CHAPTER 1

Introduction

Anwar Shah

Evaluation as a discipline, in recent decades, has progressed from providing answers to questions like economic viability, cost-effectiveness, and efficiency to questions that are critical for effective planning, financing, design, implementation, and success of a program. Today evaluation can provide valuable assistance in defining a problem, identifying program targets, designing of interventions, identifying winners and losers and organizational strengths and weaknesses, assessing the quality of interventions and performance of program delivery and its impact, suggesting modifications and alterations, and ultimately guiding a program or project to its successful end. Evaluation is no longer just a tool for program or project appraisals but also an important tool in operational management. It is not just a snapshot of the end; it is also the means to the end (see also Williams and Giardina 1993). This book provides an easily comprehensible and comprehensive survey of tools of analysis that are used in the evaluation literature to evaluate public projects, programs, policies, and policy analysis and advice. The following paragraphs provide an overview of the book.

Chapter 2 by Deb and Shah provides a brief survey of the program evaluation methods, their objectives, strengths, and weaknesses. The

A. Shah (✉)
Governance Studies, Brookings Institution, Washington, DC, USA

© The Author(s) 2020

A. Shah (ed.), *Policy, Program and Project Evaluation*,
https://doi.org/10.1007/978-3-030-48567-2_1

methods have been presented in a manner that reflects the changes in outlook towards public programs and the changing role of evaluation. Accordingly, we have first presented methods like cost-benefit analysis, cost-effectiveness analysis, the social marginal cost of funds analysis, and data envelopment analysis that are appropriate for guiding efficient allocation and utilization of resources. This is followed by a discussion on multiple-objective evaluation, which is much more holistic. Apart from efficiency, it addresses issues like the relevance of a program, effectiveness of a program in achieving its objectives and sustainability of the program benefits. Newer multi-criteria approaches such as the Iron Triangle, Alternate Service Delivery Framework (ASDF), and the Results-Oriented Management and Evaluation (ROME) are briefly sketched. The Iron Triangle notes three important constraints faced by public managers—time, cost, and quality. The ASDF brings into sharper focus the role of government as a catalyst in managing and coordinating service delivery by government and beyond government providers. ROME is particularly noteworthy for recognizing the role of citizens as the principals and various orders of governments as their agents. It provides an integrated approach to managing for results, citizens-based evaluation of those results, and the processes to hold the government to account by citizens. Finally, the authors outline the theory-based evaluation approach. It is a relatively new concept in evaluation literature where the focus is not just on whether a program succeeds or fails but also on how and why a program succeeds or fails.

The chapter concludes that the methods discussed serve different purposes. Methods like cost-benefit analysis, cost-effectiveness analysis, and data envelopment analysis address the question of efficiency in the allocation and utilization of funds. In multiple-objective evaluation, the emphasis is more on accountability in public sector programs, the effectiveness of programs, and the sustainability of program benefits. The evaluation is concerned with issues like identification of program beneficiaries, assessing their requirements, tailoring of interventions to meet those requirements, monitoring of interventions to ensure that the appropriate interventions are being delivered to the participants, and finally the overall effectiveness of the interventions in achieving the program objectives. So, a multiple-objectives evaluation plays a big role in program planning, design, and implementation. A theory-based evaluation assigns an even bigger role to an evaluation in public programs. It goes deeper into the mechanism through which the interventions bring about the desired effects. It analyzes the causal links between interventions and outcomes. So, instead of passing a summative judgment on whether a program succeeded or failed in achieving its objectives, the theory-based evaluation shows *why* it succeeded or failed. Thus, it contributes to the development of more effective programs in the future.

The authors argue that evaluation is different from other social research in the sense that it derives its questions from policymakers, program sponsors, program managers, and stakeholders. So, the applicability of any specific evaluation method depends on the questions that the evaluator has been asked to address. When the evaluation question is deciding upon alternative interventions aimed at producing similar effects, cost-effectiveness analysis might be more suitable. But it is not useful at all when the problem is prioritizing among different programs addressing different problems. A cost-benefit analysis will be more appropriate in that case. Similarly, data envelopment analysis might not have the valuation problems associated with the cost-benefit or cost-effectiveness analysis. But its applicability is limited to comparing efficiencies of similar programs only. For programs whose efficacies have already been established, multiple objective evaluations might be enough for performance evaluations. However, for pilot studies or for programs that have not been tested before, a theory-based evaluation is much more desirable. But it is also more time consuming and more expensive than any other evaluation methods. So, the choice of the evaluation method would also depend upon the availability of time and resources.

Chapter 3 by Robin Boadway summarizes the principles used to evaluate projects (such as individual investment projects, general expenditure programs, and the implementation of government policies) from an economic point of view. Following the principles of welfare economics, the objective of project evaluation is to measure the costs and benefits to individuals in society. The chapter begins with some theoretical perspectives on cost-benefit analysis and then proceeds to discuss the difficulties of implementing such an analysis and practical ways of dealing with these problems. He describes project evaluation as "an art, though one with scientific underpinnings."

Broadly, the measurement of costs and benefits amounts to a measurement of individuals' "willingness to pay." Two methods that have been developed in the theoretical literature are "compensating variation" and "equivalent variation," depending on whether one wishes to use final or initial prices (respectively) for goods affected by the project. In addition, one must consider that projects that have costs and benefits spread over time must use a common set of prices to adjust for inflation and the time value of money. Policymakers may also wish to add distributive weights—in other words weighting more heavily improvements in the incomes of the poorest. Once the discounted stream of costs and benefits is summed (yielding the Net Present Value, or NPV), the decision rule for project approval is simply determined by a (positive) sign of the NPV. Finally, the risk and uncertainty of outcomes should also be included. Other, similar techniques for project evaluation include the benefit-cost ratio and the internal rate of return methods. However, these alternative measures have some problems and may rank projects differently than the NPV criterion.

While the NPV method is in principle the same as is used by the private sector to guide the choice of investment decisions, the implementation of this procedure differs in some important ways when applied to the public sector. This is because the public sector must take into consideration: the marginal social values or shadow prices (rather than just market prices, since markets may be distorted) of inputs including labor, capital, and foreign exchange, and the impact of externalities such as pollution, general equilibrium effects of the project, valuation of intangible benefits and costs (such as time saved due to public transport), excess burden of public financing (due to distortions of the tax system—the so-called marginal cost of public funds, or MCF), the social discount rate, and social considerations (such as equity or protection of special groups). Each of these issues is considered in turn in this chapter.

Chapter 4 by Bev Dahlby probes more deeply into an evaluation tool that received brief treatment in Chaps. 1 and 2—the Marginal Cost of Public Funds (MCF), or the loss to consumers and producers caused by raising an additional dollar of tax revenue. Taxes impose a cost on the economy if they alter taxpayers' consumption, production, and asset allocation decisions, leading to a less-efficient allocation of resources. Raising an additional dollar of tax revenue costs the private sector more than a dollar if the allocation of resources in the economy becomes more distorted. The marginal cost of public funds, MCF, is a measure of the cost imposed on the private sector in raising an additional dollar of tax revenue. Dahlby argues that the marginal cost of public funds should be used in evaluating the opportunity cost of financing public sector expenditures. It also provides a guide for tax reform by revealing which taxes impose the greatest welfare losses in generating additional revenues. The MCF can also be used to measure the gains from tax reforms that shift the burden from the high-cost tax bases to ones with lower costs.

In this chapter, Dahlby introduces the concept of the marginal cost of public funds (MCF), examples of how the MCFs can be measured, and examples of how they can be applied to guide tax reform and public expenditure policies. In keeping with the theme of this volume, the emphasis is on using the MCF as a tool for public policy analysis. Two practical examples demonstrate the use of this evaluation tool in public policy analysis and evaluation. The first is an assessment of the gains from a tax reform that shifts some of the tax burdens from corporate to the personal income tax base in Alberta, Canada. The second example illustrates how the MCF can be used to derive optimal matching rates for intergovernmental infrastructure grants in a federation. In presenting these applications, the author makes a convincing case of the critical importance of the MCF as a tool for the evaluation of tax and expenditure policies.

Theory-based evaluation traces the factors that contributed to a specific outcome for the project. This a challenging task for project evaluation as it requires establishing a causal chain based upon theoretical considerations and analyzing data on various links in this causal chain. Ewa Tomaszewska in Chap. 5 provides a guidebook for those interested in conducting case studies in project evaluation determining the impact of various projects in combating corruption. The author provides a conceptual framework and data requirements for such evaluation case studies for privation programs, judicial and legal reforms, civil service reforms, trade liberalization, tax administration reforms, and direct anticorruption activities (anti-corruption agencies), the Office of Ombudsman, transparency rules and decentralization.

The impact of corruption on public service delivery performance and poverty alleviation is widely recognized. A wide consensus has also recently emerged that corruption is a symptom of failed governance and hence curtailing corruption requires addressing the causes of misgovernance. Nevertheless, the menu of potential actions to curtail corruption is very large so a framework is needed that provides guidance on ordering potential actions. Prioritization of various actions depends on both the conceptual and empirical views of what works and what does not work in the context of particular countries. Such a framework is also needed for evaluating country anti-corruption programs and policies. Chapter 6 by Huther and Shah proposes a framework for such evaluations. The chapter concludes that path dependency is critical in determining the relative efficacy of various anti-corruption programs. For example, in a largely corruption-free environment, anti-corruption agencies, ethics offices, and ombudsmen serve to enhance the standards of accountability. In countries with endemic corruption, the same institutions serve a function in form only and not in substance. Under a best-case scenario, these institutions might be helpful, but the more likely outcome is that they help to preserve the existing system of social injustice. Successful anti-corruption programs are those which address the underlying governance failures, resulting in lower opportunities for gain and a greater likelihood of sanctions. Thus, programs must be targeted to a country's existing quality of governance. Past experiences of the industrialized world confirm these conclusions since, without exception, these countries did not achieve a reduction in corruption by introducing technocratic solutions but, rather, by encouraging a sense of public duty among officials through accountability for results. Such an accountability culture came about by empowering people and by decentralizing decision making. These conclusions suggest the following stylized presentation of anti-corruption measures based on the existing quality of governance. Addressing the governance failures which distort officials' cost-benefit assessment is likely to be the only route to success in countries with high levels of corruption and poor governance since direct dialogue on corruption is likely to be counter-productive (resulting in simply another level of corrupt officials under the name of anti-corruption offices). In countries with poor governance quality, external advice can promote economic liberalization, judicial reform, and greater public participation in public expenditure decisions without explicitly raising contentious issues of corruption and, one hopes, without threatening their existing relationships. In countries with modest levels of corruption and governance quality, where the existing governance structure has the capacity to reform, it is an important focus on improvements in readily identifiable output indicators rather than uncertain measures of corruption as measures of success. In countries with high governance quality, explicit efforts to reduce corruption are likely to be successful—commissions on corruption, ombudsmen, ethics offices, and the like can rely on an infrastructure of public accountability and transparency to ensure that their findings result in lower incentives to commit corrupt acts.

Public Expenditure Review (PER) is a widely used tool by the development assistance community to develop advice on budgetary institutions and allocations. This tool has also been used by both industrial and developing countries as an aid to public sector reforms. The most common format of PERs begins with a presentation of an overall picture of the country's fiscal situation. This picture typically focuses on the country's expenditure trends. This presentation provides the background, and frequently the justification, for the specific issues addressed by the review. The picture of the fiscal situation is frequently followed by an analysis of the budget process which typically provides the foil for recommendations made in the PER. In some cases, providing a picture of the country's fiscal situation and outlining the budget process may be the only tasks undertaken in the PER. In other cases, PERs review selected inter- and intra-sectoral issues. Almost all formal, and many informal, reviews also include extensive data on a country's expenditures. Chapter 7 by Huther and Shah presents a framework for evaluating the quality and timeliness of PERs as well as conducting a review of their impact. They highlight specific elements to consider and to rate in such an analysis and how to develop final cost-efficiency and benefit-cost ratings.

Given that the empirical evidence on the relationship between government expenditures and economic growth is inconclusive, Stuart Landon in Chap. 8 asks if the composition of expenditures and design of programs is a better determinant of the effectiveness of government expenditures than the size of the public sector. Although there are no universal rules, he suggests that a review of the empirical evidence drawn from a broad number of countries can help to identify the sectors that should generally receive the highest priority in government budgets.

In theory, government intervention can improve welfare in the presence of market failure (such as insufficient competition and incomplete markets arising from public goods and asymmetric information). However, the nature and magnitude of the market failure must be known when designing the program to ensure that the costs of intervention are justified and that the government has the capacity to successfully carry it out. In states with weak capacity, government intervention may be more harmful than the market failure that the program was intended to address. Given that state capacity is a critical input for the successful implementation of government programs (and therefore the effectiveness of public expenditures generally), it is worthwhile to invest in the capacity of public administration and reduce the size while improving the quality of the civil service (though this may be politically difficult) and developing a system to monitor expenditure effectiveness. Incentives for improved public sector performance may be created by increasing wages and reducing wage compression when coupled with mechanisms that create accountability for performance. Other tools for improving the public sector include subjecting the public sector to greater competition, involving the private sector for provision, implementing user fees.

Landon also deals with sector-specific issues, including law and security, military spending, infrastructure, transportation, operation and maintenance expenditures, education, health, redistribution, regulation, financial markets, state-owned enterprises, and industrial subsidies. For each of these, he discusses whether the sector should be a priority in the public budget as well as ways to make expenditures in each area more effective. He also considers the efficiency and equity effects of expenditures in addition to the likely re-distributional consequences. Based on a review of the literature, Landon assigns a low priority to subsidies to many state-owned enterprises and private industry, poorly targeted consumption subsidies, infrastructure that could be undertaken by the private sector, social security programs, tertiary education and hospital care, military spending, and extensive regulatory regimes. High priority should be assigned to develop an effective legal system to protect and enforce property rights, effective financial regulation, maintenance of existing infrastructure, and transportation, communications, and electricity infrastructure. Spending priority should also be given to improvements in the quality and quantity of primary education and basic health care, water and sanitation, and well-targeted consumption subsidy programs for the very poor.

Governments are becoming more decentralized. Political power and public decision making in many countries around the world have, to varying degrees, shifted away from central governments, particularly over the past quarter-century. This movement has been attributed to various forces; for example, the growing number of democracies, urbanization, increasing literacy, rising incomes, a growing middle class, and the failures of central governments. The World Bank has been involved in this transition in developing countries. Given the extent of the movement occurring and the Bank's initiatives, there is a natural wish to assess the Bank's activities regarding decentralization. A comprehensive evaluation of the Bank's undertakings in this area creates the opportunity to understand better the potential for and limitations of decentralization, to identify the strengths and weaknesses of Bank activities and practices relating to decentralization, and to assist in refining Bank policies concerning decentralization. Essentially, an evaluation is to generate information that will help the Bank's decentralization policies, programs, and practices to be more successful. Chapter 9 by McMillan outlines an approach in evaluating the Bank's decentralization initiatives. The basic methodology for evaluating individual decentralization projects is outlined with elaboration components of decentralization and steps to be taken in the evaluation. The method for extending the evaluation across many projects follows. The problem of selecting or sampling the projects to be evaluated is discussed. Special treatment is given to projects on community-driven development. The chapter also presents thoughts on bringing the various analyses and the analysis of a rather diverse set of projects together and putting the results into perspective.

Chapter 10 by Shah evaluates the conditionality of development assistance in terms of its intended and unintended consequences. Development assistance is motivated by altruistic, economic, political, military, and humanitarian considerations. It is used to advance wide-ranging objectives such as minimizing risks for loan repayment, efficiency, equity of the public sector, overcoming infrastructure deficiencies, promoting growth, facilitating poverty alleviation and good governance, combating terrorism, support for a specific ideology, influence peddling, and economic and political imperialism. The provision of such assistance is often conditional as even unconditional assistance almost always carries some explicit preconditions and implicit conditions. Conditions are imposed as part of lending or grant assistance unilaterally or by mutual agreement of the donor and the recipient. These conditions form the contractual terms of such assistance which bind the recipient to expected actions or results as a quid pro quo for receiving such financial assistance. These conditions can vary from being very vague to extremely clear and precise. They may impose formal binding requirements or simply indicate informal non-binding expectations. The chapter provides conceptual perspectives from game theory, public choice, fiscal federalism, political economy, new institutional economics, and New Public Management literature on the design of external assistance and its potential impacts. It shows how the neglect of these conceptual considerations can result in a lack of effectiveness of aid conditionality and waste of such assistance. It provides an overview of the historical evolution of perspectives on donor-recipient relations and on the conditionality of external assistance. It highlights the developing consensus by the development assistance community on both the instruments of development finance and associated conditions. It also briefly notes progress, or lack thereof, for practice to conform to emerging consensus. It cites examples where the inappropriate design of conditionality led to adverse consequences for project and program outcomes. The chapter provides lessons on major issues in the conditionality of development assistance.

Reference

Williams, Alan, and Emilio Giardina. 1993. *Efficiency in the Public Sector: The Theory and Practice of Cost-Benefit Analysis.* Brookfield: Edward Elgar.