SPECIAL STUDY No. 18

New Approaches to Public Decision-Making

by

Alice M. Rivlin

prepared for the
Economic Council of Canada

NEW APPROACHES TO PUBLIC DECISION-MAKING

Prepared for the

Economic Council of Canada

by

Alice M. Rivlin

January 1972

This Study was prepared as a background paper for the Eighth Annual Review of the Economic Council of Canada. Although it is published under the auspices of the Council, the views expressed herein are those of the author.

Available by mail from Information Canada, Ottawa,
and at the following Information Canada bookshops:

HALIFAX
1735 Barrington Street

MONTREAL
1182 St. Catherine Street West

OTTAWA
171 Slater Street

TORONTO
221 Yonge Street

WINNIPEG
393 Portage Avenue

VANCOUVER
657 Granville Street

or through your bookseller

Price: $1.25 Catalogue No. EC22-2/18

Price subject to change without notice

Information Canada
Ottawa, 1972

CONTENTS

FOREWORD

In its *Eighth Annual Review*, the Economic Council of Canada focused on the processes of government decision-making.

Clearly there is growing public concern for improving these processes. This concern reflects many factors, including the rapid growth in the scale and scope of government activities, the spreading recognition that government decisions now have greater consequences for good or for ill than ever before, and the heightened pace of change in our complex modern society.

In fact, many governments have already taken cognizance of this concern, and the last decade has seen new approaches to public decision-making taking shape in many different countries. Few people are better qualified to comment on these new approaches than Mrs. Alice M. Rivlin. Mrs. Rivlin, who is now a senior fellow in economics at The Brookings Institution, played a major role in implementing the planning, programming and budgeting system of the U.S. Department of Health, Education and Welfare. She was asked to set out a discussion of the subject in terms that would be reasonably clear even to those without extensive technical background.

As is the usual practice with a study commissioned by the Council, the contents are the responsibility of the author. Publication under our auspices means that the Council considers the present study a worthwhile contribution to public knowledge and the understanding of economic issues.

Sylvia Ostry, *Director*,
Economic Council of Canada

PART 1

NEW EFFORTS TO BE SYSTEMATIC ABOUT PUBLIC DECISION-MAKING

The last decade has seen new approaches to public decision-making take shape in many different countries. Essentially, they are all efforts to be more systematic and explicit about public decisions, to look at alternatives carefully, and to weigh the consequences of public actions.

The Planning Programming Budgeting Systems at the federal level in Canada and the United States are examples of this phenomenon. Many other countries, states, localities, and international organizations are also undertaking similar efforts under different rubrics. In some places, the new approach has been announced with much fanfare, high-level support, and exaggerated expectations. In others, it has been introduced more quietly or has simply evolved from existing procedures.

The essential elements of the new approach are: (1) making goals and objectives of public policies and programs explicit; (2) weighing the consequences of alternatives carefully and as quantitatively as possible; (3) setting up a systematic process for making decisions.

Why Are These New Approaches Evolving Now?

One might wonder why so much attention in so many countries is being paid to systematic decision-making at this particular juncture in history. The problem of allocating scarce resources among competing ends is not new. Indeed, scarcity is far less evident than in the past. The Western industrial nations are increasingly affluent. One might suppose that their decisions about how to spend this affluence would get easier rather than harder, since it is no longer necessary to forgo essential consumption in order to provide government services.

But affluence itself has increased aspriations and options, and economic growth has brought a whole host of demands for new services that appear to be most appropriately provided by the public sector. The public sector has grown rapidly in all Western countries, with the result that a large portion of the nation's output

has been removed from the discipline of prices and profits, and the incentives that supposedly lead the private sector to produce effectively. The mere size of the public sector now makes it important to scrutinize public decisions and try new approaches to ensure that the public is getting its money's worth.

Moreover, the vastly increased complexity of government and society as a whole makes it necessary to review public decisions carefully. With so many programs, agencies, and levels of government now in the picture, the potentialities for unintended effects of public programs and conflicts among them are very great if no attempt is made to be systematic about decisions.

Finally, the availability of new tools, especially survey techniques for collecting information and computers for processing data, has made it possible to analyse public decisions in greater depth and in a more organized way. Even a decade ago, it was hardly possible to collect, analyse, and store the information needed to evaluate the effect of public programs on people or to compare significant alternatives. Recent rapid advances in techniques, such as simulation, make it possible to explore the consequences of alternative policies and programs more rapidly and in more detail than ever before.

What Systematic Decision-Making Is Supposed To Avoid

The new approaches evolved out of a conviction that there must be a better way to make public decisions than the present one. One has only to observe the way in which budget allocations are made in most governments, even in supposedly "advanced" Western countries, to conclude that it is worth experimenting in the hope of finding better ways.

The process of formulating public policies and programs and allocating resources among them has been characterized by several factors that clearly do not conspire to enhance the effectiveness of the results.

1. *Ignorance of the effects of public programs* -- Until recently, few attempts have been made to trace the effects of public spending or to determine who benefited and by how much. The effects were generally thought to be diffuse and unmeasurable. To follow individuals

affected by government programs long enough to find out what happens to them is difficult and sometimes expensive. As a result, decisions are often made in ignorance of the current effectiveness of programs or the expected effect of new effort.

2. *Incrementalism* -- The easiest way to make up a budget is to start with last year's and add funds if necessary and where appropriate. As a result, this year's budget often looks a great deal like last year's budget. All programs are continued without any serious examination of whether they are still useful or whether a major or minor reallocation of resources among programs would lead to better results.

3. *The squeaky wheel* -- Another characteristic of the budget process has been the tendency to allocate more funds to the bureaucrats or pressure groups who are the most articulate or vociferous in defending their programs, even if the effectiveness of these programs is dubious. When subjected to serious scrutiny and evaluated in relation to other alternatives, the squeaky wheel may well be found to be less deserving of oil than some of the quiet ones.

4. *Cross purposes* -- Where programs are not subjected to systematic scrutiny and weighed against each other carefully, programs may develop in various parts of the government that duplicate or even nullify each other and thus lead to substantial waste of public resources.

5. *Locking-in* -- Small budgetary decisions in one year may lead to large future costs that eat up resources which would otherwise be available for alternative uses, even though those who made the decision did not know that this consequence would follow. Multi-year planning can reduce this hazard.

It is these undesirable features of government decision processes that led to a search for better tools -- tools that will make it possible to plan ahead, to look at objectives and the merits of the various ways of achieving them.

Making Goals Explicit

Governments, like individuals, often have diffuse objectives. They are hard to specify and to measure, and it is sometimes difficult to know whether they have been attained. Increasing education, improving health, facilitating transportation, maintaining peace and order, etc., are such vague goals that it is not always clear whether the nation as a whole is moving towards or away from them. Specific government programs are often enacted without any clear notion of how they will contribute to policy objectives. Moreover, it is sometimes impossible to determine whether such programs are successful because no operationally meaningful measure of success has been stated.

The premise of the new approaches to decision-making is that public decisions will be improved if special effort is made to state policy objectives and to be explicit about what specific government programs are to accomplish. The premise is simply common sense -- it is easier to make a decision when one knows what one wants to do. Moreover, it is easier to evaluate an action when one has a measure of what constitutes success.

Proponents of explicit goal-setting for public policies and programs argue that identification of specific objectives increases the chance that the decision will actually reflect what the participants in a decision process want to do. Being precise is an aid to rational discussion and debate. For example, one might put many programs under the general rubric of "improving justice" and find it very difficult to choose among them. When a deliberate effort is made to define what "improving justice" means -- increasing rates of apprehension of criminals, or speeding up trials, or reducing recidivism -- the chances increase that the resulting choice of program will lead to these explicit objectives.

Not everyone agrees with the proposition that the decision process can be improved by making goals explicit or by specifying objectives. It is sometimes argued that this process may paralyse action. Different segments of the population might agree on a vague general goal when they would disagree on more specific formulations of it. Most people might agree that it would be useful to improve justice but yet find it impossible to reach agreement on what aspects of justice should be altered.

Alternatively, different groups might favour the same action for different reasons. A program to subsidize food purchases for low-income people might be favoured by both farmers and poor people for different reasons. These two groups might be able to agree on the desirability of the program and yet disagree on whether the objective was to support farm prices or assist the poor. A classic example in recent U.S. history is the Elementary and Secondary Education Act of 1965 which gave federal money for the education of low-income children, including some children in nonpublic schools. The measure was supported by those who favoured improving the lot of the poor, by those who thought the program was a first step towards more general federal aid for education, and by those who thought it established the principle of aid to private and religious schools. This coalition was sufficient to ensure passage of the Act. An explicit attempt to specify the goals of the legislation might have wrecked the coalition and perhaps doomed the legislation.[1]

Another objection to explicit goal-setting is that attempts to specify goals too precisely may lead to overemphasizing measurable objectives to the detriment of nonmeasurable but perhaps equally important objectives. Specificity about objectives might lead a nation to concentrate on economic growth and neglect the less tangible element, the quality of life. It might lead to undue concentration on reducing infant mortality, which is a normally collected statistic, and neglect of health services that could increase the vigour, energy, and activity of older people in a less measurable way.

Weighing the Alternatives

The new approaches force explicit recognition of the fact that public decisions involve choices among alternative uses of scarce resources. The resources used for education cannot also be used for defence or health or road-building. Except where idle men or machines can be put to work, more public goods mean less private ones, and within the public sector, more for one kind of program means less for another.

---

[1]Charles L. Schultze, *The Politics and Economics of Public Spending* (Washington: The Brookings Institution, 1968), p. 65.

Moreover, choices must be made between now and later -- between current consumption and investment that will increase the resources available for consumption in the future; between treating disease now or engaging in research to find a way of eliminating the disease altogether.

Choices must also be made between different groups in the population. Public expenditures on things like higher education or airport safety, may primarily benefit the upper-income groups, while health clinics or food price subsidies may primarily benefit the poor.

There is nothing new about these choices or even the general public's awareness of them. What is new is a serious sustained attempt to spell them out explicitly and to give them quantitative dimensions wherever possible.

Choices among major objectives are ultimately made by a complex political process. There is no objective way to compare the value to a nation of faster transportation systems with the value of cleaner air, or the benefits of better health with the benefits of reduced crime. But the new approaches to decision-making embody the hope that the choices made by the political process will be improved if everyone knows what is being traded -- how much speed is being sacrificed for how much reduction in pollution; how great an improvement in health can be bought with the resources required to reduce crime to a lower level.

When the choices being made are among different ways of reaching the same objective, the case for weighing alternatives carefully and quantitatively seems even clearer. Sometimes the choices are conceptually simple, such as the choice between more and less expensive ways of building a road or carrying out a manpower training program, where the results are expected to be substantially the same. Often, however, the choice is more complicated. One way of cleaning up a river may be cheaper than another, but it may take longer. One way of delivering medical care may be more effective for most people, but markedly less so for a few. The specification and quantification of these alternatives do not make the choices easier, but they do make clearer to all the participants in the decision process what the dimensions of the choices are.

When the decisions involve choices among different groups in the population, the resolution must again be primarily political. But here too, knowledge helps. The distributional effects of many public programs are not obvious. Efforts to trace through the probable distributional consequences of alternative courses of action at least allow decision-makers to choose the alternative most likely to have the distributional consequences they want to achieve.

An example from the U.S. social security system illustrates the last point. For many years, those who felt most strongly about bettering the lot of the poor favoured raising the minimum social security benefit in the belief that most of the recipients of the minimum benefit were poor. When information was collected on the characteristics of beneficiaries, it was discovered that, in fact, a large proportion of the recipients of the minimum were not poor. They were secondary earners and others whose attachment to the labour force had been short-lived precisely because they had other sources of income.

Although the notion that public decisions involve choices seems trite in the abstract, experience with the new approaches reveals that many public officials are reluctant to face this seemingly obvious fact explicitly. For example, during early attempts to implement the Planning Programming Budgeting System in the United States, some of the medical personnel in charge of disease control programs argued vociferously against estimating the number of lives that would be saved per dollar expended in different disease control programs on the grounds that all human lives should be saved and that no choices are possible. They did not want to admit that none of these programs were operating at the level necessary to save all possible lives and that any budget allocation among programs implicitly involved saving some lives and not others. These medical officials had been making these decisions for years but were extremely reluctant to face the human consequences that were actually involved.

Periodic Review Process

If the new approaches to decision-making simply involved assembling more and more information on choices, decision-makers might drown in statistics without any significant improvement in their understanding of the

problems or in the decisions they make. But major emphasis has been placed on putting information into useful digestible form and especially on establishing regular periodic review procedures in which information relevant to decisions is assembled so that major options can be intelligently reviewed. Much of the discussion and controversy about the Planning Programming Budgeting System has *not* been over the basic concept, but over the usefulness of particular tools designed to bring together and display information for this periodic review.

As PPB evolved in the U.S. federal government in the late 1960's, at least three such tools were developed -- the program and financial plan (PFP), the program memorandum (PM), and the special analysis. Specifications differed somewhat over time and were variously interpreted by individual agencies, but the general notions are not hard to describe.

Each government department or agency was required to develop a PFP showing how it proposed to use its resources over a five-year period (since the total future resources available could not be known, several alternative PFP's were usually developed based on different assumptions about the total level of resources). The plan was to be reviewed and revised each year, and the revised plan reflected in the annual budget.

Several aspects of the PFP are noteworthy. First, the fact that it was a five-year plan made it necessary to spell out future costs of current decisions. It was not possible to add a small item for "planning" or start-up costs of a new program without detailing the future implications and facing the fact that such a decision would leave fewer resources for other uses in future years. The process of formulating the plan and looking ahead to future choices forced many public officials to think more clearly and to articulate, perhaps for the first time, about how their programs were expected to develop in the future. New directions and changes of emphasis were highlighted by the five-year plan and were actively debated. These changes of emphasis might not have received attention if discussion had been confined to a one-year budget in which changes were relatively small and constrained by previous commitments. Secondly, the PFP was organized by program objectives, not by the specific administrative unit

that carried out the programs. This made it possible, at least within a major department, to look at all the resources devoted to a particular objective, such as improving the health of the aged, reducing poverty, or defending against a ground attack in Europe. This mode of presentation was generally more illuminating to decision-makers than a line item budget organized by an administrative agency, which might show the number of tanks to be used by the Army without indicating what they might be used for.

Finally, the PFP was supposed to show not only financial data, such as planned expenditures by program, but also real output measures. The intention was to develop output or success measures for each program, which could be used prospectively for planning and retrospectively to evaluate whether, in fact, the program met its objectives. The hope was to produce final output indicators, such as measures of the population's health, education level, and income, or of water pollution. So far, in most programs, it has been necessary to settle for intermediate or proxy indicators, such as numbers of doctors trained, or classrooms built, or waste treatment plants constructed.

While the PFP summarized the agency's preferred five-year plan, the program memorandum explained why the decisions were made. Addressed to the President or the Bureau of the Budget, the program memorandum was supposed to focus on major changes and big shifts in the allocation of resources, to discuss alternatives considered and rejected, and to explain the basis for choosing the proposed course of action. The program memorandum was supposed to give the President or the Bureau of the Budget enough insight into the considerations that went into a particular decision to tell whether they agreed or disagreed with the proposal. In practice, the principal usefulness of program memoranda, however, was inside the agencies that prepared them. Drafting program memoranda clarified issues, made it necessary to expose the bases for decisions, and probably led to more soundly based proposals.

Special analyses were supposed to contain supporting details, special studies, and additional data. They were essentially appendixes to the program memoranda, designed for those who had a special reason to go more deeply into the details of a particular decision.

Problems of Timing

As it developed in the U.S. federal government, the Planning Programming Budgeting System was tied to the budget cycle, since the budget submitted by the President to Congress in January each year is the principal instrument by which the executive branch affects the allocation of federal resources among programs. PPB set up a yearlong decision cycle culminating in the budget decisions themselves. The cycle was supposed to begin soon after the last budget was sent to Congress with the identification of major issues likely to arise in the preparation of the next budget. Analytical work was to be organized around these issues. Options were to be laid out and the consequences of alternative courses of action were to be explored. In the light of these studies, the five-year plan was to be reviewed and altered, after which the implications of the new plan would be translated into specific budget proposals. Program memoranda and special studies would then be put in final shape to explain the basis for the decision in the proposed budget. The whole package -- budget, PFP, program memoranda, and special studies -- was supposed to arrive at the Bureau of the Budget in plenty of time for discussion and interaction before final budget decisions were made.

That was the theory. In practice, there were many problems. The data that would have been useful for making many decisions were simply unavailable. The PFP turned out to be an unwieldy document, and program memoranda tended to be long, abstruse, and difficult for decision-makers to read. Political considerations, of course, entered many allocation decisions. Moreover, the timing of the decision cycle never worked out as planned. Analytical work on important issues took more than a few months to complete and often was not completed by the time the decision had to be made. Moreover, discussion of the budget often dragged on in Congress for most of the year, so that the next year's budget had to be formulated before final decisions were made on the budget for the current year.

Despite these shortcomings and difficulties, however, PPB introduced new dimensions into the decision process. More searching questions were asked and answered; more analysis was done of alternatives; more justification was felt necessary for political decisions

that flew in the face of analysis; and far more explanation was offered of the basis for what was done. It was clear that the forms could be better, but also that it was useful to have goals made explicit in formulating a long-range plan, to analyse the relevant alternatives as carefully as possible, and to have a regular process for reviewing these alternatives before formulating the budget.

PART 2

SOME SPECIFIC TOOLS OF SYSTEMATIC DECISION-MAKING

Benefit-Cost Analysis

The development of benefit-cost analysis was a response to a problem -- the problem of choosing among alternative public investments. Governments, like private businesses, invest resources in the hope of future benefits. They have choices of alternative investments and it is clearly sensible to look carefully at these alternatives to compare their costs and their likely payoff and to choose the most profitable alternative.

A business considering new investment presumably looks at a range of possible options and tries to pick the most profitable. In the real world, of course, businesses do not "maximize profit" in the strict textbook sense. Some corporations avoid risks even if the expected profit is high. Others have short planning horizons and seek quick returns, perhaps neglecting more profitable long-run options. Some may operate under constraints, such as keeping present management or retaining senior employees, or staying in one part of the country. But subject to these kinds of constraints and within the limits of ability and imagination, good business management involves examining options and picking those expected to be most profitable over the relevant time period.

Governments also make investments. They use present resources to increase the future well-being of the population. Some investments are in physical facilities. Roads, dams, bridges, pollution control devices, all bring benefits over a period of time. Others are human investments -- education, health care, nutrition -- that increase the productivity and well-being of people over their lifetimes. Benefit-cost analysis is an attempt to look systematically at these public investment alternatives and to identify the most profitable uses of public resources.

Historically, benefit-cost analysis was applied first to physical investments. Much of the early work was done in connection with water resource projects.

Damming a river is a major investment that may be expected to yield benefits over a long period. There are large numbers of alternative dam sites and different kinds of water resource projects that could be undertaken. It seems reasonable that decisions about which, if any, of several projects should be undertaken should be influenced by comparisons of the costs and expected benefits of the different alternatives.

The problem of where to put a dam would be relatively easy to solve if such a project had only one purpose or expected benefit, such as the production of electric power. But the problem is normally complicated by the existence of several simultaneous objectives, such as power production, flood control, and recreation. There are also different kinds of costs involved. The cost of constructing the dam and acquiring the land may be fairly easy to estimate, but there may also be hidden costs involved in the displacement of people and industry -- costs that may not be reflected in the prices that have to be paid. Since the dam will last a long time, it is necessary to decide how to value future benefits and costs relative to more proximate ones and how to take account of risks and uncertainty about future developments.

These valuation problems are difficult to solve, but over the last two decades, many attempts were made -- with increasing sophistication -- to value the benefits and costs of water resource projects in dollar terms so that decision-makers could have estimates of whether a particular project would be worthwhile and how alternative projects would compare in terms of total profitability.

More recently, the basic idea of benefit-cost analysis has been transferred from physical investments, such as multipurpose river developments, to human investments, such as manpower training. Investment in a training program affects the future income of the trainees. If the project is successful, the trainees will be more skilful and productive. This increased productivity will be reflected in higher earnings per hour or per week. At higher skill levels the trainees may also find steadier employment. Besides these monetary benefits, there may be less tangible ones, such as greater job satisfaction and self esteem or lower crime rates. But there are also costs. The costs may be nonrecurrent, as in the case of initial training, or they may be continuous, as in the

case of day care for trainees' children. Without day care, the trainees would be unable to work and their training would be wasted. Benefit-cost analysis could be used to elucidate the decision about whether a particular training project is worthwhile or which of several alternative projects is the best investment.

The current debate about work and welfare in the United States has stimulated a number of estimates of the costs and benefits involved. Much of the debate focuses on the problem of whether training and day care should be provided for mothers on welfare so that they can find employment. While one might favour this course of action for a number of reasons, it is not clear that it is justifiable in benefit-cost terms to substitute work for welfare. If the mother on welfare has little education or job experience, her earning capacity is low. Giving her sufficient training to find a better job and providing day care for her children while she works are expensive public activities. If the mother has several children, estimates usually show that the benefits (increased earnings) that can be expected from enabling her to work are less than the costs (training, work-related expenses, and day care). This does not mean, of course, that encouraging work is in any sense "irrational", but it has to be defended on grounds other than profitable investment as measured by a benefit-cost ratio.

Some attempts have also been made in the United States to estimate the benefits and costs of compensatory education for poor children. The estimates are shaky, partly because the benefits of such programs lie so far in the future. Nevertheless, such evidence as gathered so far provides little support for the hypothesis that the benefits of compensatory education are high relative to the costs. These findings (or the absence of findings) have played a part in shifting the U.S. federal government away from an education strategy against poverty and towards an income-maintenance strategy. Federal expenditures for compensatory education rose rapidly in the mid-1960's but have since leveled off, while expenditures for income-support programs have continued to rise.

Another human investment question is how to spend health resources most effectively. The case for looking at health care as an investment is perhaps most obvious

in underdeveloped countries. Especially in poverty-stricken rural nations, poor health caused by malaria and other debilitating diseases greatly undermines the productivity of the labour force. Public funds spent on inoculation, sanitation, nutrition, or pest control, could be expected to reduce the incidence of disease. These measures could lead to lower costs of treating disease in the future and to increased productivity of the labour force. Benefit-cost analysis could be useful in determining whether particular disease control programs are worthwhile and how these programs compare with other possible investments in future productivity.

The potential for applying benefit-cost analysis to health problems, however, does not end with less developed countries. One can also look at choices among various disease control programs in developed countries as a benefit-cost problem. Diseases have different incidence. Some are widespread; some are rare. They attack at different ages. Some kill; some cripple or put people out of action for long or short periods. With limited public resources to be spent on disease control, it seems sensible to make an effort to see that these resources are being channeled into the highest payoff use. In general, the choice is not between controlling disease A and letting disease B run rampant. Rather, the decision is at the margin. It involves the question of whether additional special measures to control disease A may have greater or less payoff per dollar than additional measures to control disease B.

It may also be necessary to decide between spending additional resources on controlling the disease in question at the present level of medical knowledge or investing resources in research to find a cure for the disease. To consider the research option introduces a new element of uncertainty. One cannot be sure that the research will be successful or, if so, how much the success will cost and how long it will take. It may only be possible to estimate probabilities of success at various levels of resources and to use these to see how sensitive the benefit-cost relationship is to the probability of success if medical research is undertaken.

No matter what set of investments is considered, several common problems have to be resolved before benefit-cost analysis can be carried out.

1. *Whose benefits?* -- In a business decision, it is generally clear whose profits are to be counted. Government decisions, however, are different. It is not usually appropriate to focus on the benefits to the government itself, although in some cases this may make sense (for example, increasing the number of tax collectors may result in enough additional tax collections to cover the costs). In general, it is appropriate to compare the benefits to the community as a whole with the costs. The benefits may be reflected in increased productivity of the economy, or increased satisfaction of individuals, or reduction in costs that would otherwise have occurred (such as costs of crime or floods prevented).

A relatively small governmental jurisdiction -- a city or a province -- would tend to count as benefits only those that accrue to its own citizens or within its own borders and would not give much weight to benefits flowing outside its boundary. For example, it would not consider as future benefits the increased incomes of migrant labourers or others likely to leave the area. It would be unlikely to give high priority to cleaning up a river whose downstream area lies mostly in other jurisdictions, or to support scientific research, the benefits of which would mostly go to non-residents. If the benefit-cost analysis were done by a larger unit, such as the whole country, these spillover benefits would be counted, and programs for migrant labour or water pollution or scientific research would appear to have a higher value relative to their costs. The presence of such spillovers could, of course, provide some economic justification for, say, federal grants to, or cost-sharing arrangements with, other levels of government, or possibly for the transfer of a program to the government with wider jurisdiction.

2. *Whose costs?* -- Similar problems arise in making estimates of the costs of public investment. In principle, the cost estimate to be set against the benefits ought to be an estimate of the value that the resources to be used in this particular public investment would have in alternative uses.

Outlays necessary to carry out the project or program may be a sufficient measure of costs in this sense, since the price of labour and materials used reflects their value for other uses. But there may also be

hidden costs. A power plant, for example, may cause pollution, which is costly, but this does not show up in the conventional cost estimates for building and running the plant. There are, of course, spillover costs as well as spillover benefits, and an analysis done for a larger jursidiction would count more of these than one done for a smaller jurisdiction.

3. *What benefits should be counted?* -- Even within the same jurisdiction, there are many different kinds of benefits from particular public investments. It is not always clear what to count or how to measure.

Some benefits are private; they can be appropriated by individuals. Public funds spent on higher education, for example, increase individuals' incomes or add to their personal resources. Clearly, these private benefits should be counted -- although there may be justification in having the beneficiaries pay part or all of the cost. But there are also public or social benefits that cannot be appropriated by individuals. Some individuals are better off because of the public investment in *other* people or the community at large. National defence is a classic example of a pure public good from which everyone benefits; no one can be excluded. The country must be defended as a whole or not at all. Many other government investments, however, have mixed public and private benefits. An individual benefits both from his own education, for example, and from living in a community where others are educated.

Generally, it is harder to value the social benefits in a benefit-cost analysis simply because they do not have market prices. It is impossible to get an idea of what people would be willing to pay for them.

4. *How to value the costs and benefits* -- In order to compare the profitability of different investments, public or private, it is necessary to measure all the benefits and costs in the same units, usually money. But it may be difficult to derive monetary values for some of the components, especially on the benefit side. Benefits that show up in increased income of individuals are automatically measured in dollars. Those which increase the quantity of some item normally bought and sold in the market place can also be measured in dollars, since the market price reflects consumer evaluation of this item or at least approximates it in a reasonably

competitive economy. But items not normally sold are far more difficult to value -- for example, the value of a national park or clean air or a crime-free city. In some cases, it may be possible to construct plausible shadow prices that give an indication of what the market price of the item would have been if it had been normally traded. For example, differences in property values between otherwise similar high-crime and low-crime areas may give some indication of the value that is attached to this item. The price that consumers are willing to pay for access to a recreational facility may give some hint as to the value of a similar facility in another place. Nevertheless, for many intangible items -- such as freedom from fear of a disease -- there may be no plausible way to find a shadow price. These intangible benefits may have to be given entirely arbitrary weights in the benefit-cost analysis or be left out altogether.

5. *How to compare the present and the future* -- A particular benefit is worth more now that one of the same size in a year or two or three from now; but how much more? Different individuals clearly have different rates of discounting future benefits. To some extent, these rates may be related to the particular rate at which an individual can borrow or earn interest on money lent (an individual who can easily borrow at 6 per cent is presumably indifferent between an offer of $100 now and that of $106 a year from now).

In valuing the costs and benefits of a public investment, the higher the discount rate used, the greater the weight given to present, as opposed to future, benefits. But what is the appropriate discount rate for a government?

There has been considerable argument over this question, and different rates have been used. Some benefit-cost analyses have used the government borrowing rate. Others (more correctly) have used the rate of return prevailing in the private sector, since this is a measure of what the public resources could be producing if left in private hands. But the important thing in comparing different public investment projects is to use the *same* discount rate, whatever it is. If different government departments, for example, used different rates, greater weight would be given to the future in some areas than others, and comparisons across departments would not be useful.

6. *What to do about risk and uncertainty* -- The problems of risk and uncertainty plague any investment decision -- not just public ones. Both costs and benefits may be drastically altered by changes in economic conditions, behaviour patterns of individuals, or forces beyond the control of those making the decision. For example, benefit-cost ratios for a manpower training project may be estimated from the past earnings histories of persons with similar training. It must be recognized, however, that a rise in the unemployment rate or drop in demand for the skill in question could cut the earnings of prospective trainees and make the actual benefit-cost ratio far lower than anticipated. Benefit-cost analysis will be most useful if, in addition to estimates of the most probable benefit-cost ratio, estimates are made of alternative outcomes that might arise if circumstances alter. Then the user can at least see how sensitive the results of the analysis are to changes in the uncontrollable factors.

Advantages and Limitations of Benefit-Cost Analysis

The purpose of benefit-cost analysis is to bring out into the open the economic implications of spending marginal resources in different ways so that decision-makers can have some idea whether the excess of benefits over cost is greater for an additional water resource project or an additional training project, or whether more higher education has a higher payoff than more disease control.

Even the most ardent enthusiasts of benefit-cost analysis do not advocate that decisions be made on benefit-cost grounds alone or that projects or programs with the highest benefit-cost ratios automatically be chosen. Rather, the case for benefit-cost analysis rests on the importance of having before the decision-makers information on the measurable benefits and costs of alternatives so that they at least know what they are doing. If they choose a lower benefit-cost ratio, they should do it consciously and know what they are giving up in the process.

In order to make these comparisons possible across different types of projects and programs, it is clearly necessary to have a common unit of measurement. This is why it is so important to translate all the costs

and benefits into dollar terms. It may well be difficult to do this, but the enthusiasts of benefit-cost analysis argue that a guess is better than no estimate at all and that, with time and effort, more sophisticated methods of valuing even intangible benefits in dollar terms will be developed.

Benefit-cost analysis is most clearly useful and appropriate when the alternative investments being considered are primarily designed to increase future income rather than produce intangible benefits, such as freedom from fear of crime or disease. Although proxies and shadow prices can be developed for these less tangible (or less marketable) benefits, these shadow prices may not correspond very closely to true values. Moreover, the effort involved in refining them may not add much to the intuition of those involved in the political decision process.

The critics of benefit-cost analysis argue that its limitation makes it actually dangerous as a decision tool. The difficulty of valuing intangible benefits means that they will generally be left out of consideration. Decision-makers will be given studies that show high benefit-cost ratios for projects and programs whose benefits are easily measured in dollars -- notably those that contribute in a direct way to increases in future income. The danger is that they will choose those projects and programs over others with less measurable but equally important benefits. The argument here is that if only part of the benefits can be measured, it is better not to measure any of them than to become overenamoured of those that are easily quantifiable.

The more serious limitation of benefit-cost analysis, however, is that it does not help decision-makers with distributional and broad value questions. Suppose that benefit-cost analysis indicates that the payoff from additional higher education is greater than from control of tuberculosis or that mass transit projects have higher benefit-cost ratios than rural roads. It is still true that the beneficiaries of higher education are not the same people who have tuberculosis and the beneficiaries of mass transit are not those who need rural roads. The most important parts of the decision may relate to the relative priorities given to the needs

of different groups in the population, and these decisions can only be made by the political process.

Moreover, even if one is considering a program whose benefits are indivisible and accrue to the public at large (national defence, for example, or space exploration), it is still true that different people value these benefits differently. Some individuals give great weight to the importance of space exploration and others do not; some would be willing to pay a great deal for more adequate national defence and some believe defence expenditures to be largely pernicious. These differences in values and preferences for public goods can only be resolved through the political process.

Cost-Effectiveness Analysis

Many important decisions about the use of public resources simply relate to the question of how to achieve a given objective for the same people in the most efficient manner. These questions can be analysed without getting into the difficulties that plague benefit-cost analysis -- difficulties of valuing disparate benefits in dollars or dealing with the distribution of benefits among different groups in the population. A useful set of tools for dealing with this simpler subset of problems may be described as "cost-effectiveness analysis". Cost-effectiveness analysis is useful in answering questions about how to achieve a specified set of objectives at the least cost, or how to get the most effect from a given set of resources.

Some kind of cost-effectiveness analysis is involved in any resource using decision, like the construction of a road or a bridge or a school. The physical characteristics are specified, and an attempt is made to find the least expensive way to satisfy these specifications. Often the specifications have to be adjusted to fit into the budget, and the question becomes: What is the best bridge (or road or school) that can be built for the money allotted? Usually, some of the specifications can be traded for others at a given budget level. Higher speed of construction may mean less durability; greater elegance of design may sacrifice certain safety features, and so forth. It is important to know what is being traded and at what cost.

Cost-effectiveness analysis is just an extension of these basic ideas into other, sometimes more complex, public decision problems. It is simply an attempt to specify what is to be done in order to examine systematically the alternative ways of doing it.

Cost-effectiveness analysis has been developed through a high point of sophistication in military decision-making. One classic study compared the effectiveness of airlift and sealift in attaining specified objectives that involved moving men and supplies to different parts of the globe to respond to military contingencies. The extension of cost-effectiveness analysis into the domestic decision arena is relatively new, although it seems logical and not -- at first glance -- impossibly difficult. Three examples will serve to illustrate the potential usefulness of cost-effectiveness analysis -- the choice among (1) alternative techniques and methods in elementary education; (2) alternative mechanisms for health care delivery; and (3) alternative manpower programs for increasing employment.

1. Elementary Education

Any country with universal schooling uses substantial resources to put children through the first six to eight years of school. These resources are usually allocated without much systematic analysis of their effectiveness. Could a better job be done for the same cost or the same job for a lower cost?

The problem seems clearly amenable to cost-effectiveness analysis. It is possible to specify at least some of the desired outputs of elementary education. Children should be able to read, and to do arithmetic; they should have certain information and certain skills in communication and problem-solving. These outcomes can be measured -- albeit imperfectly -- by standardized tests. There are clearly different ways to accomplish these specified objectives. The ratios of teachers to children could be altered. Different kinds of buildings and equipment could be used. The menu of possible teaching methods is extensive and includes minor variations in the ways of teaching a particular skill, as well as major differences in the concept of teaching and learning. The schools could be strict or permissive, competitive or co-operative;

classes could be divided by ability or socio-economic status, or could be deliberately mixed. All these alternatives have different costs and presumably different effectiveness. Systematic comparisons might reveal that major increases in effectiveness are possible without increases in resources or, conversely, that the same job could be done for less money. In view of the volume of resources consumed by the elementary school system, substantial effort to carry out cost-effectiveness analysis would seem warranted.

Differences in the effectiveness of education need not result only from changes in the mix of inputs (changing class size or adopting a new curriculum). Changes in the incentives facing participants in the education process might also affect outputs. The thorough cost-effective analysis, for example, might consider the effect on achievement of altering the rewards facing teachers or students.

2. Health Care Delivery

Another service that absorbs major resources in the most modern countries is health care delivery. Here again, the potential usefulness of cost-effectiveness seems great. It is certainly possible to specify certain desirable characteristics of a health system. Patients with complaints should get diagnosis. Patients with diagnosed conditions should receive treatment. Those who need intensive care should receive it; those who need only bed rest should get it. Emergencies arising from accidents or other causes should receive attention as rapidly as possible (e.g., within x minutes y per cent of the time).[1]

For any specified level of care, there are various ways in which it can be delivered. Options include changing the location and size of hospitals or the availability of outpatient care, nursing homes, convalescent care, or home nursing. They involve the distribution of doctors or ratio of doctors to nurses and other health personnel. The nature of their respective duties and responsibilities may also be changed.

---

[1] We are not talking here about the state of the art of medicine or the effectiveness of care in curing disease, just about the delivery of the care itself.

Moreover, the payment mechanism used by the patient to reimburse suppliers of medical care may alter the effectiveness of care received. If only treatment costs were covered by insurance, for example, the result might be a much higher ratio of treatment to preventive care.

In view of these alternative ways of producing the specified care level, careful analysis seems warranted. The results -- as in education -- might indicate that present delivery systems are unnecessarily costly and that equally effective care could be produced in less expensive ways.

3. Manpower Programs

Here again, substantial resources have been devoted to training without systematic consideration, until recently, of the cost and effectiveness of alternatives. Again, it is possible to specify objectives, such as increasing the employment rate of a specific group to a determined level. Alternative ways to accomplish this objective might be on-the-job training, a variety of school-type training courses, transportation and associated services to encourage the unemployed to move elsewhere, and efforts to move jobs into areas of unemployment. Here again, careful analysis of the cost of alternatives might reveal that the least costly alternative is not now being undertaken.

Some Difficulties

Some of the difficulties of benefit-cost analysis are ruled out of cost-effectiveness problems by definition -- it is not necessary to value the benefits in dollars or to decide on any set of weights for valuing the benefits to different groups in the population. However, many difficulties remain besides those discussed in the next section relating to the availability of data. First, it is difficult to choose satisfactory output measures. Even if the outputs could be specified, it is often hard to measure them. The analyst often has to resort to proxy measures.

A more serious difficulty is that public programs rarely have a single output. In the education example, it may be that more emphasis on reading means less on mathematics, or more information acquired by children means less attention to problem-solving. In health care,

frequency of care may mean less thoroughness or, if one is dealing with more ultimate measures of health success, lower mortality may mean higher morbidity rates. In manpower training programs, higher employment or quicker placement may mean jobs with more limited advancement possibilities.

What is most cost-effective? The typical cost-effectiveness analysis does not yield a single answer. Rather, it attempts to define the rate at which different objectives may be traded; how much of one has to be given up in order to attain the other. Ultimately, decisions depend on the relative importance attached to these different outputs.

Simulation of Program Effects

Considerable advances have been made in recent years in estimating the distributional effects of changes in public programs. Most progress, of course, has been made on the least difficult problems -- estimating the initial effects of changes in tax rates or cash benefit programs on the distribution of income.

Where the population is large and diverse, even quite simple changes in tax rates or cash benefit schedules may have distributional results that are difficult to anticipate. For example, changes in the income tax law may be under consideration. It would be possible to raise exemptions, to change tax rates, or to allow more generous deductions for work-related expenses. Policy-makers need to know how much each of these alterations would cost (or bring in) and how each would affect the relative position of different groups in the population -- lower- and higher-income groups, older and younger families, the urban and rural population, and so forth. The ability of tax experts to answer these questions quickly and reliably has been greatly increased by the capacity of computers to handle large bodies of data.

It is now possible to take a sample of actual tax returns and enter the relevant information on computer tape. The computer can then be programmed to recompute the tax for each individual under a variety of alternative rules. It can then reaggregate the information by age or income or family type to see how different groups would be affected by various tax changes.

Both the Canadian and U.S. governments have developed capacities for simulating tax changes on computers. Indeed, an interesting example of the use of this technique involved simulating the recommendations of the Canadian Royal Commission on Taxation (Carter Commission) to see what distributional effects these recommendations would have if applied in the United States.[1]

Similar methods have been used to estimate the costs and effects of alternative proposals for reform of the welfare system in the United States. The current welfare system involves payments only to certain types of poverty-stricken families (the aged, the blind, the disabled, and mothers with dependent children). Levels of payment are set by the state rather than the federal government, although the federal treasury puts up part of the money. The result of this arrangement is that levels of welfare benefits vary widely among states, and much of the poverty population (families with male heads who work but have low earnings) is not generally eligible for welfare benefits.

The U.S. welfare system has been criticized on a variety of grounds -- inadequacy, unfairness, and the creation of perverse incentives. This criticism created an active debate in the 1960's over the merits of alternative approaches to helping the needy, such as children's allowances or a negative income tax. Much of this discussion was hampered by the difficulties of estimating the costs of alternative plans and the effects they might have on different groups.

By 1968 and early 1969, the policy debate had focused on proposals of the negative income tax type -- namely, proposals that would guarantee all families a minimum income based on family size and allow them to keep some portion of their earnings without deduction from that guaranteed income. This discussion culminated in the Nixon administration's proposal to replace the current welfare program known as Aid to Families with Dependent Children with a new Family Assistance Plan, which is essentially a negative income tax for

---

[1]Joseph A. Pechman and Benjamin A. Okner, "Simulation of the Carter Commission Tax Proposals for the United States", *National Tax Journal*, vol. 22, no. 1, March 1969.

families with children. The debate that has raged in Congress and the public press since the administration's submission of this proposal in 1969 has been concerned not so much with whether the basic concept is right, but with the difficult choices involved in setting the guaranteed level and the rate at which benefits would be reduced as earnings rose.

The budget constraints imposed by the war and inflation have made these choices particularly difficult. One objective of welfare reform was to increase the income of the poor and reduce the disparity in welfare benefits that exists between high-income and low-income states. Raising the guaranteed level would contribute to this objective. Another objective, however, was to maintain (or increase) incentives to work among the low-income population. One way to accomplish this is to increase the allowance for work-related expenses. Another is to reduce the rate at which welfare payments are reduced as earnings (after expenses) rise. An intelligent debate about these proposals required some means of estimating the costs and distributional effects of raising the guarantee to various alternative levels or setting a marginal tax rate (rate at which welfare payments are reduced as earnings rise).

Fortunately, by the time this debate over welfare reform became an important national issue, a new tool for making these estimates had evolved. It was possible to simulate the effect of the changes on a computer. The first simulations used a special census sample designed to give more accurate information on low-income families -- the Survey of Economic Opportunity (SEO). Subsequently, this sample was updated, corrected, improved, and amalgamated with other data sources. The basic idea, however, was to reproduce the sample population on tape and program the computer to compute individual benefits under a variety of payment rolls and then aggregate the information to show the total costs of the alternative programs and how the benefits would be distributed. This conceptually simple device has proved extremely useful. Literally hundreds of simulations of alternative policies were done during the course of the administration's decision-making prior to the submission of the Family Assistance Plan to Congress; many more, during the still-continuing debate in the Congress itself.

The ability to simulate these welfare program alternatives has, of course, not made the choices easier. Indeed, it has dramatized some of the difficulties. For example, the simulations demonstrate that when the guarantee level is related to family size, there will be substantial transfers from small to large families within income classes. The simulations have also dramatized the difficulties posed by interregional income differences. The interregional differences make it difficult to move to a national payment level. If the national level is low, relative to what is considered adequate in the wealthier states, it will be necessary to allow these states to supplement the welfare payments. If they do not, the current beneficiaries of welfare programs in these states will be worse off. A higher national level would reduce the disparities among states but might disrupt the economy of low-wage areas by providing benefits above prevailing wage levels to those not working.

Useful as these tax and welfare simulations are, they are primitive compared with the type of model of the population that would be really desirable if one were to formulate policy intelligently. The simulation procedures used for estimating the effective tax changes assume that the changes will *not* affect the way in which the population divides its time between work and leisure or the family groups in which people choose to live. This assumption may not be extremely unrealistic for small tax changes in the short run. However, major alterations in the structure or rate of taxation could be expected to change incentives for work or education or conceivably living arrangements as well. Similarly, the simulation procedures used to predict the costs and effects of welfare reform assume that changes in the welfare laws will not affect the working behaviour or family groupings of the population. This assumption seems even less realistic. More than trivial changes in welfare benefit levels or extension of benefits to new groups might well be expected to affect the earning behaviour and living arrangements of the poor. Indeed, much of the policy debate about welfare revolves around these specific effects. Extending aid to the working poor is advocated as a means of decreasing the incentives that men with low earnings have to leave their families. Lowering the rate at which welfare benefits are reduced as earnings rise is advocated as a method of encouraging the low-income population to seek employment.

For these reasons, policy-makers need a model of the population that incorporates what is known about the effects of tax or welfare changes on the behaviour of the population. If such a model were available, it would be possible to simulate the effects of a variety of tax and benefit changes and to obtain estimates of the costs and income distribution effects of these changes *after* the probable effects on incentives had worked themselves out. Some progress towards constructing such models has been made, although the information available for estimating the behavioural responses of individuals to changes and incentives is extremely limited.

PART 3

IMPROVING THE DATA FOR PUBLIC DECISION-MAKING

The usefulness of these new tools of systematic decision-making depends entirely on having accurate data for use in policy analysis. The requirement is not necessarily for *more* data, since much of what is collected now is not really useful to decision-makers. Rather, the necessity is to make existing data more useful (for example, by linking separate bodies of data to each other) as well as to develop new kinds of information specifically designed to answer the questions on which intelligent decision-making depends, namely:

-- What is the present state of the nation? How are resources and problems and social characteristics distributed?

-- How effective are present public programs?

-- What would happen if alternative policies were followed?

Three sources of data of special relevance to social program decisions will be discussed here. They are: (1) sample surveys -- still photographs of society at intervals; (2) longitudinal data or panel studies -- moving pictures following individuals through time; (3) social experiments -- systematic attempts to alter the services rendered to individuals or the incentives facing them and to record and analyse their behaviour.

1. Sample Surveys

Collecting information about the characteristics of the population is, of course, not a new activity. Censuses have been carried out for many years and have yielded information on the size of the population and its characteristics, such as age, sex, education, and income. It is possible to obtain from census data periodic pictures of important aspects of the society and its characteristics.

But a census is cumbersome and expensive. Only a limited number of questions can be asked each person. Responses are often inadequate, and processing the volume of data that a census yields is an overwhelming undertaking even with modern computers.

For most decision questions, a smaller sample survey is a far more useful and flexible tool than the complete census. The sample survey yields information more cheaply and quickly and is often more accurate (because more resources can be spent eliciting and validating answers from each respondent) than a complete census. Indeed, refinements in the technique of sampling and administering surveys have made a major contribution to better informed decision-making in the last two or three decades. Moderate-size surveys, like the Current Population Survey (CPS) in the United States, yield quick and reasonably reliable estimates of unemployment rates, income changes, labour force participation, school attendance rates, and similar indicators.

Recent attention to the poverty problem in the United States has led to intensive work on surveys designed to yield a more accurate view of the characteristics of the low-income population. Two Surveys of Economic Opportunity, for example, have been undertaken by the U.S. Bureau of the Census in an attempt to obtain more accurate information on the low-income population by expanding the size of the regular CPS sample and adding more detailed questions on sources of income and other subjects. The result of these and other efforts has been a major improvement in knowledge of who is poor, where the poor live, and what their characteristics and problems are. Simultaneously, the National Health Survey is yielding improved information on the health status of the U.S. population, and the National Assessment of Education is contributing to a more accurate picture of the intellectual skills and knowledge of children and young adults.

The accuracy of survey information often leaves a great deal to be desired. For example, the total amount of transfer income (welfare, social security pensions, and the like) reported by individuals on the Survey of Economic Opportunity was considerably less than the amount known to have been paid out by the government. To obtain a more accurate representation of the income

sources of low-income families (for use in simulating the distributional effects of welfare reform), it was necessary to "correct" the responses to allow for the underreporting of income that was known to be occurring.

It is possible to improve the usefulness of survey data by merging different bodies of data to obtain a more complete picture of the population. For example, families or individuals whose characteristics are known from one survey can be assigned additional characteristics based on information collected in another survey. Suppose Survey A collects information on employment and income, and several other variables, but not homeownership, and that Survey B collects information on a different set of characteristics, including homeownership as well as employment and income. If Survey B shows that employed individuals at a given income level have a 75 per cent chance of being homeowners, the two bodies of data could be merged by designating three-quarters of the employed individuals at that income level in Survey A with the status of "homeowner".

Alternatively, it is possible to retrieve data on the same people from two different sources. To match real people, it is necessary to have some kind of identifying number or code. In the United States, the Social Security number has become an almost universal identifying number for employed persons. An interview survey could collect information on health or leisure activities of employees and record their Social Security numbers. The numbers could be used to match the survey records with tax or Social Security records to obtain information on income or work experience of the same individual.

Matching individual records on the basis of an identifying number has potential dangers as well as uses. Tremendous care has to be exercised to protect the confidentiality of individuals. The spectre of a massive data bank in which a complete dossier on each individual is recorded under his identifying number frightens many people. The Social Security number was not collected on the U.S. Census of 1970 because of public concern for protection of privacy.

Cross-section surveys -- those giving a picture of population at a particular time -- are useful in assessing current needs or problems and in analysing the

distributional impact of changes in taxes or cash benefits. They are not usually useful for evaluating the effects of current government programs or estimating how people would react to changes in public policy. For these purposes, one needs longitudinal or panel data on the same individuals at different times.

2. Longitudinal Data

Evaluation of the effectiveness of a government service requires, at the very least, "before" and "after" data on the people who received the service. One might evaluate a preschool program by testing children at the beginning and the end of the course; a manpower training program, by comparing participants' incomes before training and after they returned to the labour market; a health care program, by comparing health records of individuals before and after the program.

In general, however, "before" and "after" statistics are insufficient for definitive evaluation and may even be seriously misleading. First, the "before" and "after" comparison does not allow for what would have happened to these people without the government service. Little children learn a lot without going to school; the average income of a group of people may rise in response to changes in employment rates and general wage levels even if they receive no additional training. To make these comparisons, one needs a control group of similar individuals who did not receive the service. Genuine control groups are often hard to find. If a government service is generally available, there may be no one who is not receiving it to serve as a valid control. Even if there are only a few participants, they may differ in important ways from nonrecipients. (Volunteers for a program, for example, may be more highly motivated than those who do not volunteer.)

Second, two points in time -- "before" and "after" -- are generally not enough for a definitive evaluation. One is interested in the retention rate of skills taught and in whether and how fast the effect of preschool or manpower training fades out. Hence, one needs to follow participants and controls for some time after the program is completed. Moreover, many social services are presumably cumulative in their effect. The most interesting policy questions in education relate not to whether changing a particular piece of the curriculum makes a

difference by itself, but to whether a sustained effort has an impact and, if so, what kind and under what conditions. Health care has cumulative effects; so do nutrition, housing conditions, air pollution, and so forth. Hence, evaluating public programs usually requires ways of tracking individuals over a fairly long period, recording services received or exposure encountered, and measuring their status (health status or skill level or income) at different times. Sometimes the records that would facilitate this kind of analysis are currently in existence -- in the files of schools or doctors, for example -- but not accessible for analysis. Sometimes they are not kept at all. The U.S. welfare system, for example, keeps records of cases from the time they are opened until the time they are closed. If the same person comes back on welfare later, however, he is a "new case". Hence there is now no usable information that would show how often the same people turn up on welfare, what characteristics the frequent cases have, or how services rendered to them are related to their status in the long run.

Estimating behavioural relationships also requires information on the same people at different points in time. For example, important policy questions turn on estimates of how individuals will alter their spending, saving, or working patterns in the face of changes in their incomes. Cross-sectional data showing the behaviour of people at different income levels at a single point in time are not useful in making these estimates. It cannot be assumed that people who now have $4,000 will, if given $2,000 more income, behave like the present $6,000 income group. Their behaviour may be quite different in the face of income change, and may depend on whether the change is perceived to be permanent or temporary. To find out, it is necessary to observe the behaviour of the same people as their income changes.

Policy analysis clearly would be greatly improved if accurate information were available on panels of individuals at different points in life. This kind of information, however, is more expensive to obtain than cross-section surveys, and poses more risk to individual privacy if the information is misused.

3. Social Experiments

Some important policy questions probably cannot be answered by any amount of improvement and refinement of

data collected about present government programs and their effects. To answer these questions will require actual experiments -- trying out new programs in a systematic way and analysing the results.

For example, attempts of analysts to discover more effective methods of producing education have so far been disappointing. Statistical studies have related educational inputs (teacher qualifications, student-teacher ratios, buildings, curricula, etc.) to results (test scores and other measures of student performance) but have not uncovered any strong relationships that would help policy-makers to improve the effectiveness of education. Better data over longer periods might bring these relationships to the surface -- assuming they exist. However, it may be that the natural variation in educational inputs simply is not great enough, or that the effects of this variation are indistinguishable from the effects of nonschool factors (parental influence, and the like) that affect children's performance. If this is true, it may be necessary to undertake a series of systematic experiments with radical (and less radical) departures from existing practice. It may be necessary to formulate new models and try them out in enough different places to obtain definite indications of what works best under what conditions and with what kind of children.

Another whole range of policy questions that requires experimentation is the effect of incentives on individual behaviour. For example, recent interest in the United States in negative income taxes or guaranteed income as a substitute for the present welfare system has raised some major questions: What would be the effect of a guaranteed income on the extent to which people work? The answer to this question is crucial in estimating the cost of a guaranteed income program and its impact on the economy and society at large. There seems to be no way of finding out about such probable incentive effects by examining existing transfer programs, since these are generally limited by law to persons defined as unable to work due to age, disability, or other factors. For this reason, it has been necessary to take an experimental approach. Several experiments are currently under way to test the effect of various types of negative income tax or guaranteed income on labour force participation, earnings, family stability, and other behaviour.

Another major opportunity for experimentation relates to alternative systems of reimbursing providers of health care. It is widely agreed that the present system of reimbursement in the United States is inimical to efficient use of medical resources. Hospitals, for example, are generally reimbursed by insurance companies for actual costs of treating patients and have no incentives to improve their efficiency. Alternative reimbursement schemes, such as capitation payments, are available, but their effectiveness cannot be easily judged either *a priori* or by looking at the questionably comparable experience of other countries. Some experiments are now under way, and others are needed to assess the effects of alternative medical reimbursement systems on the quality and quantity of care received.

Although an experimental approach holds promise of greatly increased knowledge of the effectiveness of public programs, it is fraught with dangers and difficulties. There are serious ethical problems involved in experiments with human beings. Lives can be damaged if an experiment fails, and inevitably some must. Alternatively, some people may suffer if they are excluded from a successful experiment in order to form a control group against which to measure the success. There are also political problems -- risks that experimental data will be manipulated to show results favourable to persons in power or that unfavourable results will be suppressed. Moreover, there are serious technical problems in designing and evaluating experiments with human participants. Questions must be raised about the validity of experiments in which the individuals involved know they are being watched or that the experimental conditions are temporary.

Despite all these difficulties, it seems likely that the responsible use of the social experimentation technique will add to knowledge and to the ability to use public resources effectively. It should be remembered that there are also risks in not experimenting. Lives are being damaged by the status quo.

The same general kinds of objections can be raised to all the techniques discussed in this paper. Systems for improving public decision-making, benefit-cost analysis, cost-effectiveness analysis, and simulation, all entail risks of misuse. There are always danger of overemphasizing the measurable, of neglecting intangible

benefits, of misspecifying inputs and outputs, and of using measuring instruments that are invalid. Nevertheless, the alternatives -- decision-making by guesswork and intuition -- also entail serious dangers. On balance, the case for responsible use of systematic techniques to improve public decision-making seems strong.

New Approaches to Public Decision-Making

by Alice M. Rivlin

Economic Council of Canada