

Research Evaluation to Increase Impact

D. J. Templeton

Australian Centre for International Agricultural Research

Research evaluation, undertaken throughout the life of the research activity, is an important tool for research managers. When correctly implemented, research evaluation can help to increase the success of existing and future research projects and programs, thereby increasing the social, economic and/or environmental impact of the research. The relationship between the life-cycle of the research project and the type of evaluation determines the appropriate type of evaluation to be undertaken. Given the value of research evaluation, and the recent development of user-friendly evaluation procedures and models, management of a R&D portfolio can benefit from increased evaluation effort at the project, program and organisational levels.

Key Words: Research evaluation, research-to-impact pathway

Introduction

It is widely recognised that growth in agricultural production leads to improved economic development in both developed and developing countries. By increasing productivity, agricultural research is a major source of increased agricultural production and income. In addition, agricultural research can have a significant impact on the distribution of income among different types of producers, processors and consumers, different income groups, and different geopolitical and agro-ecological localities.

Despite the high returns to agricultural research, available funds are limited. Consequently, measuring the level and distribution of returns to research has become increasingly important to assist management with decisions relating to the allocation of research funds. As a result, over the past 25 years, considerable effort has been devoted to evaluating research processes and to measuring the welfare effects of research. Knowledge gained through evaluation, about the merits and worth of project processes and outcomes, can help to improve existing and future research projects and programs.

Why Evaluation is Important for Research

There are four main reasons why evaluation of agricultural R&D projects is important in both developed and developing countries. One or more of the various forms of evaluation can be used to satisfy accountability requirements, as a decision-making tool for management, to increase awareness of the efficiency and implications of the project and to clearly identify the research-to-impact pathway.

- **Accountability**

One of the primary reasons for assessing the impact of completed agricultural research projects is to satisfy the growing demand for accountability. Increasingly, national and international research organisations are being asked to provide stakeholders (funding bodies, governmental departments and so on) with a description of the expected and unexpected research-induced impacts, evidence that the outcomes of completed projects have been achieved, and an estimate of the returns to the investment in the project. In turn, the impact assessments are used to underpin private and/or public support for continued investment.

- **Management decision-making tool**

Evaluation studies, in particular impact assessments, are undertaken to better inform managers on the complementarities and trade-offs between activities or projects within the investment portfolio. Knowledge on the returns to alternative projects can help managers to target limited R&D funds towards projects or programs that are expected to have high returns. Alternatively, where the economic return of the project is not considered to be the main priority, impact assessments are useful in providing information on the opportunity cost of investing in projects with lower returns.

- **Increasing awareness**

Evaluation studies are also undertaken with the aim of making the project team aware of the broader implications (or lack thereof) of their research and/or extension activities. These studies can be undertaken at the various stages of the project life cycle, from planning and implementation through to the adoption and use of the project results, which may not occur for a number of years after the project has been completed. The lessons learnt from such evaluations can help to improve the efficiency of current and future research projects.

- **Identifying research-to-impact pathway**

By undertaking various forms of evaluation, awareness of the research-to-impact pathway is increased and factors that affect the uptake of the research results are identified. In particular, by focussing on the users of the research outputs, rather than just the outputs themselves, an *ex ante* evaluation undertaken at the planning stage will highlight the pathways necessary to ensure that the research ultimately results in an impact. In addition, by undertaking M&E activities throughout the implementation of the research project, timely information on whether or not the research goals are being met will be readily available and weak links in the research-to-impact pathway will be identified. Finally, in an *ex post* impact assessment, information on the success of the project in terms of the delivered outcomes is provided, as well as the information on the economic returns to the research project.

Stages of Research and Types of Research Evaluation

In the conceptual framework proposed by Owen (1993), there are five types of evaluation. These are development evaluation, design evaluation, process evaluation, evaluation of program/project management and impact evaluation. In essence, these types of evaluation relate to the information required at the different stages of the research activity. The relationship between the life cycle of the research project and the type of evaluation is presented in Figure 1. This Figure, and the subsequent discussion, draws on work by Owen and Rogers (1993) and the Victorian Department of Natural Resources and Environment (NRE) (2002).

- **Development evaluation**

Development evaluation is undertaken before the research project commences, with the aim of clarifying the overall goals of the project. Typical questions that need to be addressed at this stage include: Who are the likely beneficiaries? What are their requirements/needs? What are the expected economic returns to the research?

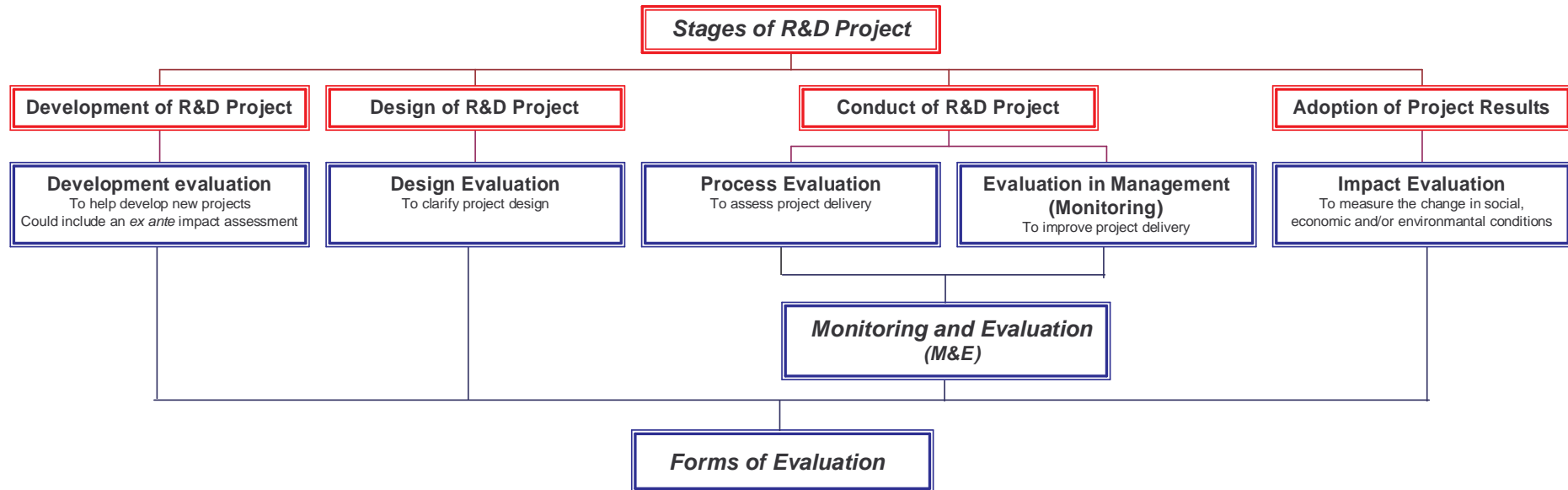
- **Design evaluation**

Design evaluation is undertaken at the project development and planning stage, and is often revisited at the start of an impact evaluation. The purpose of this form of evaluation is to determine the logic behind the project, to clarify the research-to-impact pathway and to identify any potentially weak linkages. Questions that could be asked at this stage include: Do the goals match the proposed action? What are the steps that need to be taken so the social, economic and/or environmental goals are met? What are the underlying assumptions?

- **Process evaluation and evaluation of research project management**

Process evaluation, and evaluation of the management of the research project, can be undertaken at any time in the life-cycle of the project, or it can be part of the on-going monitoring process. These two types of evaluation are often simply referred to as 'monitoring and evaluation' (or M&E). The purpose of M&E is to assess how the project is being delivered and what improvements can be made to the delivery process. The questions typically addressed in M&E activities are: Were each of the milestones met? If not, what can be done to ensure that they are met? In what ways could the project be managed better? What processes need to be changed to ensure that the ultimate goals of the project are reached?

Figure 1: Stages of Research Activity and Types of Research Evaluation



- **Impact evaluation**

The fifth type of evaluation is called impact evaluation (also commonly referred to as impact assessment). At the Consultative Group for International Agricultural Research (CGIAR) workshop on ‘The Future of Impact Assessment in the CGIAR’, an impact was defined as the broad, long-term economic, social and environmental effects resulting from research. These effects may be positive or negative and include both anticipated and unanticipated outcomes. Such effects generally involve changes in both knowledge and behaviour, which could be at the individual, organisational or institutional level. The purpose of impact assessment is to describe and assess the impact of the project on social, economic and/or environmental conditions. Impact assessments can be qualitative or quantitative, although economic impact assessments are largely quantitative providing information on the dollar returns to the investment in the research project.

Developing an Evaluation Plan

In Australia, all project proposals submitted to the Agriculture Division of the NRE are required to have an accompanying evaluation plan. This ‘evaluation culture’ is spreading throughout the rest of the Department resulting in high attendance at the evaluation workshops, which are held by the Agriculture Division’s Evaluation Unit. Drawing heavily on material provided at the Project Evaluation Training Workshop, the six steps for developing an evaluation plan (ranging from clarifying the outcomes to documenting the evaluation plan) are outlined briefly here.

- **Step 1 Clarify the Outcomes**

The first step towards developing an evaluation plan should be taken at the developmental stage of the research project. This step consists of three components. The first component is to clarify the projects goals and show linkages. Even if the goals of the project have been specified before the evaluation plan is developed, it may be worthwhile revisiting the project goals during the evaluation planning process. In addition, it is useful to show how the goals of this project could be linked to higher-order objectives. Take for example a research project on breeding heavier pigs. The aim of this project may be to increase pork production, while a higher-order objective maybe to increase the exports of pig meat and an even higher-order objective may be to increase the welfare of poor pig producers.

The second component is to identify the key users (or target group(s)) and their requirements. For example, the target group may be either extension workers, farmers, processor, exporters or policy makers, or it maybe a combination of these groups. In the evaluation literature, the key users are generically referred to as ‘next users’ and may have specific requirements that need to be met before they can learn about and adopt the outputs of the research project. For example, a smallholder may need to (a) be aware that a specific scientific project has developed pigs which produce more meat, and (b) have access to financial resources, before being able to take advantage of the research results.

The final component of this first step to developing an evaluation plan is to obtain a clear picture of what the success of the project will look like. A clear understanding of what *change* the project is trying to achieve is essential not only for R&D evaluation but also for agricultural research management. For example, in the case of the pig research project, the desired change may not be just to produce pigs with an increased carcass weight, but rather to increase the income levels of poor rural pig producers.

- **Step 2 Draw up the Program Logic**

Program logic identifies the cause-effect relationships between inputs, outputs, intermediate outcomes and final impacts. It is described as a form of design evaluation by Owen (1993), although it is commonly referred to as program logic in the international literature on evaluation even though it can be applied at the project or sub-project level.

A program logic model is developed for two main reasons. First, to clarify and evaluate the logic of the research project, often done in the developmental stage. Second, to provide a framework to evaluate the performance of the project. Drawing up a program logic model can help to clarify the aims of the project, identify and describe the major elements of the project, identify the expected cause-effect relationships and identify the key areas for evaluation. As such, a program logic model is a very useful tool in agricultural research management.

While there are various methods of program logic, each with a slightly different emphasis, the preferred form used in the Agricultural Division of the NRE is Bennett's Hierarchy. This is because Bennett's Hierarchy requires researchers to think of the expected users of the technology, their characteristics and how these characteristics may influence technology uptake. In other words, Bennett's Hierarchy is very much people (or goal), rather than output, focussed. It is also relatively easy to use and widely applicable. It can be used throughout the project life cycle and provides a way of aggregating information from sub-project level to project level to program or strategy level. Given the value of Bennett's Hierarchy as a research management tool, it is discussed in greater detail in the following section.

- **Step 3 Develop Key Evaluation Questions**

Key evaluation questions are focussed and open questions, and form the basis of the data requirements for an evaluation study. For example, in the case of the research project on breeding larger pigs, key evaluation questions could include: What are the returns to the research? How much better off are smallholder pig producers as a result of this project?

Not surprisingly, when developing key evaluation questions, the evaluation audience or stakeholders and their requirements need to be considered.

- **Step 4 Complete the Evaluation Schedule**

This step involves recording the key evaluation questions and sub-questions. At this stage revisions to the key evaluation questions may be necessary.

It also involves selecting a suitable method or methods for collecting the data necessary to answer the key evaluation questions. Ultimately the method(s) chosen will not depend only on the key evaluation questions but also on other factors that need to be considered when managing any research project. These factors include the cost of the method, the timeframe of the research project and the resources and skills available to the project manager. In addition, the biases of the alternative methods need to be considered. Some commonly used methods for collecting the necessary information include; goal attainment scaling, observation, questionnaires, focus groups, and semi-structured interviews. Clearly, it is much easier to evaluate a research project, and to undertake an *ex post* impact assessment, if consideration of the data required for evaluation is given throughout the life of the project.

The final requirement of this step is to determine who will conduct the evaluation work and when. For example, M&E can readily be undertaken by the project team throughout the project. On the other hand, *ex post* economic impact assessments are undertaken after the project is completed and often by an external analyst.

- **Step 5 Develop Management and Evaluation Strategy**

This step involves considering how the evaluation will be managed and how to maximise the use of the evaluation. The management of project evaluation should be governed by the principals and practices of good project management. Special considerations relevant to the management of evaluation include; developing an evaluation plan, considering who should undertake the evaluation, considering what resources are required, considering the timing of the evaluation, and reviewing the evaluation.

A strategy aimed at maximising the use of the evaluation should also be developed as part of the evaluation plan. One of the biggest failings in evaluation is that the findings are frequently under-utilised (NRE 2002). Given that evaluation studies can provide the project team, program managers and other important stakeholders with valuable information, within the framework of sound research management, information gained from the evaluation study should be reported in a way that will ensure it is used.

- **Step 6 Document the Evaluation Plan**

The final step is to document the evaluation plan. This task, which is relatively simple and just as important as documenting the project proposal, consists of four sections. The first section is the introduction. In this section, the project is described, the key users are identified and a description what success looks like is given. The second section covers program logic, where

the theory of action for the project is presented and any assumptions are highlighted. The third section is the schedule. In this section the key evaluation questions are documented, the data needs and data collection methods are described, and the audience and reporting requirements are discussed. In the final section, the management and utilisation strategy is presented.

Bennett's Hierarchy

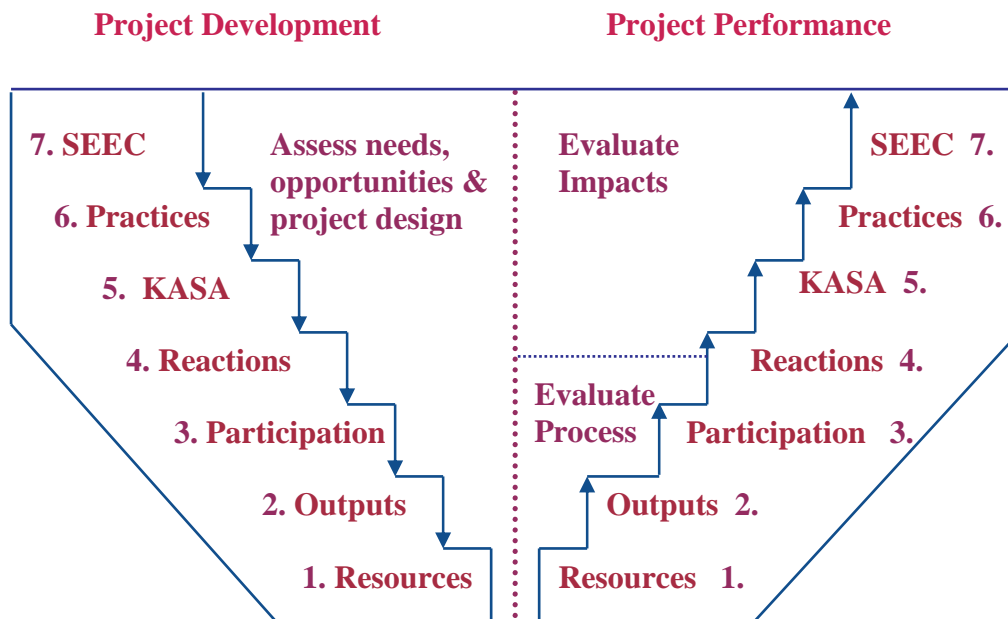
Bennett's Hierarchy is a goal-based approach to evaluation. It was originally developed in 1977 as an 'outcome' hierarchy for extension activities (Bennett 1977). It is now a commonly-used framework in extension evaluation and, as stated earlier, it is the preferred form of program logic used by the Agricultural Division of the NRE. While Bennett's Hierarchy was created for extension projects, modified versions of Bennett's have been developed within the NRE so it can be applied to projects that are not extension focused.

There are four main reasons why Bennett's Hierarchy is a commonly used form of program logic. First, it directs attention towards the expected users of the project results, and how their characteristics (such as their attitudes, knowledge and skill levels, values and so on) are likely to influence technology adoption. Second, it is relatively easy to use. Third, it can be used in planning, guiding evaluation and project reporting. Finally, it offers a means of aggregating information from sub-project to key project and from key project to program or strategy level.

Bennett's Hierarchy describes a simplified cause-and-effect chain of events through project identification at seven levels. At the first level, the resources expended by the project are depicted. These resources produce research findings (in the case of an R&D project) and/or activities (such as dissemination or extension) (level two) which involve people (level three) with certain characteristics. These people will react (level four) to the project, resulting in a ***change*** in their knowledge, aspirations, skills and/or attitudes (level five). These changes could result in a ***change*** in farm practices (level six) thereby achieving the end result or ultimate goal(s) of the project, which is a ***change*** in economic, social and/or environmental conditions (level seven).

Bennett and Rockwell (2000) describe how to use the hierarchy for both planning and evaluation purposes. As can be seen in Figure 2, the hierarchy depicts seven levels or steps to project development and project performance. By considering each of these at both development and performance, the research-to-impact pathway is explicitly accounted for throughout the project life-cycle. In addition, process and impact evaluation become an integral part of managing the research project.

Figure 2: *Bennett's Hierarchy - Integrated Project Development with Process and Impact Evaluation*



While Bennett's Hierarchy is relatively easy to implement, there are some traps that analysts can fall into. To reduce the number of potential errors in implementation, there are four tips that are worth bearing in mind. First, the next users of the project results need to be correctly identified and depicted in level three. People that need to collaborate as part of the project are not necessarily the next users, and as such should be documented in level two. Second, it is common that there is a huge leap between the practice change (level six) and the change in economic, social and environmental conditions (level seven). If this is the case, it may be necessary to document several steps as part of level seven. Third, make sure that enough detail is given at each level. Where appropriate, specify expected outcomes in terms of when, where, and how. Conversely, avoid too much detail so the hierarchy can fit on one page, if possible. Finally, avoid jargon so the hierarchy is understood not only by the project team but also by other stakeholders, such as the program manager, policy makers, funding bodies and so on.

Evaluating the Impact of Research within the Economic Surplus Modelling Framework

Introduction

In recent years, substantial effort has been devoted to measuring the welfare effects of agricultural research within a partial-equilibrium framework that uses the concept of economic surplus. Procedures and formulae for measuring the size and distribution of returns to agricultural research, in terms of changes in economic surplus, have largely been derived from welfare economics theory and the literature on modelling market

displacements. The literature now contains a wealth of information on how to examine the economic impact of a research-induced technical change for a range of market situations. The widespread applicability of the economic surplus approach is evident from the number of empirical studies in which it has been used. (For details see Alston, Norton and Pardey 1995)

In this section, a description of the set of principles that underlie the economic surplus approach is provided. Because of the extensive literature available, this description is brief and is limited to outlining the basic economic surplus model. A user-friendly impact assessment software package called DREAM (Dynamic Research Evaluation for Managers) is presented in the next section.

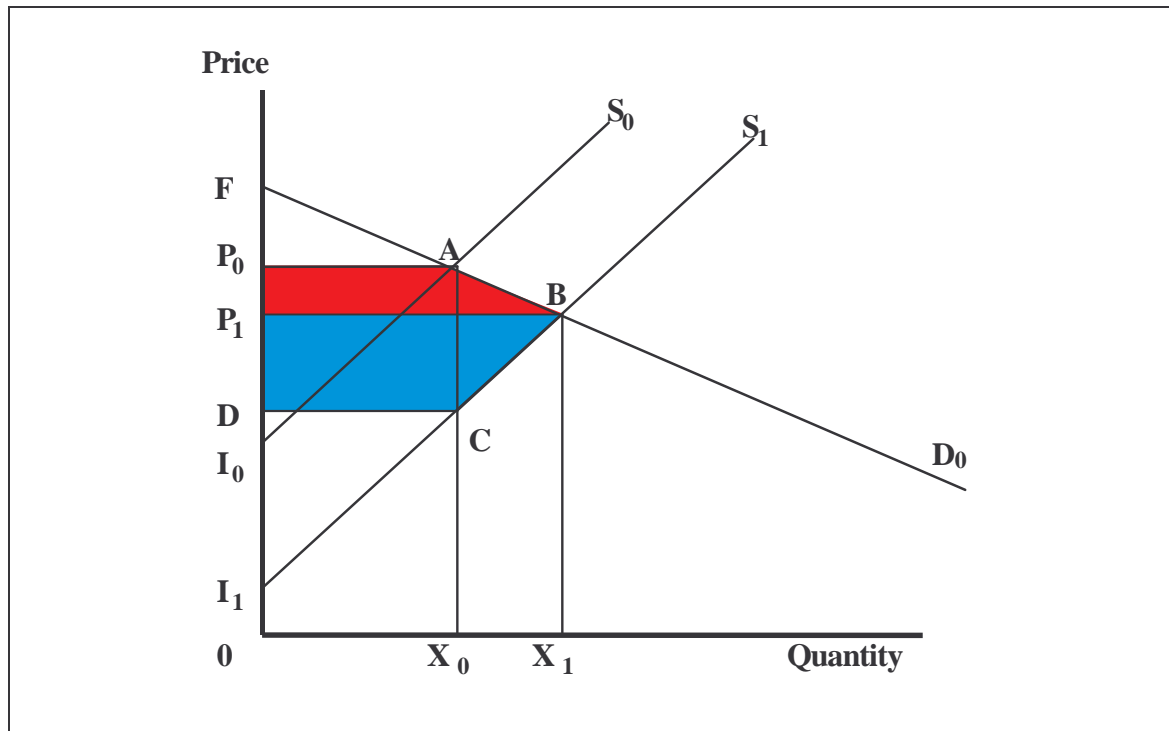
The Basic Economic Surplus Model

The economic surplus approach to welfare measurement is based on the theory and assumptions underlying welfare economics (Alston 1991). The concept of consumer surplus was first defined by Dupuit (1844, p. 29) as 'the difference between the sacrifice which the purchaser would be willing to make and the purchase price he has to pay in exchange'. Following this definition, consumer surplus is the area below the demand curve for the commodity and above the market price line.

Marshall (1893) first introduced the concept of producer surplus in the late 19th Century. He formalised the notion that when a seller makes a sale, the individual generally receives a surplus from that transaction. In other words, by selling a particular commodity, the seller obtains something, which is of greater (direct or indirect) utility to the seller than the utility that the seller would have derived if the commodity had not been sold. Producer surplus is measured as the area below the market price line and above the commodity supply curve. Total economic benefits are measured as the summation of the consumer surplus and producer surplus areas.

In addition to providing information on equilibrium prices and quantities and economic welfare, an economic supply and demand model can also be used to show the effects of research on the level and distribution of the change in economic surplus or welfare. The basic economic surplus model, presented in Figure 3, is used to show the impact of a research-induced shift in supply on producer and consumer welfare. The initial equilibrium price and quantity are P_0 and X_0 , respectively. Consumer surplus is equal to FAP_0 and the producer surplus is equal to P_0AI_0 (total revenue, P_0AX_0 , less the total cost of production, I_0AX_0). Total welfare is equal to FAI_0 , which is the sum of consumer and producer surplus.

Figure 3: Welfare Effects in a Single Market Context



Cost-reducing (or yield-increasing) research will result in a rightward shift in the supply curve, say from S_0 to S_1 , resulting in a new equilibrium price (P_1) and quantity (X_1). Because of the changes in the equilibrium prices and quantities, there will also be changes in the level of welfare accruing to producers and consumers and, therefore, a change in total economic welfare. The change in consumer surplus (ΔCS) from the research-induced supply shift is equal to the area P_0ABP_1 . The change in producer surplus (ΔPS) is equal to the area P_1BI_1 minus the area P_0AI_0 , which, in the case of linear supply curves moving in a parallel fashion, is equal to the area P_1BCD . The change in total economic surplus (ΔTS) is equal to the area I_0ABI_1 that, in the case of a parallel supply shift, is equal to the area P_0ABCD .

So long as the demand curve is less than perfectly elastic, consumers will gain from a cost-reducing technical change because they consume more of the product at a lower price. Even when the demand curve is perfectly elastic, consumers will be no worse off. Again in the case of a parallel supply shift, producers will also gain from the adoption of a cost-reducing technology so long as the reduction in the price of their product is more than offset by the vertical downward shift in the supply curve (i.e, the reduction in the cost of production). That is, so long as demand is not perfectly inelastic, in which case they would be no worse off. Even when the demand curve is inelastic, Wohlgenant (1997) has shown that industry marginal returns to yield-increasing research can be positive.

The changes in economic welfare can be expressed algebraically as follows:

$$\Delta CS = P_0X_0Z(1+ 0.5Z\eta)$$

$$\Delta PS = P_0 X_0 (K - Z)(1 + 0.5Z\eta)$$

$$\Delta TS = \Delta CS + \Delta PS = P_0 X_0 K(1 + 0.5Z\eta)$$

where K is equal to the vertical research-induced shift in the supply curve measured as a proportion of the initial equilibrium price, $Z = K\varepsilon/(\varepsilon+\eta)$ and ε and η are the elasticity of supply and the absolute value of the elasticity of demand, respectively (Alston, Norton and Pardey 1995, p. 211).

A number of additional points regarding the nature of the economic surplus model warrant mentioning. First, it is a static model. Therefore, while successful investment in a particular agricultural project will usually result in a stream of benefits over time, the research benefits, measured as changes in economic surplus from the model, represent gross returns to research for a specified time period. As such, they do not provide complete answers to questions regarding the flow of net returns to research. A complete analysis of the flow of net benefits over time requires that estimates of research costs, adoption rates and the life span of the new technology also be included in the analysis (Alston, Norton and Pardey 1995).

Second, the gross surplus changes are commonly measured off medium- or long-run supply curves, with the dynamic aspects such as lagged adjustment in supply to changes in prices largely being ignored.

Third, the economic surplus model presented above is based on the premise that research-induced price and quantity changes occur only in the market of interest. In other words, it is assumed that research on a particular commodity (say, pork), will only result in a change in the production and price of that commodity and will not affect the production or price of any other commodity (for example, beef).

The DREAM Model for Impact Assessment

While the basic economic surplus model can adequately deal with issues regarding the level and distribution of benefits from research in a single market setting, questions regarding the distribution of those benefits between disaggregated groups of producers, suppliers of inputs, consumers, regions or nations cannot be answered. To disaggregate the measures of benefits from research, the analysis needs to be extended to account for either vertical or horizontal market integration, or both. This can prove to be a non-trivial task, particularly for complex market scenarios. Fortunately, a user-friendly system for estimating the level and distribution of the economic benefits of agricultural R&D, called DREAM, is available from the International Food Policy Research Institute web site (<http://www.ifpri.cgiar.org/dream.htm>) free of charge. DREAM is based upon the concepts and methods described in Alston, Norton and Pardey (1995).

DREAM is designed to evaluate the economic impacts of agricultural R&D for a broad range of policy, market, technology, and adoption conditions. The objective is

to provide R&D analysts with a practical means of generating information to support strategic decision making.

DREAM focuses primarily on the evaluation of new technologies or practices applicable at the farm level. But while the immediate impacts of R&D often arise from technology-induced changes in outputs and costs at the farm level, the broader economic effects also depend upon a range of biophysical, social, and market factors, for which DREAM requires the user to provide quantitative estimates. Like any model, the results obtained from DREAM will only be as good as the data put in. (For a good discussion on data-collection methods see Alston, Norton and Pardey 1995.)

DREAM can handle very simple to quite complex problems. Analytical options include multiple regions, supply and demand dynamics, and a range of options to represent technology transfer and adoption. Thus, DREAM provides a framework for exploring various kinds of policy, technology, extension, and trade issues. The challenge for the analyst is to develop a clear understanding of the model's capabilities, assumptions, and limitations. Without such an understanding it is difficult to formulate real-world problems in relevant analytical terms and to properly interpret the resulting outputs.

Summary and Conclusions

Given that the funds available to undertake agricultural research are limited, being able to evaluate the merit and worth of project process and outcomes, in a theoretically-consistent manner, is important. Following the work by Owen (1993), it is now recognized that there are five types of research evaluation that could be applied to various stages of a project from planning through to post-completion. The five forms of evaluation are development evaluation, design evaluation, process evaluation, evaluation of research management and impact evaluation. As the names suggest, each type of evaluation relates to various stages in project development.

Evaluation of agricultural R&D projects and/or programs is important to meet growing accountability requirements, as a research management tool, to increase awareness of the implications of the project, and to identify the research-to-impact pathway. Developing an evaluation plan is an important first step to evaluation. It aids in clarifying the outcomes, drawing up the program logic, developing the key evaluation question and developing management and evaluation strategies.

A program logic model is a useful component of project evaluation and an important part of a sound research-management strategy. Program logic identifies the cause-and-effect relationship between inputs, outputs, intermediate outcomes and final outcomes or impacts. A widely-used, and widely-applicable, form of program logic is the goal-based approach to evaluation called Bennett's Hierarchy. It can be readily used in project planning, for guiding evaluation and for reporting purposes.

In line with the increased demand for accountability, substantial effort has been devoted to measuring the impact of agricultural research on the level and distribution of consumer and producer welfare. Evaluating of returns to research within an economic surplus framework is now commonplace. The DREAM model has been

developed as an impact-assessment tool, which can be used to aid decision-making in agricultural research management.

In sum, given the value of research evaluation, and the recent development of user-friendly evaluation procedures and models, it is clear that research management will benefit from increased evaluation effort at the project, program and organisational levels.

References

- Alston, J.M. 1991, 'Research benefits in a multimarket setting: a review', *Review of Marketing and Agricultural Economics* 59(1), 23-52.
- Alston, J.M., Norton, G.W. and Pardey, P.G. 1995, *Science Under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting*. Cornell University Press, Ithaca, New York, (published in cooperation with the International Service for National Agricultural Research).
- Bennett 1977, *Analysing impacts of extension programs*. Extension Service, US Department of Agriculture, Washington DC.
- Bennett, C.F. and Rockwell C. 2000, *Targeting Outcomes of Program (TOP)*, <http://deal.unl.edu/TOP/synopsis.html>.
- Dupuit, J. 1844, 'On the measurement of the utility of public works', *Annales des Ponts et Chaussées*, Second Series, Vol. 8; translation reprinted in D. Munby, *Transport*, 1968.
- Marshall, A. 1893, 'On Rent', *Economic Journal*, 3 March.
- NRE 2002, *Project evaluation training workshop: part A*, Daylesford, 21-23 May 2002.
- Owen, J.M. 1993, *Program Evaluation*. Allen and Unwin Pty Ltd, Sydney
- Owen, J.M. and Rogers, P.J. 1999, *Program Evaluation- Forms and Approaches*. Allen and Unwin Pty Ltd, Sydney
- Wohlgenant, M.K. 1997, 'The nature of the research-induced supply shift', *Australian Journal of Agricultural and Resource Economics*, 41(3), 385-400.