

---

## SUMMARY

---

Agency and program administrators and decisionmakers responsible for implementing early childhood intervention programs are becoming more interested in quantifying the costs and benefits of such programs. Part of the reason for this is that foundations and other funders are putting more emphasis on results-based accountability. At the same time, arguments for the value of early childhood intervention are being made within the public sphere on the basis of published estimates of costs and benefits. Program implementers are naturally attracted by statements that a certain intervention produces \$4 in savings for every \$1 it costs and would like to make similar statements about their own programs. Meanwhile, decisionmakers without particular interest in any given program would like more quantitative decision aids when it comes time to choose among a variety of possible program models or program improvements to implement.

Our objective here is to offer assistance to decisionmakers and program implementers considering an assessment of costs and outcomes. We do not offer a specific step-by-step manual, but we discuss the kinds of issues that must be taken into account and why. We do so in enough detail that readers can decide if this type of quantitative analysis is the right course for them and, if so, can knowledgeably interact with an expert cost-outcome analyst. While we understand that some readers will want to undertake analysis of costs and outcomes to justify a program in which they have a special interest, we take the viewpoint here of an unbiased allocator of funds. What evidence should such a person want to see before concluding that a particular intervention is a wise investment? That sort

of evidence is what the implementer seeking to justify further funding will need to present.

We begin by setting the conceptual framework within which program costs and outcomes may be understood. We then draw out some of the implications of that general framework for the analysis of early childhood interventions in particular. After reviewing some examples of such analyses, we apply the methodology to an actual case in which a consortium of program funders must decide whether to proceed with an assessment and, if so, what kind of assessment to undertake. The consortium is led by the U.S. Substance Abuse and Mental Health Services Administration and the Casey Family Programs, and the intervention of interest is the *Starting Early Starting Smart* program.

## THE COST AND OUTCOME ANALYSIS FRAMEWORK

Decisionmakers and program implementers just beginning to think about analyzing costs and benefits are often surprised to learn that several analytic avenues are open to them. Which one or ones they choose will have important implications for what they learn and how much they must spend to learn it. Among the choices are these:<sup>1</sup>

- **Cost-benefit analysis** (or benefit-cost analysis) entails comparing a program's benefits to a stakeholder with its costs to that stakeholder. Such a comparison requires putting benefits and costs in comparable terms, and the terms conventionally chosen are dollars. Benefits that cannot be expressed in dollar terms cannot be compared in this manner and are included only in associated qualitative discussion. Cost-benefit analysis seeks to help in deciding whether a program is of value to the stakeholder. Often cost-benefit analysis is conducted from the perspective of society at large.<sup>2</sup>

---

<sup>1</sup>Terminology in this field has not been standardized, and these terms appear in the literature with a variety of different meanings. We have chosen typical definitions.

<sup>2</sup>Of the four analytic approaches listed here, cost-benefit analysis is subject to the greatest challenges in execution and interpretation. That is because benefits must be denominated in dollars, and that adds another source of uncertainty and potential

- **Cost-savings analysis** is restricted to the costs and benefits realized by the government as a whole or a particular funding agency. Only the costs to the government are taken into account, and the benefits are those expressible as dollar savings somewhere in the government. This kind of analysis is used to determine whether a publicly provided program “pays for itself” and is thus justified not only by whatever human services it may render but also on financial terms alone.
- **Cost-effectiveness analysis** determines how much must be spent on a program to produce a particular outcome (or, what is equivalent, how much of a particular type of benefit will result from a given expenditure). While this can be done for multiple outcomes, no attempt is made to sum the complete array of benefits into a single aggregate measure.
- **Cost analysis** alone (no measurement of benefits) can be useful to decisionmakers for a variety of purposes, for example, discovering which factors need to be considered in replicating a program elsewhere or for informing budget projections.

In deciding which avenues to pursue, the decisionmaker or implementer must choose what he or she wishes to learn and consider the funds available for undertaking the analysis. The analyses above are ordered in terms of how much attention must be paid to quantifying outcomes and expressing them in dollar terms (from a lot at the top to none at the bottom). Other variables being equal, the resources and calendar time devoted to the analysis will drop with each successive approach down the list.

As we describe them here, these cost and outcome analysis methods are used only as components within a broader decision support framework that we call *policy analysis* or *policy scorecard analysis* (the latter term derives from the use of a tool called the *scorecard*).<sup>3</sup>

---

disagreement over quantities. For some benefits, dollar conversions are not really feasible. Cost-benefit assessments can thus rarely be comprehensive.

<sup>3</sup>The term *policy analysis* was originally adopted by RAND analysts and others to describe an approach for quantitatively analyzing management problems. Today, the term is used even more broadly to characterize a wide range of quantitative and qualitative approaches to addressing policy issues. Hence, we will employ the more focused term *policy scorecard analysis* for the remainder of this summary.

Despite the name, it does not pertain only to high-level public policies but also to decisions made regarding specific strategies and programs. Policy scorecard analysis offers a framework within which to consider multiple benefits, as required in the first two approaches listed above, and multiple costs, as required by all four. Policy scorecard analysis also entails consideration of alternative programs. This is important for benefit and cost analysis. In trying to determine whether the numbers emanating from these analyses support (further) investment in the program, funders will be asking, “Compared with investment in what else?” A benefit-cost ratio of 1.5 to one (\$1.50 of benefits for every dollar of costs) may not be good enough if an alternative with similar objectives has a ratio of two to one. Decisionmakers will thus be considering a range of alternative interventions or at least a choice between funding the program in question and some default course of action (which could be leaving things as they are).

The results of a policy scorecard analysis can be summarized in a simple tool called a *scorecard*. The scorecard lists benefit and cost categories down the side, together with the program design features influencing them, and lists the alternative courses of action across the top. Thus, each cell in the scorecard gives a particular cost or benefit (or design feature) for a particular program. In identifying the row and column heads and filling in the cells—that is, in conducting the policy scorecard analysis—several guidelines must be kept in mind:

- **Designate which benefits and costs accrue to which stakeholders.** If you say that a program generates more savings than costs, people will want to know, savings to whom? And costs to whom?
- **Define explicitly the period over which the analysis applies.** If the purpose of the analysis is to determine whether a program has a favorable benefit-cost ratio or pays for itself in government savings, it is better to look well into the future. No one period or duration is correct, however. The choice depends on the patience of the decisionmaker in question, with individuals typically having shorter planning horizons than society as a whole. This distinction makes a difference because the costs of early intervention programs typically accrue over a matter of months or a few years, whereas the benefits are often not fully realized

until the participating children age into adulthood. Counting such benefits directly entails long-term follow-up of program subjects, though some future benefits can be predicted on the basis of shorter-term trends.

- **Discount future costs and benefits.** Although it is important to count future benefits (and costs), they cannot be counted at full, nominal value. People discount future benefits and costs: getting a \$1,000 benefit five years in the future does not look as attractive as getting it now; having to pay \$1,000 five years in the future does not seem as onerous as having to pay it now. A real annual discount rate of 3 percent to 6 percent is typically applied to future benefits and costs.
- **Record cost elements as resource quantities.** Until the figures are added up at the end, costs should be recorded in terms of resource quantities—hours of labor, square footage of rental space, etc.—rather than in dollar terms. Prices for these resources can vary from one site to another, and on-budget dollars in particular do not always reflect total costs. A physician may donate time on the weekends, but from society’s point of view, that time is not “free”; perhaps it could have been put to another, more beneficial use.
- **Address uncertainty.** Future benefits and costs cannot often be predicted with great confidence. Where a range of values is plausible, that range should be made explicit in the analysis. Likewise, structural uncertainty (e.g., about possible future changes in laws relevant to a program) should also be considered.

The final step in the cost and outcome analysis is to add up all the benefits (or savings) and add up all the costs and compare them across programs. The four analysis methods listed above offer alternative ways of performing this step. Cost-benefit and cost-savings analysis each provide a single measure of merit for each alternative; the alternative with the greatest merit according to this measure is declared the winner. Cost-effectiveness analysis provides multiple measures of merit. They can be combined into a single measure (e.g., the ratio of effectiveness to cost, if a single effectiveness measure dominates), which will be used in the same way as a cost-benefit or cost-savings measure. Or they can be used to define a different

kind of selection rule, one that deems “best” the policy that achieves a specified level of effectiveness at lowest cost (a *constant effectiveness analysis*) or that achieves the greatest effectiveness for a given cost (a *constant cost analysis*).<sup>4</sup>

Comparing costs and benefits may not produce a single “answer” that one program is obviously preferable to another. One program may produce a net benefit to one group of stakeholders, while another benefits a second group. The net benefit of one program may be somewhat higher than that for another, but the uncertainty ranges may overlap so much that the advantage cannot be asserted with confidence. Some change in the institutional environment, e.g., tax reform, could shift benefits and costs enough to change the advantage from one program to another. Such possibilities do not subtract from the value of the cost and outcome analysis. On the contrary, some of the most valuable insights are suggestions for policy changes that reallocate benefits across stakeholder groups so that all of them gain and thus have no incentive to block a program.

In most studies, the majority of the analytical effort will come from learning about the domain, structuring the models of how the intervention works, collecting and cleaning data, etc. In short, filling in the scorecard is challenging. Given that groundwork, computing the summary evaluation metrics is straightforward, whether that metric is a benefit-cost or a cost-effectiveness ratio.

Hence, instead of suggesting that one must choose to implement one of these four analysis approaches, it is more accurate to say that one must choose whether or not to conduct a careful, quantitative summation of the effects of the program. If the answer is yes, then there follows a choice of whether to present the results of that analysis to decisionmakers, as a benefit-cost ratio, cost-effectiveness ratio, cost-savings ratio, cost-only analysis, or some combination thereof.

It is thus important to keep cost-benefit analysis, cost-savings analysis, and other forms of cost and outcome analysis in their place. In

---

<sup>4</sup>The latter is sometimes called a *constant budget analysis*, but this is only appropriate if all the costs appear in the budget of the agency making the decision. In many programs, costs may be distributed across many stakeholders. They will not all appear in any single party’s budget.

any decision, some factors can be resolved only through a decision-maker's values and subjective judgment or through negotiation among stakeholders. Likewise, the public quantifying of decision factors may occasionally be problematic (e.g., when an auto manufacturer compares the cost of a safety improvement with the dollar-equivalent benefit of the lives that could be saved by that design change). Nevertheless, these methods can provide valuable input to choosing among different programs, demonstrating a program's worth, improving programs, and replicating them.

### **APPLYING THE FRAMEWORK TO EARLY CHILDHOOD INTERVENTIONS**

Early intervention programs attempt to improve child health and development by providing young children and their families various social services and supports. Such programs can have effects in four domains: emotional and cognitive development, education, economic well-being (in terms of public assistance, income, and crime), and health. Specific examples of possible benefits within each of these categories are given in Table S.1. Which benefits are measured depends on the purpose of the analysis. Cost-benefit and cost-savings analyses typically seek a comprehensive accounting of the benefits to society or to government (respectively), although many benefits are difficult to express in dollar terms and therefore cannot be aggregated in the cost-benefit assessment. While cost-effectiveness analysis can in principle be performed for any outcome, it is often the case in practice that a single benefit or a narrow set receives most of the attention. A full analysis of the benefits of an early intervention program should include collection of data on as many potential benefits as the analyst's resources permit.

Note that early childhood interventions can benefit parents and other caregivers while simultaneously helping children. It is important to measure benefits to caregivers, because these are often realized over much shorter time periods than are those accruing to children. Ignoring these benefits means underestimating a program's benefit-cost ratio or its potential net savings to government, particularly over the short term—and for some analyses, it will only be feasible to make short-term measurements.

**Table S.1**  
**Early Childhood Intervention Program Benefit Domains**  
**and Illustrative Measures**

Benefit Domain	Illustrative Measures for:	
	Child	Parent/Caregiver
Emotional and cognitive development	Socioemotional and behavior scores IQ test scores Teacher's ratings	Quality of parent-child relationship Quality of home environment
Education	Achievement test scores Grades <i>Grade progression (repetition)</i> <i>Participation in special education</i> <i>Educational attainment</i>	<i>Educational attainment</i>
Public assistance receipt, income, crime	<i>Receipt of public assistance</i> <i>Employment</i> <i>Earnings/income</i> <i>Criminal activity</i> <i>Contact with criminal justice system</i>	<i>Receipt of public assistance</i> <i>Employment</i> <i>Earnings/income</i> <i>Criminal activity</i> <i>Contact with criminal justice system</i>
Health	Physical and mental health status <i>Child abuse and neglect</i> Substance abuse Fertility control <i>Emergency room visits</i> <i>Other health care use</i>	Physical and mental health status Family violence Substance abuse Fertility control

NOTE: *Italics* indicate measures more easily expressed in dollar terms.

Any analysis of benefits of a program under way *must* include a comparison group. This is a group of children and caregivers not enrolled in the program but similar in as many ways as possible to the program participants and whose progress along the various benefit measures is tracked.<sup>5</sup> Children in particular have a tendency

<sup>5</sup>Ideally, one should randomly assign children and caregivers to program participation versus the comparison group. This ensures that the participation and comparison

to improve along various measures of development as they grow. Evaluators must take care to ensure that the program benefits they measure are net of what would have occurred naturally or what children would realize anyway from outside influences without the program. Measurements of the comparison group provide estimates of benefits that would have accrued in the program's absence.

Data on progress along benefit measures can be collected by survey questionnaires, tests, or other means of direct interaction with the children and their caregivers. For some benefit types (e.g., reductions in involvement with the criminal justice system), administrative data may be available. When only a few years of data collection are feasible, a glimpse into the future can be obtained through mathematical models that can predict future criminal activity or future earnings on the basis of childhood information. (This cannot of course be done with confidence for any given child, but results obtained for a group of children may be sufficiently reliable for the purpose.)

As with benefits, the cost elements to be included in an analysis depend on its purpose. For example, costs that accrue to society but not to a funding agency are included in a societal cost-benefit analysis but omitted from a cost-savings analysis. Regardless of the analysis to be performed, program costs must be estimated as net of those accrued by comparison group children for similar services. For example, if an intervention is intended to increase prenatal care, the analysis should include only the resources devoted to the visits and services received by program participants in excess of what they would have received anyway (i.e., in excess of those received by the comparison group).

Estimation of costs should follow the general guideline given above regarding the need to estimate resource quantities instead of dollars and to account for "opportunity" costs and other off-budget resource expenditures. Costs borne by participants should also be included, as well as costs borne by other agencies or service providers. Collect-

---

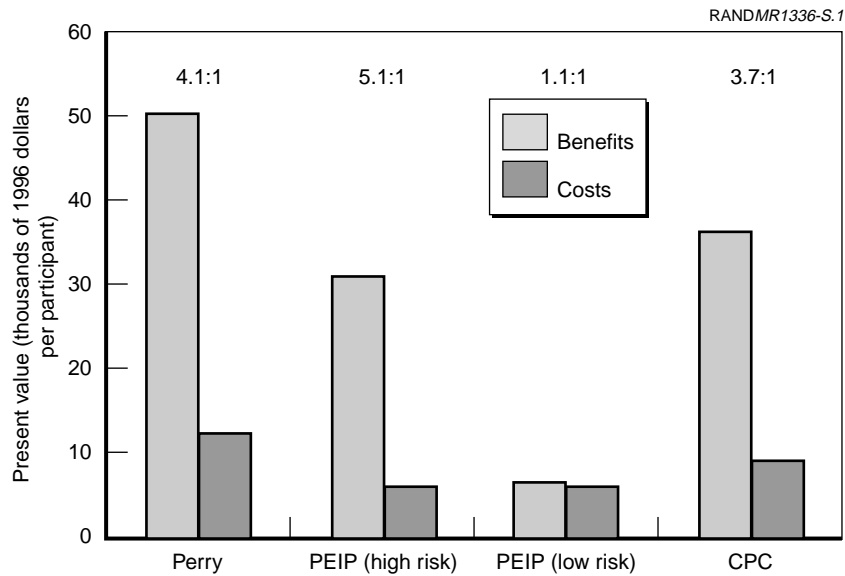
groups are (statistically) identical in both measured and unmeasured characteristics. When the comparison group is selected by random assignment, it is often called a control group. When random assignment is not feasible or desirable, a comparison group can still be chosen, by identifying children and caregivers who are similar in various measured ways to the program participants.

ing cost data for the same set of service providers for both the treatment and control groups allows the analyst to detect both cost shifting (e.g., from one payor to another) and cost offsets (e.g., reduced utilization of services in one area as a result of increased service use in another). In implementing a program, it may also be useful to distinguish between the fixed costs that are not dependent on the number of children served and the variable costs that are. The split between fixed and variable costs will influence the calculation of benefit-cost ratios, net savings, and cost-effectiveness ratios for programs when scaled up to serve larger numbers of children.

### **SOME ILLUSTRATIVE ANALYSES**

Given the challenges and requirements outlined so far, it should not be surprising that not many scientifically sound cost-benefit and cost-savings analyses of early childhood intervention programs with long-term follow-ups have been conducted. Among those recently analyzed or reanalyzed are the following:

- The **Perry Preschool** program provided center-based classes and teacher home visits for one or two school years to 58 children ages three or four in Ypsilanti, Michigan, from 1962 to 1967. Benefits were tracked for both the participants and the comparison group (65 children) through age 27. Benefits included better school performance, higher employment, less welfare dependence, and lower involvement in criminal activity on the part of participants. The most recent cost-benefit assessment evaluates benefits expressible in monetary terms at \$50,000 per child, half of that in the form of savings to government, versus a program cost of \$12,000 per child (see Figure S.1).
- In the **Prenatal/Early Infancy Project (PEIP)** in Elmira, New York, nurses started visiting mothers when they were pregnant and continued until their child was age two. The objective was to improve pregnancy outcomes and parenting skills and link the mother with social services. Between 1978 and 1980 the program reached 116 first-time mothers. They and another 184 in the control group have been followed through age 15 of the first-born child. Benefits for the mothers included better pregnancy behaviors and less child abuse in the short term and lower wel-



**Figure S.1—Some Early Childhood Interventions Have Been Shown to Have High Benefit-Cost Ratios**

fare participation and criminal behavior in the long term. The children benefited as well in several domains. For the higher-risk portion of the sample (unmarried mothers with low socioeconomic status), benefits amounted to almost \$31,000 per mother-child pair, with almost half of that in the form of a reduction in welfare received by the mother. For the lower-risk portion of the sample, however, benefits came to only \$6,700. Program costs were about \$6,100.

- The **Chicago Child-Parent Centers** have promoted reading and language skills, provided health and social services, and promoted parent involvement for children in preschool through third grade. A cohort of 989 children completing kindergarten in 1986 was tracked to age 20 and compared with a no-preschool group of 550 children. The program resulted in long-lasting educational-achievement benefits. Higher between-grade promotion rates, reduced special-education use, increased earnings expected as a result of better educational performance, and low-

er involvement with the juvenile justice system translated into about \$35,000 in benefits per program participant. The program cost nearly \$10,000 per participant.

These analyses demonstrate that early childhood interventions can generate savings to government and benefits to society that exceed program costs. Indeed, for most of the samples reported above, benefits were a multiple of costs, and all of these programs resulted in benefits that could not be translated into costs and were thus omitted. Therefore, decisionmakers and implementers thinking about performing analyses of costs and benefits should not give up merely because they don't see how some of a program's principal benefits can be converted to dollar terms.

Two further lessons for cost-benefit analysis may be drawn from these examples. First, many important benefits can only be captured through an extended time horizon. The savings from Perry Preschool, for example, did not accumulate to match the level of program costs until the participants were 20 years old. Some of these benefits can be predicted on the basis of shorter trends, but not all can, and confidence in predicted results increases as follow-up periods lengthen.

Second, programs can be beneficial to caregivers as well as to children. In fact, when time is lacking for lengthy follow-ups or when they are not feasible, measuring benefits to caregivers can result in early favorable benefit-cost ratios and net savings. The Elmira program was the only one of those summarized that measured caregiver benefits, and, in that case, savings sufficient to balance costs were tallied within two years of the end of program services.

## **FRAMING A POLICY SCORECARD ANALYSIS FOR A SPECIFIC PROGRAM**

The *Starting Early Starting Smart* (SESS) program is intended to test the effectiveness of integrating mental health services and substance abuse prevention and treatment into early childhood education or primary health care for children from birth to age seven. The program is under way at 12 sites nationwide, seven using the early childhood (EC) education model and five using the primary care (PC) paradigm. (See the appendix for a description of each state.) Most of

the sites serve between 100 and 300 children, and comparison groups average out to similar numbers.

By “effectiveness,” the program means increased access to, use of, and satisfaction with behavioral health services and increased social, emotional, and cognitive functioning on the part of served children. Data on these benefit measures are being collected over an 18-month follow-up period at intervals that average six months (PC sites) or nine months (EC sites). No cost data are being gathered in this first phase of the program, but a second phase is being planned, and part of that planning is to assess the feasibility of cost and outcome analysis.

SESS program implementers are wise to take cost and benefit evaluation issues into account in the planning stage. Too often, evaluation is considered only after program design has been finalized along lines that preclude sound cost and benefit assessment. SESS’s Phase I design raises issues that need to be resolved for Phase II if cost and outcome analyses are to be possible. One issue, for example, is that some sites did not use random assignment (primarily EC sites), which raises concerns about the validity of the treatment group versus comparison group difference as a measure of the true effects of the program. Future demonstration sites should aim for random assignment if at all possible. Another concern is that a few sites are experiencing relatively high dropout rates, which could bias benefit estimates if those who are lost to follow-up are different from those who remain in the study and if they differ in important ways that cannot be observed. Obtaining a consistently high follow-up rate across sites would need to be a priority in Phase II. Also, Phase I has been characterized by between-site variations in services. This is problematic from an evaluation standpoint for a couple of reasons: It complicates interpretation of results, and it complicates the design of comparison groups.

The design of comparison groups for SESS offers lessons for other programs. Because SESS attempts to integrate behavioral health services into existing early childhood and primary care settings, only the benefits of the new, integrated services plus increases in the “dosages” of existing services may be credited to SESS, not the full benefits realized from participation in the early childhood program and primary care. Similarly, only the costs associated with these

incremental activities should be considered. Therefore, the comparison groups must be designed to isolate the SESS effects by including everything except SESS. The appropriate comparison groups for this evaluation would consist of children involved in early childhood and primary care programs without the integrated SESS services, not children receiving no services at all.

In the policy analysis scorecard, then, the columns would correspond to the early childhood program without SESS, primary care program without SESS, and then the integrated EC plus SESS and PC plus SESS interventions, along with whatever variants are retained. The rows would be the program descriptors and cost and benefit categories. The program features reported would be those having implications for costs or benefits, e.g., population served, eligibility criteria, age of children at enrollment, qualifications of program personnel, types and “dosages” of services rendered, transportation provisions, and so on. In future demonstrations, this information can be collected through site visits and other mechanisms currently being used in the evaluation of Phase I.

Cost estimates would begin with the cost of serving one child (or child’s caregiver) in terms of labor hours expended with the child and in preparing for the session and in terms of materials consumed. These would then be multiplied by dosage per child and number of children served. Fixed costs unrelated to number of children served, such as space rental, would then be identified. Multiplication by unit costs to convert to dollars would be done last. Ultimately, the cost information should be as comprehensive as possible and comparable across demonstration sites.

Benefit measures now being collected for SESS include information on child problem behavior and social skills, child cognitive development, parent-child interaction, caregiver stress and negative or positive behaviors, caregiver mental health problems, caregiver education and employment, and home environment. As discussed above, the emphasis on both child and caregiver benefits will be important to making the short-run benefit tally as complete as possible. Almost all of these measures, however, are within the domain of emotional and cognitive development and are not easily expressed in dollar terms. This makes a formal cost-benefit or cost-savings analysis problematic in that only a limited set of outcomes might possibly be

valued in dollar terms to be compared with program costs. Unless the program impact for those outcomes valued in dollar terms is very large and favorable, so that sizable dollar benefits are generated, a cost-benefit analysis would be unlikely to show a favorable outcome for the SESS program based on the information available after two years.

While not the program's main intent, other benefits could result from it. Some of these benefits, in such areas as physical health, labor market outcomes, and involvement with the criminal justice system, could be more easily expressed in dollar terms than those now being measured. These outcomes could be collected for parents or caregivers in the short term, and with longer-term follow-up, for the participating children. If behavioral changes are large in these areas as a result of the SESS intervention, they can produce sizable dollar benefits that, even when discounted, will be a large offset to the costs of the program. This is especially relevant for changes in parental behavior that can be measured even in the short run. Improvements of adult economic and health outcomes have been demonstrated to produce substantial short-run benefits in other early childhood programs.

Costs and outcomes would be measured for both the participant and comparison groups, with the difference between the two constituting the incremental cost and benefits from implementing SESS. To compare the present values of all costs and benefits, it will be important to predict how they will accrue over time. Costs and benefits should also be categorized according to which groups incur them. It will be of interest, for example, to know how much the intervention costs and benefits participants, the agency implementing the program, other agencies, and society as a whole.

Taking all these steps would be sufficient to support as full a cost-benefit or cost-savings analysis as is likely to be feasible given the current state of the art. If SESS decisionmakers wish to be able to say something about the value the program returns to society relative to its costs, the preceding array of evaluation tasks and program design modifications would be required. If they decide it is enough to be able to say how much the program saves the government relative to what it costs, then some elements—costs to participants or losses to

crime victims, for example—can be omitted. The overall level of effort required, however, is not likely to change very much.

If SESS funders or implementers would like instead to focus on one or a few prominent measures of effectiveness to compare the different SESS variants with each other, a cost-effectiveness analysis should be sufficient. By collecting cost data, along with data on that one or those few benefits, it would be possible to say, for example, how much child problem behavior decreased (relative to no SESS) per thousand dollars spent on SESS plus EC or SESS plus PC. No conversion of the benefit to dollar terms would be necessary.

Finally, if the purpose was to find out how much program modifications or proliferation of sites would cost, no benefit data would be necessary at all. Clearly, program decisionmakers may have to make trade-offs between what they might like to achieve and how much of a resource commitment they are willing or able to make.

## CONCLUSIONS

The recommendations we offer specific to the SESS program may be framed as a set of more-general guidelines for decisionmakers considering cost and outcome analysis of an early childhood intervention program. In particular, among the recommendations that can be applied more broadly are the following:

- Regarding the design of a program evaluation and cost and outcome analysis:
  - Specify the explicit goals of the cost and outcome analysis to guide the scope of cost and benefit data collection and analysis.
  - Identify comparison groups and track the same cost and outcome measures for both comparison and participant groups. If possible, use random assignment to define comparison groups to provide a more valid test of intervention program effects.
  - To minimize attrition in a longitudinal study, devote resources to retaining study subjects.

- Collect information on program features through site visits and other mechanisms to accurately characterize features of the intervention models as they are implemented and to ensure fidelity to the program model.
- Regarding the collection and analysis of cost data:
  - Collect cost information for both treatment and comparison groups at each site where the intervention program is implemented.
  - Ensure that the cost information is as comprehensive as possible: Costs borne by various parties should be differentiated, the period in which costs are incurred should be identified, and direct and indirect costs, fixed and variable costs, and goods and services provided in-kind should be measured.
  - Plan for proper training and technical support of implementation sites and any cross-site data collection organizations to ensure uniformity in the collection of cost data. Collect information on the cost of data collection, training and support, and the related analyses of the data.
- Regarding the collection and analysis of outcome data:
  - If cost-benefit or cost-savings analysis is the goal, include in the outcome data information for parents and other caregivers in the short term and long term and for children in the long term in those domains with outcomes that can be readily evaluated in terms of dollars and can produce large dollar benefits. The choice of specific outcome measures should be guided by findings from related evaluation studies whenever possible.
  - Obtain information from participants that facilitates collection of administrative data and allows effective tracking of individuals to increase response rates at later follow-ups.
  - Where possible, collect complete histories using retrospective survey questions or administrative data for outcomes that may generate a continuous flow of dollar benefits (e.g.,

labor market outcomes, social welfare program use, use of costly health or education services).

- When supported by other empirical evidence, project future benefits based on observed outcomes. Consider additional method development that would permit such forecasts for a broader range of outcomes.

While we believe these principles are quite general, ultimately these recommendations should be viewed as guidelines that may need to be tailored to the specific circumstances of a given intervention program and its evaluation design. In the end, the objectives of a program's decisionmakers will dictate the shape of the analysis.

The general policy scorecard analysis tools considered in this report, and those specific to cost and outcome analysis, have great promise for improving decisionmaking with respect to such investment programs as the early childhood interventions represented by SESS and its counterparts. When used with skill and judgment, the application of these methods to other programs, such as SESS, will further broaden our base of knowledge regarding the value of these investments and aid decisionmakers in their choice among program alternatives.