such strategic evaluations could, for example, be a regime of completely decoupled direct income payments, guaranteed for a limited period of time.

Evaluations of significant policy changes under the CAP (type B) should be done, from an *ex ante* perspective, whenever such changes are being considered, and definitely before they are proposed by the Commission. After such changes have been made, they should also be evaluated *ex post*. Such evaluation should occur once, say three years after the change was implemented. As a concrete example, the MacSharry reform would, according to this proposal, have been evaluated, in a formal sense, *ex ante* before the proposal for this reform was tabled. The *ex post* evaluation of this policy change, which was fully implemented in 1995, would then be provided in 1998.

Evaluation of specific programmes (type C) is an ongoing business that should be done within a systematic cycle so that each major programme is reviewed, say, every three years. As the purpose of this evaluation is to see whether the policy measures concerned have operated satisfactorily, the perspective adopted should be *ex post* (or, using the terminology of the evaluation guide issued under the SEM 2000 initiative⁸⁶, the perspective is "intermediate" as the policy is still in place). As an example, intervention buying and export restitutions under the CAP would be evaluated every three years.

Operational evaluations of programme administration (type D) also are an ongoing business, to be conducted on a regular basis. A three-year cycle may be appropriate for these evaluations as well, and the perspective also has to be *ex post* (or "intermediate").

⁸⁶ European Commission, 1997a.

External versus Internal Evaluation of the CAP

A major issue in all policy evaluations is whether they should be done in-house in the institution responsible for administering the measures concerned, or by external experts. Both solutions have advantages and drawbacks. Internal evaluators have optimal access to all technical information required, they can closely co-ordinate with the operational staff responsible for administering the measures, and can involve them in the evaluations, thus generating learning-effects that may be very valuable for future policy implementation. External experts, on the other hand, are more independent from the political considerations that have resulted in the establishment of the policies concerned, and may feel less constrained by considerations regarding political "feasibility" of alternative regimes. They may also command specific methodological expertise, for example in the form of analytical tools (such as quantitative models) that are time-consuming to develop. Given such comparative advantages and drawbacks of internal and external evaluators, it is probably best to combine their contributions, though the "optimal" combination may again have to vary between the different types of evaluations.

The occasional strategic evaluations (type A), which are supposed to address the fundamental justification of policies, are probably best done by external experts, who may feel sufficiently free to question the validity of the overall policy regime. As far as the CAP is concerned, such external evaluations have always been provided, from time to time, by external experts, mainly from academia, on their own responsibility. However, the provision of such external evaluations has generally been unsystematic and lacked regularity and co-ordination. It is therefore suggested that the Commission should have funds at its disposal for commissioning, on a regular basis, such strategic evaluations of the CAP. The Commission's role in this business should be to co-ordinate such external evaluations, to make sure that all important questions get looked at, and that best evaluation practice is spread.

Evaluations of policy changes (type B) and of specific ongoing programmes (type C), on the other hand, are probably best done in-house in the Commission. Type B evaluations, in particular the *ex ante* evaluations of policy changes to be proposed, have to be closely integrated into the process of generating policy proposals in the Commission, and given the political sensitivity of policy changes to be proposed, the involvement of external evaluators would be rather difficult. Type C evaluations of ongoing programmes require a lot of technical information not easily available to external experts. The new evaluation unit established in DG VI has an important role to play in such evaluations. However, as discussed above in Chapter 3, it is also desirable that operational staff in the Commission also gets involved to some extent in these evaluations. While internal evaluation is probably the best approach in these cases, it is suggested that from time to time the in-house evaluations provided in the past are assessed by outside experts, to make sure that the approaches adopted and results achieved are not too much constrained by considerations regarding political "feasibility" of alternative policies. The Commission should, then, have funds available for commissioning such regular outside assessments of its own internal evaluations.

Operational evaluation of the administration of policy measures to make sure they are cost-effective and safe from corruption (type D) is again best done by outside evaluators. This is the domain of the European Court of Auditors.

It might also be considered to involve, to some extent and from time to time, non-governmental organisations in CAP evaluation. This could be particularly useful where issues such as implications

for the environment and regional development are concerned. One way of doing this could be to invite such organisations to comment on the results of evaluations done in the Commission or by external experts.

The Role of the Member States in CAP Evaluation

It is natural that Union policies, in particular those pursued through regulations, should be evaluated at the Union level. Thus evaluation of market and price policies under the CAP should be organised, and where appropriate done, by the Commission. The results of these evaluations should, though, be available to the Member States, and the Member States should be provided an opportunity to comment on the results of these evaluations, on a regular basis and in a structured process.

At the same time the Member States of course have the right, and should use the opportunity, to engage in their own evaluation of the CAP, including those agricultural policy measures that are pursued at the Union level. After all, the Member States are paying for these policies, in economic and budgetary terms, and they (should) have an interest in evaluating the extent to which these policies meet their objectives. Indeed, it is desirable that an atmosphere of "competitive evaluation" is generated, through which the CAP is evaluated at different levels and from different perspectives. This would help both to provide a more comprehensive picture of how the CAP operates, and to strengthen the methodological basis of policy evaluation in the EU. Of course, different Member States would probably come up with differing evaluation results, depending on their specific objectives and circumstances. However, if organised appropriately the resulting process of debate about the results of evaluation pursued at both the Union and the Member State levels can create a wealth of information on the implications of the CAP and on policy alternatives. From everything that has been said above in this study it should be clear that there is not one simple "truth" on policy effects that could possibly result from evaluation. However, if the results of alternative evaluations are confronted with each other, the overall process of multi-dimensional policy evaluation may iterate towards the "truth".

Moreover, if in the future the CAP progresses in the direction of becoming an integrated rural policy, rather than an agricultural policy based on commodity market intervention, and if this means that a significant part of the budget is devoted to agri-environment and rural development, then this has important implications for evaluation. Such policy will inevitably be defined and implemented - and maybe even funded - more at the Member State and regional level. Therefore evaluation will have to be done more at that level too.

The Role of the European Parliament in CAP Evaluation

In addition to its budgetary competence and responsibility, the European Parliament should also have a role in the process of more general evaluation of the CAP as outlined in this study. Like the Member States, the European Parliament should have access to the results of all CAP evaluations of the different categories discussed above. Moreover, the European Parliament should have the right to request specific evaluations to be done either by the Commission or by outside experts. For this purpose, it is suggested that a given share of the total budget made available for CAP evaluation (say, ten per cent of that sum) is earmarked for evaluations specifically requested by the European Parliament.

Creating a Better Basis for CAP Evaluation

For policy evaluation to be effective, and its results to be useful, the foundations for evaluation have to be laid already when policy decisions are prepared and taken. In particular, the objectives pursued should be defined clearly, and determination of the indicators to be used in future evaluation of the policy should not be simply left to the evaluator. The discussion of objectives and indicators in Chapters 5 and 6 above has shown that in this regard much remains to be done. At the same time, the results of existing policy evaluations should play a more prominent and effective role in decisions on future policies.

It is therefore suggested that in its proposals for regulations and directives, the Commission should in future specify objectives more clearly than is currently the case in the whereas clauses. Moreover, just as the expected financial implications have already to be indicated in a financial statement accompanying Commission proposals, all policy proposals should in future be accompanied by a list of indicators to be used in later evaluation. In its debates about policy proposals, the European Parliament and the Council should explicitly address the whereas clauses and lists of indicators, and by deciding on a policy also decide on them. If this is done, not only will it create a much better basis for future policy evaluation, but also it will enhance transparency and accountability of policy decisions.

At the same time, Commission proposals for policy changes of type B should be accompanied by the results of prior *ex ante* evaluation, and Commission proposals for changes of policies falling in category C should also explicitly make reference to results of *ex post* evaluations of the policies concerned. A procedure like this would create a desirable feed-back between evaluation and policy decision.

Methodological Issues in CAP Evaluation

Policy evaluation is, to some extent, an art that escapes fixed rules. However, there are certain fundamental methodological issues that can and should be considered, as discussed above in Chapter 7. Most of these issues are of a more technical nature, and they do not have to be mentioned here again. However, there is one issue that is often overlooked in the practice of policy evaluation, and where improvement is desirable.

In policy evaluations, both *ex post* and *ex ante*, there should always be an explicit description of what the reference situation is meant to be, i.e. which policy is the benchmark (no policy, unchanged policy, or some specified alternative policy), and the outcome which would have been, or is expected to prevail, under that benchmark policy. Logically, the effects of policy can only be established through a comparison with a well defined reference situation. In *ex post* evaluations, this reference case necessarily is a counterfactual situation. Looking at changes of variables over time is simply not sufficient, because it is likely to create a wrong impression of the magnitude and possibly even the direction of the policy effects.

The reference situation does not always have to be established through formal model analysis, it can also be based on other approaches such as trend extrapolation. If no other approach is feasible, even pure guesswork can be used to establish the reference situation. However, it is important that evaluators make it explicit what they think the situation would have been, or might be in the future, in the absence of the policy to be evaluated. This is crucial where evaluation deals with the

quantitative effects of policies. At the same time, analysts should say clearly against which reference policy they are evaluating the current policy. If in any way possible, not only the old policy should be used as the reference policy, but evaluation should also be made in comparison with the conceivably most cost-effective policy. This approach should, in particular, be followed in evaluations of type A.

In *ex post* evaluations, evaluators should use the objectives and indicators defined by the policy makers (see previous section), so as to assess whether the policy has achieved the objectives it aimed to pursue. However, they should also be free to base their evaluation on additional objectives, and indicators derived from them, in order to provide a broader perspective.

Resources to be Made Available for CAP Evaluation

Analysis of the CAP, of the economic, social and political circumstances in which it operates, of its regional and environmental implications, of market developments both in the EU and at the international level, of agricultural issues involved in EU enlargement, of the way in which the CAP interacts with other countries' agricultural policies, of its relationship with the international trading order, and of many related issues, has always played an important role in the Commission. In particular, analytical work done in DG VI A.I has proved an invaluable source of information on which the Commission could base both the implementation of existing and proposals for future policies. There is no doubt that such analytical efforts should play an important role in agricultural policy making. Agriculture, taken together with the upstream and downstream activities directly linked with it, is an important agricultural sector, with major significance for the many and large rural areas in Europe. A public policy that intervenes so strongly in such an important sector, and which consumes such a large amount of resources, has to be based on extensive and sound analysis. It would simply be irresponsible vis-à-vis the general public to pursue this policy in an *ad hoc* manner. This is true not only in the EU, but also in other countries. Indeed, a number of countries in other parts of the world maintain a much larger analytical capacity directly linked to the institutions responsible for agricultural policy making and implementation. For example, the Economic Research Service of the United States Department of Agriculture is a body much larger than the analytical capacity available to DG VI of the European Commission. Of course, a direct comparison is difficult for many reasons, not the least because of the joint responsibility of the Union and the Member States for EU agricultural policies, and because of many other institutional differences between the United States and the EU. However, it is significant that on a number of occasions, not the least in the analysis of international markets for agricultural and food products, the EU Commission has to rely on work done in ERS and other institutions outside the EU, such as the World Bank, OECD and FAO. A political and economic power like the EU should be able to rely on its own analysis, and make its commensurate contribution to the global body of analytical information on agricultural issues.

Based on these considerations it is recommended that the EU invests considerably more resources than in the past into the analysis of agricultural policies and related issues. Not all of these resources would have to be made available for formal policy evaluation in the sense discussed in this study, though more comprehensive analytical groundwork would certainly make an important contribution to creating a better basis for the different categories of formal CAP evaluation outlined above.

At the same time, an explicit decision should be taken on the magnitude of resources to be made

available for CAP evaluation. There are no objective criteria on which to base this decision. However, an argument could be made that the budgetary resources to be invested in CAP evaluation should be linked to expenditure on the policy itself. As a starting point for discussion, a magnitude of one half of one tenth of one per cent of EU expenditure on the CAP is suggested as the budget to be made available for CAP evaluation. For the time being this would mean an annual budget for CAP evaluation in the order of magnitude of Euro 20 million. It is suggested that this amount is budgeted for CAP evaluation at the Union level. Evaluation pursued by the Member States would come on top of this amount of resources, and be financed directly by the Member States.

It should be repeated, that the suggestions for the future of CAP evaluation made in this Chapter are offered to provide a starting point for discussion at the political level. They are based on the assessment presented in the study and thus have some analytical base. However, they also, of course, reflect the subjective views of the authors. These suggestions should also be seen in the context of some of the comments made, in Chapter 3 above, on the current practice of CAP evaluation at the level of the European Union. In any case, much would be gained if a constructive discussion on these and related issues were held in the institutions concerned.

9. Conclusions

This study has discussed the purpose and methodology of evaluating the CAP. At a fundamental level, the purpose of evaluating EU agricultural policy, like any other public policy, is reasonably straightforward. It is justifying the necessity of the policy in the first place, and then managing justified policy efficiently. Evaluation should force the policy process to be more open, objective and self-justifying. While this is the general aim of all policy evaluation, in the practice of evaluating the CAP it makes sense to distinguish different types of evaluation activities, with different somewhat more specific purposes. In particular, the purpose of what can be called strategic evaluation includes an examination of the objectives of the CAP and of their coherence one to another and with respect to other objectives of the EU and its Member States, as well as an assessment of the question whether these objectives are achievable at all. At a somewhat lower level, regular evaluation of policy changes and of existing programmes will have to take the objectives as largely given and concentrate on the extent to which the measures to be evaluated achieve them in a cost-effective manner. Finally, audits of the operational implementation of the CAP have the purpose of assessing financial management, legality and regularity.

This study has not aimed at an actual evaluation of the CAP. However, its mandate to discuss the issues to be considered in evaluating the CAP has required it to take a look at the way the objectives and costs of the CAP can be interpreted in an evaluation of the policy. The conclusions to be drawn from that part of the study are rather sobering. The CAP's objectives lacked clarity and coherence from the beginning, are interpreted in different ways in the Member States, kept changing over time, and were gradually complemented by second-order objectives formulated with the goal of reducing problems caused in pursuit of the original objectives. For CAP evaluation, this creates a difficult problem. If the objectives are unclear, incoherent or conflicting and costs to be considered are not well defined, then no amount of evaluation is going to be very helpful in guiding us to a conclusion on whether the CAP has achieved its objectives and whether it has done so in a cost-effective way. One conclusion to be drawn regarding the purpose of CAP evaluation is that strategic evaluation, including an examination of the objectives, should be given high priority. Another conclusion is that

in future decisions on the CAP a much more determined effort should be made to specify the objectives pursued, as well as the indicators translating these objectives and the costs to be considered into empirically observable variables that can be used in policy evaluation. The study has recommended that this be made a formal requirement in the process of taking political decisions on the CAP.

As another conclusion regarding objectives and costs, this study has strongly argued that the costs of the CAP should be defined from a broad perspective, and should include not only public expenditure. Consumer burden, distortions of resource use, macro-economic implications, international repercussions, environmental effects and other items should all be considered along with budget expenditure. Indeed, in order to reduce the overall cost of the CAP it may be necessary, under some circumstances, to invest more budget expenditure in a reform of the CAP. This is not to say that budget expenditure under given policies should not be scrutinised very carefully. However, evaluation of the CAP would take too narrow a view if it were to assess cost-effectiveness of the policy only on the basis of budget expenditure.

As far as methods and approaches used to evaluate the CAP are concerned, this study has, among others, reviewed the current practice of CAP evaluation at both the Union level and in four selected Member States, on a limited scale. Quite apart from the substance of the findings in each particular case, one major conclusion to be drawn from that part of the study is that the institutional arrangements established, the procedural approaches followed, the amount of resources invested, the intensity of analysis provided, the links created with future policy decisions, and many other dimensions differ strikingly among the Member States included in this brief review, and also between them and the Union. In some cases, in particular in the European Commission and in the Netherlands, new evaluation procedures have been established recently and have not yet had time to bear fruit. Moreover, only a small number of Member States could be considered in this study. Given the limited amount of information that could be collected within the scope of this study, the authors have deliberately steered clear of judging on best practice of CAP evaluation as it is currently done in the Member States and the Union. However, given the importance of the subject it may be a good idea to engage in a study that systematically brings together such information for all Member States, looks into the results of evaluations done at both the Member State and Union levels, and draws conclusions regarding best practice that can help to improve future systems of CAP evaluation.

On the analytical methodology of evaluation, this study has tried to make a strong point for logical consistency, and in particular for the necessity of establishing the counterfactual reference situation against which current policy is evaluated. Inappropriate specification of that reference situation, in particular use of a simple before-after comparison is the single most frequently committed mistake in policy evaluation, and it should be avoided in future evaluations of the CAP. Establishment of the counterfactual reference situation does not necessarily require the use of sophisticated econometric models, it can also be done through pure guesswork. However, it should always be an explicit element of policy evaluation. At the same time the study has pointed out that a wide variety of analytical tools for quantitative analysis is available that can, and should in future more intensively, be used in evaluating the CAP.

This study has, finally, made a number of recommendations regarding the institutional arrangements and the amount of resources to be made available for CAP evaluation. It has argued for a more

prominent role of the Member States in CAP evaluation; for more involvement of the European Parliament in CAP evaluation; for a well-defined role of external expertise in CAP evaluation; for considerably more resources than in the past to be invested into the analysis of agricultural policies and related issues; and for a budget for CAP evaluation at the Union level in the order of magnitude of one half of one tenth of one per cent of EU expenditure on the CAP (around Euro 20 million).

Overall, the picture emerging is that a lot of work has always been done in the broad area of analysing and evaluating the CAP. The accelerating speed of CAP reform to an extent mirrors the results of these efforts. However, compared to what is done in other parts of the world in terms of laying the analytical groundwork for agricultural policies, the EU should intensify its work in this area. Moreover, the formal process of CAP evaluation can be improved in a number of regards. A public policy that consumes budgetary and other resources at a level found in the CAP should be well justified and thoroughly scrutinised. In that context, evaluation of the CAP can make a significant contribution.

REFERENCES

- Bouckaert G. (1996), Measurement of Public Sector performance: Some European Perspectives, in: Halachmi A. and Bouckaert G., *Organisational Performance and Measurement in the Public Sector, Toward Service, Effort and Accomplishment Reporting*, London: Quorum Books.
- Bovaird T., Gregory D. (1996), Performance Indicators: The British Experience, in: Halachmi A. and Bouckaert G., *Organisational Performance and Measurement in the Public Sector, Toward Service, Effort and Accomplishment Reporting*, London: Quorum Books.
- Brouwer, F.M., and S. van Berkum (1996), *CAP and Environment in the European Union: Analysis of the Effects of CAP on the Environment and an Assessment of Existing Environmental Conditions in Policy*. Wageningen: Wageningen Press.
- Buckwell *et al.* (1997), *Towards a Common Agricultural Policy for Europe*, European Economy reports and studies No 5, Commission of the European Communities DGII, Brussels.
- Bundesministerium der Finanzen (1995), *Vorläufige Verwaltungsvorschriften zur Bundshaushaltsordnung*. Bonn.
- Carter N. (1991), Learning to measure performance: The use of indicators in organisations. *Public Administration*, Vol. 69, Spring 1991, pp. 85-101.
- Commission Européenne (1998), *Bilan d'Évaluation 1996/1997. Avant-projet de budget général des Communautés Européennes pour l'exercice 1999*, Section III-Commission. Document de travail, Mai 1998. Brussels: European Commission.
- Commission of the European Communities (1988), *The future of Rural Society*, Com(88)371 final, Brussels.
- Commission of the European Communities (1998), *Proposals for Council Regulations (EC) concerning the reform of the common agricultural policy*, COM(1998) 158 final, Brussels.
- Committee for Reforming the Common Agricultural Policy (1997), *Food and the Environment: Swedish Strategy for the future of EU agriculture*. SOU 1997:151. Stockholm: Swedish Government Official reports, Ministry of Agriculture.
- Council Regulation (EEC) No 1079/77 of 17 May 1977 on a co-responsibility levy and on measures for expanding the markets in milk and milk products, *Official Journal* L 131, 26/05/1977, p. 0006.
- Council Regulation (EEC) No 1822/77 of 5 August 1977 laying down detailed rules for the collection of the co-responsibility levy introduced in respect of milk and milk products, *Official Journal* L 203, 09/08/1977, pp. 0001-0004.
- Council Regulation (EEC) No 2040/86 of 30 June 1986 laying down detailed rules for the application of the co-responsibility levy in the cereals sector, *Official Journal* L 173, 01/07/1986, p. 0065.
- Council Regulation (EEC) No 2052/88 of 24 June 1988 on the tasks of the Structural Funds and their effectiveness and on coordination of their activities between themselves and with the operations of the European Investment Bank and the other existing financial instruments, *Official Journal* L 185, 15/07/1988, pp. 0009-0020.

- Council Regulation (EEC) No 2078/92 of 30 June 1992 on agricultural production methods compatible with the requirements of the protection of the environment and the maintenance of the countryside, *Official Journal* L 215, 30/07/1992, pp. 0085–0090.
- Council Regulation (EEC) No 2080/93 of 20 July 1993 laying down provisions for implementing Regulation (EEC) No 2052/88 as regards the financial instrument of fisheries guidance, *Official Journal* L 193, 31/07/1993, pp. 0001–0004.
- Council Regulation (EEC) No 2081/93 of 20 July 1993 amending Regulation (EEC) No 2052/88 on the tasks of the Structural Funds and their effectiveness and on coordination of their activities between themselves and with the operations of the European Investment Bank and the other existing financial instruments, *Official Journal* L 193, 31/07/1993, pp. 0005–0019.
- Council Regulation (EEC) No 2082/93 of 20 July 1993 amending Regulation (EEC) No 4253/88 laying down provisions for implementing Regulation (EEC) No 2052/88 as regards coordination of the activities of the different Structural Funds between themselves and with the operations of the European Investment Bank and the other existing financial instruments, *Official Journal* L 193, 31/07/1993, pp. 0020–0033.
- Council Regulation (EEC) No 2083/93 of 20 July 1993 amending Regulation (EEC) No 4254/88 laying down provisions for implementing Regulation (EEC) No 2052/88 as regards the European Regional Development Fund, *Official Journal* L 193, 31/07/1993, pp. 0034–0038.
- Council Regulation (EEC) No 2084/93 of 20 July 1993 amending Regulation (EEC) No 4255/88 laying down provisions for implementing Regulation (EEC) No 2052/88 as regards of the European Social Fund, *Official Journal* L 193, 31/07/1993, pp. 00039–0047.
- Council Regulation (EEC) No 2221/88 of 19 July 1988 amending Regulation (EEC) No 2727/75 on the common organization of the market in cereals, *Official Journal* L 197, 26/07/1988, pp. 0016.
- Council Regulation (EEC) No 4253/88 of 19 December 1988, laying down provisions for implementing Regulation (EEC) No 2052/88 as regards coordination of the activities of the different Structural Funds between themselves and with the operations of the European Investment Bank and the other existing financial instruments, *Official Journal* L 374, 31/12/1988, pp. 0001–0014.
- Council Regulation (EEC) No 856/84 of 31 March 1984 amending Regulation (EEC) No 804/68 on the common organization of the market in milk and milk products, *Official Journal* L 090, 01/04/1984, pp. 0010–0012.
- Cour des Comptes Européenne (1998), Avis No. 10/98 de la Cour des comptes européenne relatif à certaines propositions de règlements dans le cadre de l' Agenda 2000. Luxembourg: Cour des comptes européenne.
- De Goerter, H. and Tsur Y. (1991), Explaining Price Policy Bias in Agriculture: The Calculus of Support-Maximising Politicians. *American Journal of Agricultural Economics*, No.73(4), pp. 1244-1254.
- Diederens, P., N. van den Eeden and J. Kuper (1998), *Economische en bestuurlijke evaluatie van de Bijdrageregeling innovatieprojecten*. The Hague: LEI-DLO.

- Directie Financieel-Economische zaken (1998a), *Beleidsvaluatie-onderzoek binnen LNV, "de stand van zaken"*, notitie nr.1. The Hague: Ministerie van landbouw, natuurbeheer en visserij, 14 januari 1998.
- Directie Financieel-Economische zaken (1998b), *De visie op beleidsvaluatie-onderzoek binnen LNV*. The Hague: Ministerie van landbouw, natuurbeheer en visserij, 10 juli 1998.
- Directie Financieel-Economische zaken (1998c), *Plan van aanpak beleidsvaluatie-onderzoek binnen LNV*. The Hague: Ministerie van landbouw, natuurbeheer en visserij, 10 juni 1998.
- Economic Research Service (ERS) (1996), *Provisions of the 1996 Farm Bill - the Federal Agriculture Improvement and Reform (FAIR) Act*, Special Supplement, AGO-961, Washington, D.C.: United States Department of Agriculture.
- Europäische Kommission (1998), *Die Finanzierung der Europäischen Union*. Bericht der Kommission über das Funktionieren des Eigenmittelsystems. Brussels: European Commission.
- European Commission (1994), *The Agricultural Situation in the Community*, annual report, Brussels.
- European Commission (1995a), *Agricultural Situation and Prospects in the Central and Eastern European Countries*. (Ten country reports and one summary report.) Brussels: European Commission.
- European Commission (1995b), *Study on Alternative Strategies for the Development of Relations in the Field of Agriculture between the EU and the Associated Countries with a View to Future Accession of these Countries*. Document CSE(95)607. Brussels: European Commission.
- European Commission (1996a), *Evaluation: Concrete steps towards best practice across the Commission*, Communication to the Commission by Mr Liikanen and Mrs Gradin in agreement with the President, Internal document SEC96/659 final, Brussels: European Commission.
- European Commission (1996b), *Structural Funds and Cohesion Fund 1994-99 - Regulations and Commentary*, Luxembourg: Office for Official Publications of the European Communities.
- European Commission (1996c), *The Cork Declaration - The European Conference on Rural Development*, at Cork, Ireland from 7th to 9th November, 1996, hypertext version, DGVI website.
- European Commission (1997a), *Evaluating EU Expenditure Programmes: A Guide. Ex post and Intermediate Evaluation*. First edition, January 1997. Brussels: European Commission.
- European Commission (1997b), *Long Term Prospects - Grains, Milk and Meat Markets*. CAP 2000 Working Document. Brussels: European Commission.
- European Commission (1998a), *Prospects for Agricultural Markets 1998-2005*. Brussels: European Commission.
- European Commission (1998b), *CAP Reform Proposals: Impact Analyses*. Brussels: European Commission.
- European Court of Auditors (1998), *Observations of the Court on the envisaged reform of the Common Organisations of the Market in olive oil*. Luxembourg: European Court of Auditors.

- Fennel, R. (1987), *The Common Agricultural Policy of the European Community. Its institutional and administrative organisation*. 2nd Edition, Oxford, London, Edinburgh: BSP Professional Books.
- Folmer, C., M.A. Keyzer, M.D. Merbis, H.J.J. Stolwijk, and P.J.J. Veenendaal, *The Economics of the Common Agricultural Policy (CAP)*. European Economy, Reports and Studies, No. 5, 1994. Brussels: European Commission.
- Government of the Netherlands (1996), *European Expenditure*, Part I. Interdepartmental Policy Study, European Expenditure, Part 1: Agricultural Expenditure. Budget preparation 1996 – Report nr. 1. The Hague: Government of the Netherlands.
- Grainer M. (1996), Positioning performance Measurement for the Twenty first Century, in: Halachmi A. and Bouckaert G. *Organisational Performance and Measurement in the Public Sector, Toward Service, Effort and Accomplishment Reporting*, London: Quorum Books.
- Gray A., Jenkins B. and Glynn J. (1992), Auditing the Three Es: The Challenge of Effectiveness, *Public Policy and Administration*, Vol. 7, No. 3, Winter 1992, pp. 56-70, reprinted in *Public Sector Management*, Sir Bourn J. (Ed.) (1995), Vol. II, The International library of Management, Dartmouth, Aldershot, pp. 601-615.
- Hill, B. (1996), *Farm Incomes, Wealth and Agricultural Policy*, Second Edition, Perspectives on Europe, Contemporary Interdisciplinary Research, Avebury, Aldershot, Brookfield USA, Hong Kong, Singapore, Sydney.
- HM Treasury (1988), *Policy Evaluation: a guide for managers*, London: HMSO.
- Josling, T. and S. Tangermann (1989), Measuring Levels of Protection in Agriculture: A Survey of Approaches and Results, in: A. Maunder and A. Valdés (eds.), *Agriculture and Governments in an Interdependent World*, Proceedings of the XX International Conference of Agricultural Economists. Dartmouth, Aldershot, 1989.
- Koester, U. and S. Tangermann (1976), *Alternativen der Agrarpolitik. Eine Nutzen-Kosten-Analyse im Auftrag des Bundesministeriums für Ernährung, Landwirtschaft und Forsten*. Münster-Hiltrup: Landwirtschaftsverlag.
- Koester, U., *Policy Options for the Grain Economy of the European Community: Implications for Developing Countries*. IFPRI Research Report No. 35. Washington, D.C., 1982.
- Larsen, A., et al. (1994), *EC Agricultural Policy for the 21st Century*. European Economy, Reports and Studies, No. 4, 1994. Brussels: European Commission.
- Levy, Roger P. (1995), *Expert paper on Evaluation of the Common Agricultural Policy Expenditure*, prepared for DGXIX of the European Commission, Centre for Public Policy and Management, School of Public Administration and Law, Robert Gordon university, Aberdeen.
- MAFF (1998a), *Economic Evaluation of Policies*, Research Booklet April 1998, Economics (Resource Use) Division, Whitehall Place, London.
- MAFF (1998b), *Policy Evaluation Guidance*, Economics (Resource Use) Division, Whitehall Place, London.

- Moyer H. W. and Josling T.E. (1990), *Agricultural Policy Reform. Politics and Process in the EC and the USA*, New York, London, Toronto, Sydney, Tokyo: Harvester Wheatsheaf.
- National Audit Office (1996), *State Audit in the European Union*. London: NAO.
- National Audit Office (no date), *A framework for Value for Money Audits*, London: NAO.
- OECD (1988), *New Trends in Rural Policy Making*, Paris: OECD.
- OECD (1991), *Changes in Cereals and Dairy Policies in OECD Countries. A Model-based Analysis*. Paris: OECD.
- OECD (1994), Performance Management in Government: Performance and Allocating Resources. *Public Management Studies*, No. 3, 1994.
- OECD (1996), *Agricultural Policies, Markets and Trade in OECD Countries - Monitoring and Evaluation 1996: Summary and Conclusions; Main Report*. Paris: OECD.
- OECD (1997a), *Environmental Indicators for Agriculture*, Paris: OECD.
- OECD (1997b), *Agricultural Policies in OECD Countries: Monitoring and Evaluation 1997*. Paris: OECD.
- OECD (1998a), *Adjustment in OECD Countries: Reforming Agricultural Policy*. Paris: OECD.
- OECD (1998b), *Agricultural Policies in OECD Countries 1998: Volume I - Monitoring and Evaluation; Volume II - Measurement of Support and Background Information*. Paris: OECD.
- OECD (1998c), *Agriculture and the Environment: Issues and Policies*. Paris: OECD.
- Overberg, B. (1996), *Die Auswirkungen der Europa-Abkommen mit den mitteleuropäischen Ländern auf den Agrarhandel*. Bonn: Köllen Druck+Verlag.
- Oxley H. and Martin J.P. (1991), Controlling Government Spending and deficits: Trends in the 1980s and Prospects for the 1990s, *OECD Economic Studies*, No. 17, Autumn 1991, pp. 145-180, Paris: OECD.
- Pelkmans J. and Labory S. (1998), *European Regulation and Cost Benefit Analysis - A Methodological guide for Non Specialists*, report prepared for DGIII, CEPS, Brussels.
- Petit, M. (1985), *Determinants of Agricultural Policies in the United States and the European Community*. International Food Policy Research Institute. Research Report 51, Washington, D.C.: IFPRI.
- Petit, M., de Benedictis M., Britton, D., de Groot M., Heinrichsmeyer W., Leshi F. (1987), *Agricultural Policy Formation in the European Community: The birth of Milk Quotas and CAP Reform*, Amsterdam: Elsevier.
- Riksrevisionsverket (RRV) (1998a), *Pre-study for the Arable Area payment Scheme*. RiksRevisionVerket, 1998(40), Stockholm: Riksrevisionsverket .
- Riksrevisionsverket (RRV) (1998b) *The Arable Area Payments. Interim report on the Income Objective, 1988-11-20*, Stockholm: Riksrevisionsverket

- Sohn, W. (1984), *Versorgungssicherung als Argument in der Agrarpolitik - Ökonomische Analyse mit einer empirischen Untersuchung für die Bundesrepublik Deutschland und die Europäische Gemeinschaft*. Münster-Hiltrup: Landwirtschaftsverlag.
- Swinnen, J. (1994), A Positive Theory of Agricultural Protection, *American Journal of Agricultural Economics*, No. 76, pp.1-14.
- Tangermann, S. (1991), A Bond Scheme for Supporting Farm Incomes, in: Marsh, J., B. Green, B. Kearney, L. Mahe, S. Tangermann and S. Tarditi, *The Changing Role of the Common Agricultural Policy. The Future of Farming in Europe*. London and New York: Belhaven Press, pp. 95-96.
- Tangermann, S. (1996), Implementation of the Uruguay Round Agreement on Agriculture: issues and prospects, *Journal of Agricultural Economics* 47, No 3, pp. 315-337.
- Tangermann, S. and Josling T. (1994), *Pre-Accession Agricultural Policies for Central Europe and the European Union*, report prepared for European Commission, DGI, Brussels.
- Tomkins C.R. (1987). *Achieving Economy, Efficiency and Effectiveness in the Public Sector*. Institute of Chartered Accountant in Scotland Edinburgh: Chartered Accountants of Scotland.
- Wissenschaftlicher Beirat beim Bundesministerium für Ernährung, Landwirtschaft und Forsten (1994), *Vorschläge für eine grundlegende Reform der EG-Zuckermarktpolitik*. Angewandte Wissenschaft, Heft 430. Münster-Hiltrup: Landwirtschaftsverlag.
- Wissenschaftlicher Beirat beim Bundesministerium für Ernährung, Landwirtschaft und Forsten (1996), *Zur Neuorientierung der Landnutzung in Deutschland*. Angewandte Wissenschaft, Heft 453. Bonn: Köllen Druck+Verlag.
- Wissenschaftlicher Beirat beim Bundesministerium für Ernährung, Landwirtschaft und Forsten (1997a), *Kompetenzverteilung für die Agrarpolitik in der EU*. Angewandte Wissenschaft, Heft 468. Bonn: Köllen Druck+Verlag.
- Wissenschaftlicher Beirat beim Bundesministerium für Ernährung, Landwirtschaft und Forsten (1997b), *Zur Weiterentwicklung der EU-Agrarreform – Entkoppelung der Preisausgleichszahlungen und Umsetzung der GATT-Beschlüsse*. Angewandte Wissenschaft, Heft 459. Bonn: Köllen Druck+Verlag.
- Wissenschaftlicher Beirat beim Bundesministerium für Ernährung, Landwirtschaft und Forsten (1997c), *Die Entwicklung der Landwirtschaft in Mitteleuropa und mögliche Folgen für die EU-Agrarpolitik*. Angewandte Wissenschaft, Heft 458. Bonn: Köllen Druck+Verlag.

* * *

OFFICIALS INTERVIEWED

We greatly appreciate the information received from the following officials:

Jean-Louis CHOMEL,	European Commission, DG VI
Jörgen HOLMQUIST,	European Commission, DG VI
Svend JAKOBSEN,	European Commission, DG VI
Mathias LÖSCH,	European Commission, DG VI
Asuncion RUBIRALTA,	European Commission, DG XIX
Mr. STUYCK,	European Commission, DG XIX
Hilkka SUMMA,	European Commission, DG XIX
Jacques VONTHRON,	European Commission, DG XIX
H. SCHLAGHECK,	Federal Ministry of Food, Agriculture and Forestry of Germany
Albert VERMUE,	Ministry of Agriculture, Nature Management and Fisheries of the Netherlands
David THOMPSON,	UK Ministry of Agriculture, Fisheries and Food, Director of Economics and Statistics,
David CAWLEY,	UK Ministry of Agriculture, Fisheries and Food, Economics (Resource Use) Division, Branch B.
Nigel ATKINSON,	UK Ministry of Agriculture, Fisheries and Food, Economics (International Policy) Division.
Pamela SMITH,	UK National Audit Office, Agricultural Policy Group
Meta BOETHIUS,	Swedish National Audit Office (RRV)
Goran HYLINDER,	Swedish National Audit Office (RRV)
Thomas HAGMAN,	Swedish Ministry of Agriculture, Division for EU Co-ordination and International Affairs

* * *