

**Promoting Readiness of Minors in
Supplemental Security Income
(PROMISE):**

**Recommendations of the
Technical Advisory Panel
Regarding the Use of Incentive
Payments and the Evaluation
Design**

Final Report

February 28, 2012

Thomas Fraker
Todd Honeycutt



MATHEMATICA
Policy Research

Contract Number:
SS00-06-60110

Mathematica Reference Number:
06970.300

Submitted to:
Social Security Administration
Office of Retirement and Disability Policy
Office of Program Development and
Research
496 Operations
Pole 3-C-25
Baltimore, MD 21235

Project Officer: Jeffrey Hemmeter

Submitted by:
Mathematica Policy Research
1100 1st Street, NE
12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763

Project Director: Thomas Fraker

**Promoting Readiness of Minors
in Supplemental Security Income
(PROMISE):**

**Recommendations of the
Technical Advisory Panel
Regarding the Use of Incentive
Payments and the Evaluation
Design**

Final Report

February 28, 2012

Thomas Fraker
Todd Honeycutt

MATHEMATICA
Policy Research

CONTENTS

I.	OVERVIEW OF PROMISE AND THE ROLE OF THE TECHNICAL ADVISORY PANEL	1
	A. Overview of the PROMISE Project	1
	B. The PROMISE Technical Advisory Panel	3
	C. The Purpose of this Report	4
II.	INTERACTIONS BETWEEN SSI INCENTIVE PAYMENTS AND THE PILOT PROJECTS.....	7
	A. How Should Incentives Be Structured to Serve as a National Model?	8
	B. What Are the Interrelationships Between the Evaluation and the Incentive Payments?	11
	C. How Should Incentive Payments to Awardees Be Structured for Intermediate Outcomes?	13
III.	DEVELOPING AND MEASURING INTERVENTIONS AND PARTNERSHIPS	17
	A. What Are the Objectives and Components of a Process Study of PROMISE?.....	18
	B. How Could an MIS Be Incorporated in PROMISE?.....	19
	C. What Consideration Should Be Given in the Process Study to the Partnership Aspect of PROMISE?.....	21
	D. What Elements of the Process Study Should Be Incorporated in the RFA for PROMISE Projects?	22
IV.	EVALUATION DESIGN	23
	A. What Are the Key Research Questions for the PROMISE Evaluation?	24
	B. Should an Experimental or Nonexperimental Design Be Used?	28
	C. What Data Collection Challenges Is the Evaluation Likely to Face?	30

V.	EXTERNAL VALIDITY.....	33
	A. Who Should Participate in PROMISE?.....	34
	B. What Features of the State and Local Environments Will Help Projects Achieve Their Specific Objectives and Reduce SSI Dependency?	36
	C. What Information Should the Evaluation Provide to SSA and Its Partner Agencies?	37
VI.	OTHER CONSIDERATIONS FOR PROMISE	39
	A. The Family Is a Prominent Aspect of PROMISE, but the Role of the Family in the Intervention Has Not Been Fully Developed and Is Potentially Problematic.....	39
	B. TAP Recommendations About Aspects of the PROMISE Interventions	40
	C. The Effect of the Age-18 Redetermination Process Should Be Fully Considered for PROMISE.....	41
	D. Incorporating Incentive Payments by ED in PROMISE	42

APPENDICES BOUND UNDER SEPARATE COVER

APPENDIX A: TAP MEMBER COMMENTS ON PROMISE EVALUATION CRITERIA	A.1
APPENDIX B: PROMISE TECHNICAL ADVISORY PANEL MEETING NOTES	B.1
APPENDIX C: PROMISE TAP MEMBER POSTMEETING COMMENTS	C.1

I. OVERVIEW OF PROMISE AND THE ROLE OF THE TECHNICAL ADVISORY PANEL

A. Overview of the PROMISE Project

PROMISE—Promoting Readiness of Minors in Supplemental Security Income (SSI)—is a joint initiative of the Social Security Administration (SSA), the U.S. Department of Education (ED), the U.S. Department of Health and Human Services (DHHS), and the U.S. Department of Labor (DOL) to fund pilot projects in states to promote positive changes in the outcomes of youth with disabilities who receive SSI and their families. Broadly, the purpose of PROMISE is to (1) fund pilot projects that have the potential for transformational impacts through making better use of existing resources, improving coordination among programs, and achieving better outcomes for SSI youth and their families; (2) provide evaluation results to inform policy decisions; and (3) generate knowledge regarding program delivery and outcomes. ED will be responsible for project implementation (such as awarding the grants to pilot sites) and will have an anticipated budget of \$30 million. SSA will be responsible for project evaluation and incorporating incentive payments based on reduced dependency on SSI; its anticipated budget is \$10 million. It should be noted that these agencies have not yet allocated these funds to PROMISE. DHHS and DOL will provide input on the project design and services to project participants.

Three distinguishing features of PROMISE are family involvement, multiple agency involvement, and the use of incentive payments to promote outcomes. First, the project is unique in that it will allow awardees to provide services to family members, rather than SSI recipients only. Given the limitations of household resources for this population, PROMISE could benefit youth by promoting the employment and training of their parents. This feature of PROMISE will require the involvement of programs that have not typically served youth with disabilities. Second, the project will involve multiple agencies at multiple levels of government. At the federal level, four agencies are participating in the design and evaluation of the project, bringing their expertise in specific areas to inform the project and identify ways that states can leverage existing resources. At the state and local levels, awardees (which we assume will be state agencies) will collaborate with other agencies and private providers on delivering services to participating youth and their families. Third, the project will provide awardees with incentive payments based on actual savings to SSA from reduced benefit payments to beneficiaries. Additional features of PROMISE include a planning stage, during which input into the project's design is solicited from stakeholders; the incorporation of agency-level waivers to provide flexible access to services for pilot participants; and a rigorous evaluation design to guide implementation, gather evidence, and validate incentive payments. These features of PROMISE will ensure that the project will address many of the barriers in serving youth with disabilities, such as a fragmented service environment, inefficient handoffs to adult services, and inadequate funding of services.

PROMISE pilot projects will be awarded to state agencies with the expectation that they will deliver more-intense and better-coordinated services than would otherwise be the case. SSI youth and their families who participate in these projects will be able to access services funded by each of the four federal partners and by multiple state and local agencies and organizations. Ideally, each project will promote improved identification of eligibility for and delivery of services to participants, such as ensuring that SSI youth who receive high school special education services are connected in a timely manner with vocational services. More specifically, youth and families could be connected to ED's state vocational rehabilitation (VR) agencies and DOL's One-Stop Career Centers for employment services, as well as to more specialized programs through State Councils on

Developmental Disabilities and local child care and housing supports (both funded through DHHS). The precise services that the pilot projects will offer are currently unknown, as are the structures of the operating organizations and their partner agencies. SSA's Youth Transition Demonstration (YTD) projects are seen as potential models for the delivery of services to youth.¹ With PROMISE intended to be managed at the state level and to operate at the local level, communication and management will be primary concerns in implementing the pilots.

Another potential component of PROMISE is the waiver of certain federal agency rules, which will provide awardees with flexibility in delivering services. The waivers might reduce the funding requirements of the projects and spur innovation. During the PROMISE planning stage, the consortium of federal agencies will identify administrative, regulatory, and legislative barriers to serving youth with disabilities and their families. These agencies currently have some authority to minimize such barriers, though additional waiver authority might be needed to substantially enhance the projects. For example, the agencies could seek the authority to pool a portion of the funds that they provide to the states under PROMISE to improve the targeting and coordination of services for youth with disabilities and their families. However, the agencies are sensitive to the need for any waivers or administrative changes to minimize risks for youth with disabilities as well as for taxpayers.

The application and award process for pilot projects is planned for late in fiscal year (FY) 2013 or early in FY 2014. ED will release a request for applications (RFA) with a two-stage selection process. First, an expert panel will review and score the applications. Second, officials will conduct face-to-face interviews with the project teams proposed by the applicants with the highest stage-one scores. It is anticipated that ED will enter into cooperative agreements with the top-scoring stage-two applicants. This funding mechanism would provide ED with more flexibility than traditional grants to renegotiate the terms of the awards to enhance the prospects for project success.

The PROMISE design features currently under consideration include the possibility that awardees would receive initial funding from ED for designing, coordinating, and delivering services, and also be eligible for incentive payments from SSA for SSI-related outcomes achieved by participating youth and possibly their families. The long-term objective for the pilot projects will be reductions in SSI benefits, for which the awardees would receive incentive payments, but SSA may decide to also make incentive payments based on intermediate outcomes. These would be outcomes—such as completion of education and training programs and employment in a paid job—that would be expected to lead to reduced SSI benefits in the longer term. ED may also designate some of its \$30 million for this purpose.

PROMISE incentive payments are viewed both as a means of promoting outcomes and as a source of ongoing funding for projects after the start-up funds have been exhausted. Ideally, the savings to the federal government from reductions in SSI benefits for participants would be large enough to generate incentive payments that could fund a project's services while also providing

¹ Martinez, John, Michelle S. Manno, Peter Baird, Thomas Fraker, Todd Honeycutt, Arif Mamun, Bonnie O'Day, and Anu Rangarajan. "The Social Security Administration's Youth Transition Demonstration Projects: Profiles of the Random Assignment Projects." Document No. PR08-56. Washington, DC: Mathematica Policy Research, December 2008. Available at <http://www.ssa.gov/disabilityresearch/docs/YTD%20Profiles%20Report%20final12-11-2008.pdf>. Accessed December 27, 2011.

some savings to SSA. This approach to incentive payments draws on the social impact bond model (described in Chapter II) and is an innovative way to encourage states to restructure their service environments to achieve outcomes that benefit SSI youth (through increased independence), state agencies (through additional federal funding), and SSA (through savings in SSI benefits net of the cost of the incentives).

B. The PROMISE Technical Advisory Panel

The PROMISE technical advisory panel (TAP) was constituted to provide expert input into the designs of the project's incentive payments and its evaluation. SSA, the federal entity responsible for the evaluation component of the project, funded the TAP and oversaw its formation and functioning. SSA charged the TAP with several tasks: to develop criteria for the evaluation of PROMISE pilot projects; to explore methods for constructing reliable treatment and comparison groups; to define the evaluation's outcome measures for tracking participants' progress toward independence; to define the preliminary research questions, concepts to be evaluated, and evaluation designs to assess the impacts of pilot projects; and to define the mechanism for setting appropriate incentive payments reflecting SSI savings. Aspects of the intervention design, such as the services to be delivered and the agencies and programs to be involved, were not under the TAP's purview. SSA awarded a contract in September 2011 to Mathematica Policy Research to develop and manage the TAP and to prepare this report on its recommendations.

Seven TAP members of varied backgrounds were selected to provide expert opinions and recommendations on PROMISE issues. Per SSA's specifications, the TAP included experts from academia, nonprofit agencies, and the federal government. The TAP members held relevant graduate degrees and had multiple years of experience in a range of fields relevant to PROMISE: SSA and disability policy, research design and evaluation, statistics, research ethics, consumer concerns, rehabilitation and employment services, state education systems, and social impact bonds. The TAP comprised the following individuals (their organizational affiliation and areas of expertise are given in parentheses):

- Burt Barnow, Ph.D. (Trachtenberg School of Public Policy and Public Administration, George Washington University; public policy, program design and evaluation, and employment services)
- Hugh Berry (Department of Education, Office of Special Education and Rehabilitative Services, National Institute on Disability and Rehabilitation Research; youth transition, disability policy, and rehabilitation services)
- Mark Donovan (vice chairman of the Marriott Foundation for People with Disabilities; employment services for people with disabilities)
- David Johnson, Ph.D. (University of Minnesota; evaluation design, disability policy, and state education systems)
- Jamie Kendall (U.S. Department of Health and Human Services, Administration of Children and Families, Administration on Developmental Disabilities; disability and youth transition)
- Jeffrey B. Liebman, Ph.D. (John. F. Kennedy School of Government, Harvard University; social impact bonds and SSA policy)

- Pamela J. Loprest, Ph.D. (Income and Benefits Policy Center, the Urban Institute; disability and youth transition)

The TAP members provided their input orally during a one-day, in-person meeting and in writing both before and after the meeting. The meeting was held in Washington, DC, on December 6, 2011, and included all seven TAP members. Representatives of the four funding agencies and Mathematica also attended the meeting. To prepare for the meeting, the TAP members reviewed a summary of PROMISE and provided written responses to a series of questions about designs for the evaluation and the incentive payments. They also provided written comments after the meeting on issues that they felt were most important for the project. The input by the TAP members in these three formats—premeeting comments, the in-person meeting, and postmeeting comments—constitutes the basis for the findings in this report.

C. The Purpose of this Report

The purpose of this report is to provide SSA and its consortium partners with independent guidance on the designs for PROMISE incentives and for the PROMISE evaluation. Proper designs for both will bolster the success of the PROMISE pilot projects and strengthen the lessons that can be learned from them. The findings in this report reflect the input of the TAP members and the experience of the Mathematica project team in evaluating large-scale demonstration projects, particularly those involving youth with disabilities.

We took an analytical approach of reviewing the TAP member comments to identify common themes, which we present in this report as TAP recommendations for PROMISE. The recommendations reflect areas in which the TAP members were largely in agreement. Because the TAP consisted of seven individuals with strong and well-reasoned views on the design for PROMISE, there were many areas in which they were not in agreement. In addition, there were other areas in which not all members provided input, though we recognize that the lack of input did not necessarily represent a lack of either opinion or agreement with other members. We devote the majority of each chapter to presenting the broadly supported recommendations and the underlying reasoning. We also present other ideas about PROMISE for which there was no general agreement among the TAP members or that reflect the input of only one or two members. We supplement the TAP recommendations with analyses based on our experience with evaluation design and SSA demonstration projects. In some instances, we present our views as recommendations for PROMISE in the context of TAP member comments. We identify these as our recommendations rather than those of the TAP.

This report is structured as follows. Chapter II presents information about possible approaches for SSA to incorporate incentive payments in the PROMISE project. Chapter III identifies issues specific to the measurement and assessment of the services delivered by awardees and the partnerships that are expected to arise. Chapter IV explores evaluation design issues for the project, including the key research questions that the PROMISE evaluation should answer. Chapter V reviews concerns surrounding the external validity of findings from the evaluation. The final chapter addresses four issues that were beyond the TAP's mandate yet were important enough that many members chose to provide their input: the role of the family in PROMISE, the nature of the interventions, the age-18 redetermination issue for youth SSI recipients, and individually based incentive payments that could be funded by ED. Each chapter begins with an introduction to the issue area and a summary of the recommendations by the TAP, followed by any Mathematica

conclusions or recommendations. Both of these are presented in a brief, bulleted format. The chapter then provides a more detailed discussion of the specific issues, recommendations, and conclusions flowing from the TAP's input.

This report is accompanied by a separate document consisting of three appendices that present the comments of the TAP members. These comments are not attributed to specific individuals. At SSA's request, we provide them to assist others in thinking about PROMISE design issues and the recommendations presented in this report. Appendix A contains the premeeting written comments of the TAP members, organized by the questions that we initially posed to them to stimulate their thinking about PROMISE. Appendix B contains edited notes of the comments that were made orally during the meeting itself by the TAP members and other attendees. These are organized such that none of the comments are attributed to specific individuals. Appendix C contains the postmeeting written comments of the TAP members, organized by (unnamed) individual.

II. INTERACTIONS BETWEEN SSI INCENTIVE PAYMENTS AND THE PILOT PROJECTS

Incentive payments based on actual SSI savings could ideally fund state PROMISE projects in the long term. However, during the limited period of performance for the PROMISE pilot projects (anticipated to be five years), SSA may decide to make incentive payments based on outcomes that could be achieved within a shorter time frame and that are known precursors to reductions in SSI benefits. To be effective incentives, these payments must be designed to have desired effects on the behaviors of the awardees, service providers, and SSI youth and their families. Poorly designed incentives could encourage unintended or undesirable outcomes. If SSA is to make payments on the basis of actual or anticipated savings from reductions in SSI benefits, then the agency must establish objective criteria for those payments and the PROMISE projects must be designed to meet them. Factors that should be considered in designing the incentive payments include the entities (individuals and/or organizations) to which the incentives will be offered, specification of criteria and associated outcome measures, consideration of outcomes for counterfactual groups of youth and families, and administrative mechanisms for making payments.

Though the design for PROMISE incentive payments is still under development, one approach being considered would have SSA and ED each providing incentives with distinct structures and objectives. Under this potential design, SSA would provide incentives to awardees based on actual or anticipated reductions in SSI benefits at the aggregate level and ED would provide incentives to awardees, service providers, and/or youth and families based on more-immediate individual-level outcomes that are documented precursors to benefit reductions. These agencies could allocate portions of their funding for PROMISE to cover the costs of the incentives. This chapter presents the TAP's advice regarding incentive payments by SSA. Its advice regarding incentive payments by ED is included in Chapter VI.

The TAP made the following recommendations regarding SSA incentive payments based on SSI benefit reductions:

- In the long term, incentive payments could be an ongoing funding stream for services provided by PROMISE or a successor program.
- Incentive payments should incorporate a long-term perspective and be based on group outcomes: the outcomes achieved by all youth (and possibly their families) receiving the intervention.
- To facilitate administration and minimize associated costs, the TAP recommended that SSI-related incentive payments be based on administrative data that are already being collected (for example, SSA administrative records).
- As part of the PROMISE evaluation, the TAP members recommended using a comparison group to determine whether the intervention had an effect on benefits beyond what would have occurred in the absence of the intervention.
- Incentive payments based strictly on observed benefit reductions might not be realistic in the context of a limited-duration demonstration. For the purposes of the demonstration only, an alternative approach would be for SSA to provide incentive payments based on intermediate outcomes achieved by participants.

- Intermediate outcomes that may be the basis for incentive payments should be ones that have been found to be precursors to SSI benefit reductions, which are the ultimate objective.
- The TAP members supported the designation of a common core set of intermediate outcomes across PROMISE awardees, but they also recommended allowing leeway for additional outcomes specific to each awardee that would be based on the awardee’s specific interventions and logic models.
- To the extent feasible, SSA’s mechanism for incentive payments based on intermediate outcomes should mirror the mechanism the agency would use for incentive payments based on SSI benefit reductions.

We make the following additional recommendations based on issues raised by the TAP members:

- The TAP members disagreed on who should receive the incentive payments for reductions in SSI benefits received by PROMISE participants—awardees, service providers, or the participants themselves. We recommend that these be provided to the awardees due to administrative simplicity and the potential for payments in this form to fund overall project operations.
- The appropriate time frame for observing reductions in SSI benefits due to increased earnings of PROMISE participants, and issuing the associated incentive payments, will depend on the population targeted by and the service model of each project.
- Although this issue was not mentioned by TAP members, incentive payments based on reductions in SSI benefits could also reflect any concurrent changes in Social Security Disability Insurance (SSDI) benefits.
- Determining incentive payments based on the changes in benefits for PROMISE participants between the preintervention period and a follow-up period would probably result in excessive payments. We present two possible models for using the evaluation and its associated comparison group to provide SSA with guidance on how to structure incentive payments in a national program, both of which would provide SSA with information to develop a national incentive payment policy that would save costs.

This chapter explores three aspects of designing and implementing PROMISE incentive payments: (1) developing a structure for the incentive payments that could serve as a model for incentives under a hypothetical national program with many features of PROMISE; (2) in the context of the demonstration, interrelationships between the evaluation and the incentive payments; and (3) consideration of intermediate outcomes (other than SSI benefit reductions) that might be the basis for incentive payments.

A. How Should Incentives Be Structured to Serve as a National Model?

The critical ultimate objective for PROMISE—the reduction of SSI benefits among youth and their families—could be the basis for incentive payments that would fund the continuation of project services after start-up funds are exhausted. This approach has the potential to satisfy two competing demands: (1) improving services and outcomes for youth receiving disability benefits and (2) reducing federal spending, in that the reduction in federal SSI benefits could pay for the services

that generate those reductions while also providing some net savings to the government. Social impact bonds, the cornerstone of the Obama administration's Pay for Success initiative, are a variant of this approach.² With social impact bonds, private providers are paid only when they achieve specific, measurable outcomes that result in savings to the government, thus minimizing the financial risk to taxpayers.

While the use of social impact bonds is barely under way in the United States, the United Kingdom is conducting an experiment that tests their capacity to reduce prisoner recidivism. This example might help to illuminate the mechanism of the PROMISE incentive payments.³ The United Kingdom has contracted with a private organization to arrange for services to prisoners released from a specific prison. The organization will receive payments only if the recidivism rate for those prisoners is at least 7.5 percent less than that for prisoners released from a comparison group of prisons. The contract is structured so that a greater difference in the recidivism rates will result in larger payments, up to a specified maximum. The awarding of payments will begin four years after the commencement of project services, to provide time for the delivery of services and the observation of outcomes.

The incentive payments under PROMISE are not likely to be structured strictly as social impact bonds, because the pilot projects will probably be awarded to state agencies rather than private entities and the awardees will receive initial funding for implementation. However, the potential for the awardees to receive large incentive payments based on participant outcomes, as under social impact bonds, might encourage them to propose and implement innovative solutions to a vexing problem—how to promote the independence of youth with disabilities.

SSA's Ticket to Work (TTW) program currently makes incentive payments to service providers based on outcomes achieved by participants in the program; however, the structure for incentive payments in the PROMISE demonstration could be different. The TTW program pays providers for various outcomes obtained by individual participants (for example, a month of earnings above a threshold). Incentive payments in PROMISE could instead be based on group outcomes—the outcomes achieved by all participants in a pilot project—and priority would be given to long-term outcomes. As noted, the primary outcome would be SSI benefit reductions. The long-term perspective would be intended to encourage the continued provision of services to youth into adulthood. If implemented in PROMISE, this structure for incentive payments would have the potential to foster the development and delivery of services that would lead to extended benefit reductions for the target population while avoiding a misplaced emphasis on narrowly defined outcomes that fail to result in the economic independence of SSI youth or savings for SSA.

The TAP members disagreed on who should receive the incentive payments for the reductions in SSI benefits received by PROMISE participants—awardees, service providers, or the participants themselves. We note that under the social impact bond model, incentive payments serve a dual

² Liebman, Jeffrey B. "Social Impact Bonds: A Promising New Financing Model to Accelerate Social Innovation and Improve Government Performance." Washington, DC: Center for American Progress, 2011; Nonprofit Finance Fund. "Pay for Success: Investing in What Works." New York, NY: Nonprofit Finance Fund. Available at http://nffsib.org/sites/default/files/pay_for_success_report_2012.pdf. Accessed February 6, 2012.

³ Liebman, Jeffrey B. "Social Impact Bonds: A Promising New Financing Model to Accelerate Social Innovation and Improve Government Performance." Washington, DC: Center for American Progress, 2011.

purpose: they incentivize the desired outcomes and they fund the services that generate those outcomes. Furthermore, it would be desirable for the PROMISE incentives to be structured so as to encourage all involved in the pilot projects to work in ways that are aligned with the ultimate objective of reducing SSI benefits for youth and families. Given these considerations, we recommend that incentives based on SSI benefit reductions be provided to the awardees due to administrative simplicity and the potential for payments in this form to fund overall project operations. Awardees would be free to allocate the funds to partner agencies and service providers as they see fit (possibly in the form of incentive payments), based on their models of service provision. They could even use some of the funds to offer incentive payments to participating youth and their families. We also note that other possible types of incentive payments in PROMISE, such as those that ED may make (see Section VI.D), could provide opportunities for payments to a wider range of recipients.

SSA's mechanism for making incentive payments to PROMISE awardees could be designed to be relatively simple (for example, providing 50 percent of SSI savings to the awardees) and replicable in a national program. However, some TAP members argued that providers may have a better understanding than awardees of the needs of PROMISE participants and that the providers may also need more incentives and resources to provide better services. A mechanism for payments to providers might be administratively complicated. For example, if participating youth and their families were to receive services from multiple providers, there would be a question of how payments should be allocated among the providers. Other TAP members suggested that providing payments directly to youth and families might be beneficial, giving them added inducements to achieve the desired outcomes. Likely, SSA will need more information about the interventions in order to make well-informed decisions about how and to whom payments should be made. The agency could use the demonstration to test different payment strategies.

The appropriate time frame for observing a reduction in SSI benefits due to increased earnings of PROMISE participants and for issuing the associated incentive payments will depend on the population targeted by each project and the service model of each project. SSI benefit reductions in the short term might result from services delivered to parents, whose increased earnings could reduce their benefits and/or those of the target youth through deeming.⁴ However, some TAP members expressed their concern that the linkage between such short-term savings and the long-term independence of youth might be tenuous; that is, increased parental earnings might not affect benefit receipt by SSI youth when they become adults. Benefit reductions in the short term might also result from improved educational and vocational services for older members of the target population (youth ages 16 and 17), which could improve their chances of finding employment relatively quickly. The TAP members were in general agreement that a relatively long follow-up period, perhaps 7 to 10 years, would be needed to observe the full effects of the intervention on SSI benefit receipt, though impacts might well persist beyond that time horizon. This is particularly so because the PROMISE target population includes minors who are many years away from adulthood. Incentive payments tied to benefit reductions over a long follow-up period could encourage

⁴ Household members' earnings and assets are included in determining the SSI benefit amounts for their children, even children who receive benefits as adults (ages 18 and older).

PROMISE projects to deliver intensive services that might result in long-term independence for substantial proportions of the target population.⁵

Although this issue was not mentioned by TAP members, incentive payments based on reductions in SSI benefits could also reflect any concurrent changes in SSDI. Some participating SSI youth could work enough to qualify for SSDI, especially given that it is easier for individuals in their late teens and twenties to qualify for SSDI (fewer quarters of covered earnings are required than for older applicants). For these youth, the increase in SSDI benefits could more than offset the reduction in SSI benefits. Conversely, PROMISE could result in lower SSDI benefits if participants were diverted from applying for SSDI because of their increased earnings. With modifications to reflect differences in ages, these same points could apply for parents on SSI. SSA could structure its incentive payments under PROMISE so as to not reward cost-shifting from SSI to SSDI by including SSDI benefits in the calculation of benefit savings.

To facilitate administration and minimize associated costs, the TAP recommended that SSI-related incentive payments be based on administrative data that are already being collected (for example, SSA administrative records). This would have four advantages. First, it would provide complete information (that is, no data would be missing) on project participants. Second, because these records show actual benefit payments, they would provide the true value of the outcome, rather than an estimate (as would be the case with measures obtained through a survey or reports submitted by awardees). Third, the cost of obtaining and processing the data would be small—SSA staff could conduct the analysis at specified times during the intervention. Fourth, this approach to calculating incentive payments could be replicated on a larger scale, such as for a possible subsequent national program with some of the key features of PROMISE.

B. What Are the Interrelationships Between the Evaluation and the Incentive Payments?

As part of the PROMISE evaluation, the TAP members recommended using a comparison group to help determine whether the intervention had an effect on benefits beyond what would have occurred in the absence of the intervention. Some portion of PROMISE participants would leave the SSI rolls without an intervention. For example, 42 percent of SSI minors stop receiving benefits at the initial stage of the age-18 benefit redetermination process (under which youth are assessed against the adult standards), though about one-quarter of those whose benefits cease either successfully appeal the initial determination or reapply for benefits within four years.⁶ Additional youth would leave as a consequence of their employment or that of their parents. Some might accumulate enough quarters of covered earnings to qualify for SSDI. The TAP members recommended that such reductions in SSI benefits not generate incentive payments.

A simplistic methodology for determining incentive payments would be to base them on the changes in benefits for PROMISE participants between the preintervention period and a follow-up period. This approach would probably result in excessive payments because it would not adequately

⁵ Under a long-term perspective, the cumulative savings over multiple years are potentially quite large, justifying incentive payments that would be sufficient to fund intensive services.

⁶ Hemmeter, Jeffrey, and Elaine Gilby. “The Age-18 Redetermination and Postredetermination Participation in SSI.” *Social Security Bulletin*, vol. 69, no. 4, 2009, pp. 1–25.

account for transitions such as those described above—that is, awardees would receive payments for reductions in benefits that would have occurred regardless of the intervention. Such excess payments could be avoided by incorporating a comparison group into the methodology for determining incentive payments. While the evaluation of the PROMISE demonstration will most likely include a comparison group, a nationwide rollout of a PROMISE-like program very likely would not. Here we present two possible models for using the evaluation and its associated comparison group to provide SSA with guidance on how to structure incentive payments in a national program:

- In the first model for incentive payments, SSA would set a benchmark for determining incentive payment amounts based on changes over time in benefits received by participants. For example, SSA could provide payments based on every dollar of reduction in aggregate benefits received by the participant group beyond some benchmark, such as a 5 percent reduction relative to the aggregate benefits in the preintervention period. The benchmark, which would be based on an analysis of SSA administrative data on benefits, would represent the benefit reduction that would be expected for participants absent the intervention. At various points in the demonstration, the evaluator or SSA could assess the cost-effectiveness of the resultant incentive payments by comparing the SSI benefit reductions of the participant and comparison groups. A cost-effective approach (from the perspective of SSA) would be one in which the cumulative aggregate incentive payments were less than the cumulative difference in aggregate benefit reductions between the participant and comparison groups. SSA could use the results from the cost-effectiveness analysis to adjust the benchmark for incentive payments to increase the likelihood that future payments would be cost-effective. Similarly, SSA could use the findings from this approach to establish payment benchmarks for a hypothetical future national program with many of the key features of PROMISE.
- The second model for incentive payments would leverage the comparison group to ensure that the payments are continuously cost-effective during the demonstration. This model would base incentive payments on contemporaneous differences in outcomes between the participant and comparison groups for the PROMISE evaluation. SSA could provide payments to awardees based on a proportion of the difference in aggregate benefit reduction between the participant group and the comparison group. For example, if the proportion were set at 90 percent, then SSA would pay 90 percent of the benefit savings to the awardee in the form of incentives and retain 10 percent as net savings for the agency. The benefit and payment calculations could be conducted periodically, such as on a quarterly basis. A drawback of this model relative to the first model is that it could not be directly implemented in the context of a national program, but it has the advantage of ensuring that the incentive payments in the demonstration would not exceed SSA's benefit savings. This approach would generate findings that could be the basis for simulations of cost-saving payment strategies for a national program.

Both of these models would provide SSA with detailed information about two parameters in a national incentive payment policy that would save costs: the benchmarks for local entities to receive payments and the level of payments.

In addition to the above, the evaluation can provide SSA with opportunities to assess other aspects of the mechanism involved in making incentive payments. The TAP members recommended a payment mechanism that would allow payments to be made quickly and with a minimum of documentation on the part of the awardee. The evaluation could assess whether its mechanism met those criteria, as perceived by the awardees (through surveys or qualitative interviews). The evaluation could also identify barriers to obtaining payments, show how recipients used the payments, and explore issues in the use of administrative data for the payments (such as the amount of effort needed to distribute payments and how quickly payments could be made).

C. How Should Incentive Payments to Awardees Be Structured for Intermediate Outcomes?

The TAP members recognized that incentive payments based strictly on observed benefit reductions might not be realistic in the context of a limited-duration demonstration. Significant benefit reductions might occur only in the long term, particularly for projects that emphasize benefit savings only when the participating SSI youth become adults. The PROMISE projects are expected to operate over a five-year period that will encompass project start-up, service delivery, and close-out activities. Assuming that the projects begin enrolling and serving participants in FY 2014, the assessment of SSI benefit savings for the older participants as they become adults (about 7 to 10 years after enrollment) would occur in FY 2021 at the earliest, and later if enrollment activities were to extend over several years. For projects serving younger members of the PROMISE target population, an even longer period would be needed to assess their SSI savings as adults. Additionally, depending on the ages of the enrolled youth, a focus on near-term benefit reduction could be counterproductive in the long run. It might result in an overemphasis on employment for younger enrollees, to the exclusion of services and activities that could have even larger positive effects on long-term employment outcomes, such as completing high school and enrolling in postsecondary education.

For the purposes of the demonstration only, an alternative approach would be for SSA to provide incentive payments to awardees based on intermediate outcomes achieved by participants. Such payments would reflect a two-tier payment structure in which awardees would be eligible first for payments based on goals that might be achieved more quickly than benefit savings and then for payments based on actual benefit savings, which would likely take longer to be realized. This structure could bolster efforts of those involved in the PROMISE projects and encourage the achievement of more immediate objectives for youth, particularly if the first-tier payments are substantial and the awardees and their partners view the milestones for those payments as being fair and achievable. This two-tiered payment structure might incentivize the projects to provide more-intensive services to individual youth and their families over a longer period than they might otherwise do, thus increasing the likelihood of long-term reductions in disability benefits.

The TAP members emphasized that intermediate outcomes that may be the basis for incentive payments should be ones that have been found to be precursors to benefit reductions, which are the ultimate objective. Such outcomes would likely involve education and employment, with education emphasized over earnings from employment, given the young ages of many in the target population of minors on SSI. Few of these youth could be expected to have significant earnings in the first few years of the demonstration; however, it would be reasonable to expect some of the older ones to have part-time jobs and other work-based experiences. The intermediate outcomes and associated milestones should be carefully selected and structured to minimize unwanted and possibly counterproductive distortions of provider services and participant behaviors. For example, if high

school enrollment were an outcome eligible for intermediate incentive payments, schools might alter their procedures to keep more participants enrolled after age 18. This might have little impact on the participants' independence or later benefit receipt, though the awardee would receive payments related to the outcome. High school graduation might be a preferable intermediate milestone, one more likely to be associated with higher earnings and lower benefits in the long run.

The TAP members supported the designation of a common core set of intermediate outcomes across PROMISE awardees, but they also recommended allowing leeway for additional outcomes specific to each awardee that would be based on the awardee's specific interventions and logic models. Some suggested that applicants propose the additional intermediate outcomes for which they could receive incentive payments in their applications. Presumably, these would reflect what the applicants believed their projects could achieve in working with SSI youth and their families that would ultimately lead to benefit reductions. It is likely that the proposed outcomes would take into account both the nature of the planned services and the characteristics of the youth and families that would be served. For example, the intermediate outcomes for youth enrolled in high school might be very different from those for younger participants or for parents.

The TAP recommended that, to the extent feasible, SSA's mechanism for incentive payments based on intermediate outcomes mirror the mechanism the agency would use for incentive payments based on benefit reductions. In particular, these payments should reflect group-level (rather than individual-level) outcomes, the outcomes should be clearly defined and measurable, and the outcomes should be available in administrative data files. As with incentive payments based on benefit reductions, there should be no restrictions on how awardees use the payments received for achieving intermediate milestones; they should be free to (1) use them to directly fund additional services for project participants, (2) pass them along to their PROMISE partners or participants, or (3) use them to support other programs. By implementing incentive payments based on intermediate outcomes in this way, SSA would gain experience with the mechanism for providing incentive payments based on SSI benefit reductions.

Possible mechanisms for incentive payments based on intermediate outcomes follow the two models for payments based on benefit reductions that were presented in Section B of this chapter. Recall that the first model relies on *benchmarks*, while the second model leverages the anticipated existence of a *comparison group* in the PROMISE evaluation design. Here we elaborate on each of these models in the context of incentive payments based on intermediate outcomes:

- Incentive payments based on benchmarks would entail assessing aggregate intermediate outcome measures for PROMISE participants (for example, the proportion of participants receiving VR services) relative to externally defined benchmarks. Incentive payments would be based on either the achievement of the benchmarks or the gaps between the participant outcomes and the benchmarks. The benchmarks could be set based on published reports of these outcomes for similar groups of youth (such as reports based on the 2012 National Longitudinal Transition Study), on original analyses by the evaluator of other survey or administrative data on comparable youth, or on the outcomes for participants at the time of enrollment (adjusted by some multiplier reflecting the desired improvement due to PROMISE). We are concerned that benchmarks based on these sources may not be fully appropriate for the PROMISE population and could lead to excessive or insufficient incentive payments.

- Under the comparison group approach, aggregate intermediate outcomes for participants would be assessed relative to outcomes for a comparison group, and the incentive payments would be based on differences between the two. This approach to incentive payments would be similar to that used on projects entailing social impact bonds. It would ensure that the incentive payments would be based on documented interim impacts of the intervention. However, this approach would probably not be feasible under a national rollout of a PROMISE-like program due to the probable absence of a comparison group.

Incentive payments based on intermediate outcomes may be necessary only in the context of a time-limited demonstration. In a full implementation of a PROMISE-like program, a longer time horizon would probably apply, which could eliminate the need for payments based on intermediate outcomes. Given this, the second of the two approaches described above—the comparison group approach—would appear to be the preferred design for incentive payments based on intermediate outcomes in the PROMISE demonstration.

Regardless of the mechanism for incentive payments based on intermediate outcome measures, a determination would have to be made regarding the source of data on the measures. The evaluation may include survey and management information system (MIS) data that could provide detailed information on the intermediate outcomes; however, such systems are not likely to be incorporated in a national program. State or federal administrative data, such as state VR and education records, could be useful in this regard, as repeated extracts could be obtained and the data could encompass both participant and comparison group members, and similar mechanisms could be used for a national program. However, administrative data sources might provide only limited measures that would be appropriate bases for intermediate incentive payments. If payments based on intermediate outcomes are unlikely to be elements of a future national PROMISE-like program and if SSA is confident that some incentive payments based on benefit reductions will be made during the PROMISE demonstration, then the intermediate payments in the demonstration could be based on survey or MIS data. The experience gained during the demonstration in making payments based on benefit reductions would probably suffice to inform the development of a mechanism for such payments in a national program.

Some TAP members voiced concerns about the risks in providing incentive payments based on intermediate outcomes. They had three specific concerns. First, awardees and service providers might focus on the intermediate outcomes to the detriment of PROMISE's ultimate goal, which is to reduce disability benefits. Second, there may be no good evidence linking intermediate outcomes to longer-term benefit reductions, although there may be evidence of linkages between potential intermediate outcomes and improvements in employment and earnings, which should be accompanied by reductions in benefits. Third, success in the short term may not be sustained in the long term. It is not unusual for social welfare interventions to have larger impacts based on early follow-up data but smaller or no impacts on similar outcomes based on longer-term follow-up data.⁷

⁷ Schochet, Peter Z., John A. Burghardt, and Sheena M. McConnell. "Does Job Corps Work? Impact Findings from the National Job Corps Study." *American Economic Review*, vol. 68, no. 5, December 2008, pp. 1864–1886.

III. DEVELOPING AND MEASURING INTERVENTIONS AND PARTNERSHIPS

The anticipated PROMISE projects will likely differ with respect to the services they will deliver to youth and their families and the state and local agencies and providers involved. This will necessitate an evaluation that includes a solid process study, but the variation in projects, perhaps even across sites within a state, might limit the degree to which the measures for the process study can be uniform. Tracking the services delivered to participants and the partnerships that develop across agencies involved in PROMISE can provide important information for understanding the projects. SSA and the other federal agencies in the PROMISE consortium will have to decide how much detail to collect about service delivery, partnerships, and outcomes. In addition to supporting the evaluation, this information could be critical to the provision of effective technical assistance. By closely tracking services and outcomes, feedback could be provided to the projects to improve their delivery of services, leading to substantially better outcomes in SSI benefit receipt. In general, the more information collected by the process study, the more the funding agencies will know about the projects and the more insight they will have into the outcomes observed. However, there is a trade-off: the more information, the greater the cost of collecting it, including the burden on awardees and local participating entities.

The PROMISE TAP made the following recommendations for developing and measuring interventions and partnerships:

- The evaluation's process study should assess the types and intensity of services delivered to youth and families, the characteristics of the state and local environments, aspects of state and local agency and provider partnerships, and the fidelity of the implemented projects to the PROMISE model and the awardees' proposed interventions.
- The process study should assess the impacts of the projects on the receipt of services by incorporating a counterfactual component.
- MIS data should be a key element of a comprehensive process study.
- As partnerships may be important aspects of the PROMISE projects, the process study should assess how much state and local agencies and providers actually work together as opposed to simply voicing support for improved cooperation.
- The PROMISE request for applications (RFA) process provides an opportunity to orient applicants to the demonstration objectives and to assess their capacity to design and run successful projects. The RFA could specify, as concretely as possible, the long-term outcomes expected for participating youth and their families, while inviting applicants to propose intermediate outcomes that are precursors to the ultimate outcomes. Applicants could also be instructed to provide evidence of linkages between their proposed intermediate outcomes and the long-term objectives of PROMISE.
- The RFA could require applicants to present logic models showing how their proposed interventions would generate the intended long-term outcomes and the related intermediate outcomes.
- The RFA could require applicants to describe their current use of management and reporting systems as an indicator of their ability to incorporate an MIS in their proposed projects.

We make the following additional recommendation, based on TAP member comments and what is currently known about the design for PROMISE:

- Although the TAP members generally agreed that having the PROMISE projects adopt a common MIS would benefit the process analysis, our experience suggests that it would be unrealistic to require a uniform, or largely uniform, MIS. An alternative approach that would be largely responsive to the concerns of the TAP members would be for the funding agencies to specify the parameters for the data to be captured by an MIS but to allow the awardees to determine what system to use to implement those specifications.

This chapter first reviews the critical objectives and components of the PROMISE process study. A discussion of issues related to developing and using an MIS across multiple PROMISE projects follows in the next section. The third section identifies concerns involved in developing an understanding of the partnership aspect of PROMISE projects. The chapter concludes with a consideration of RFA requirements for awardee support of the process study.

A. What Are the Objectives and Components of a Process Study of PROMISE?

The TAP members agreed that the fundamental objective of the PROMISE process study should be to understand the interventions that the awardees will implement. The awardees are likely to propose different sets of services with varying partners, providers, and target populations, so it will be important to document the interventions that each awardee implements. In addition, the process study should inform the disability research and policy community about promising practices in working with youth with disabilities and their families.

The core components of the process study identified by TAP members included the types and intensity of services delivered to youth and families, the characteristics of the state and local environments, aspects of state and local agency and provider partnerships, and the fidelity of the implemented projects to the PROMISE model and the awardees' proposed interventions. Each of these components would capture an important aspect of the projects, contributing to a comprehensive understanding of the PROMISE demonstration. The interventions will likely vary in the types and amounts of employment, education, benefits planning, and other services delivered, depending in part on the needs of their specific target populations. The projects will have different state and local economic, political, and service environments, which could affect the design and implementation of the interventions, particularly in regard to the participating agencies and service providers. The emphasis on partnerships and systems change will likely necessitate that the evaluation assess how agencies work together to design and deliver services to youth and families. Finally, as each awardee will have proposed an intervention based on the PROMISE model; assessing the extent to which the awardee provided or arranged for services that reflected what was proposed is likely to be important. These core components of the process study would provide a basis for understanding the outcomes for youth and families that arise from the projects.

The TAP members also identified additional components of the process study that would enhance understanding of the PROMISE projects. An analysis of participation in the projects by eligible youth and their families would document their interest in the services offered and the effectiveness of project outreach efforts. An assessment of the receipt of incentive payments would illustrate the administrative challenges and successes of using such a system. If waivers and blended funding are important aspects of the PROMISE model, then the process study may need to

document the use of waivers and the sharing of resources among partner agencies. Finally, measuring the satisfaction of youth and families with the interventions could reveal whether they had good rapport with project staff and felt the services were worthwhile.

In addition to providing an understanding of the interventions that the PROMISE awardees will implement, the TAP members emphasized the need for the process study to assess the impacts of the projects on the receipt of services. Although youth and families in the evaluation's counterfactual groups will not be eligible for PROMISE services,⁸ many of them will receive services through other sources, such as secondary schools and VR agencies. Some might even receive services from the same providers involved in the PROMISE projects. Conversely, some youth and families in the evaluation's treatment groups may receive no or few services. A survey of both types of groups could provide comparable data on the receipt of project services and counterfactual services from the perspective of the youth and their families. This approach would enable the evaluator to assess the average impact of the intervention on the receipt of services. Absent such a method, an alternative approach would be to assess this impact through administrative data available from certain providers (such as VR records). However, this approach might be too narrow given the number of providers accessed by and the depth of services received by SSI youth and their families.

Data sources for the process study could include MIS data, site visits, semistructured interviews with project staff, participant focus groups, survey data on youth and their families, and other data as appropriate. MIS data could show the delivery of services to participants from the perspectives of the awardees or service providers and could be a critical source of information on participants' outcomes during their participation. Site visits and semistructured interviews with project staff and the staff of other service providers could inform the evaluator about qualitative aspects of the interventions, including the counterfactual service environment and relationships among partner agencies. Focus group discussions with project participants could provide useful information about youth and their families' perceptions of services and staff, as well as the facilitators of and barriers to success from their perspectives. As noted in the preceding paragraph, surveys of members of the evaluation's treatment and counterfactual groups could provide comparable data on the receipt of services, as well as detailed information on project participation and satisfaction with services from the perspective of the youth and their families.

B. How Could an MIS Be Incorporated in PROMISE?

The TAP members were in agreement that MIS data should be a key element of a comprehensive process study of PROMISE. For the evaluator and the funding agencies, an MIS could document the services delivered by the projects and track the immediate outcomes—such as high school graduation, waiver use, and employment—of participating youth and their families. For providers, it could document the services that they delivered and the outcomes that their participants achieved. An MIS could also be used to generate periodic reports that the funding agencies could use to monitor fidelity to the intervention model and outcomes achieved. These

⁸ In a given evaluation site, the counterfactual group would consist of either youth (and their families) who had been randomly assigned to a control group or youth who had been selected into a comparison group by some means other than random assignment. In either case, the members of the counterfactual group would not be eligible for PROMISE services. See Chapter IV for additional details.

reports could identify areas in which technical assistance could help each project achieve its desired outcomes.⁹

The importance of MIS data to the process study is underscored by the fact that there are no good alternatives for developing a solid understanding of PROMISE services. Although useful in understanding certain aspects of project implementation, a survey of youth and their families (if included in the PROMISE evaluation design) would be unlikely to provide as much detail on services as an MIS, and the data that it would provide might be subject to considerable recall error. Site visits and staff interviews could provide useful qualitative information on project services, but these are unlikely to yield the detailed quantitative data on services that could be obtained from an MIS.

Although the TAP members generally agreed that having the PROMISE projects adopt a common MIS would benefit the process analysis, our experience suggests that it would be unrealistic to require a uniform, or largely uniform, MIS. Applicants may propose very different approaches to achieving similar objectives. Awards may be given to different types of agencies with different types of partners. Multiple agencies and providers are likely to be involved in each project, with some being responsible for services to youth and some to parents, and each having an existing MIS for recording and managing services. An additional uniform MIS across all of the awardees, agencies, and providers could be complicated to design and burdensome to implement, resulting in incomplete and unreliable MIS data. Considering these factors, the cost of such a system might be prohibitive, given the level of funding anticipated for the PROMISE evaluation.

An alternative approach that would respond to the concerns of the TAP members would be for the funding agencies to specify the parameters for the data to be captured by an MIS but to allow the awardees to determine what system to use to implement those specifications. The parameters could include a broad array of services and outcomes. Within these parameters, the awardees would identify the specific measures to be captured. These should be relevant for the awardee and the proposed intervention, be clearly defined to promote consistent data entry, and entail minimal administrative burden. Examples of potential parameters include an indicator of whether a participant's family ever attended one of the project's benefits counseling sessions (for an intervention that includes benefits counseling for parents) and a measure of the number of sessions on résumé writing and/or work site tours that a participant attended (for an intervention that includes such sessions). Measures conforming to such parameters that could be easily collected within the structure of existing agency and provider systems could provide a solid foundation for understanding the services delivered to project participants. The measures could also be ones that awardees and providers would view as critical to their management and monitoring of project activities.

⁹ Under the Youth Transition Demonstration evaluation, the reporting functions of an MIS were critical to the design and delivery of well-focused technical assistance to the demonstration projects. See Martinez, John, Thomas Fraker, Michelle Manno, Peter Baird, Arif Mamun, Bonnie O'Day, Anu Rangarajan, and David Wittenburg. "The Social Security Administration's Youth Transition Demonstration Projects: Implementation Lessons from the Original Projects." Washington, DC: Mathematica Policy Research, February 2010. Available at http://www.ssa.gov/disability_research/docs/YTD%20Special%20Process%20Report%20202-22-2010.pdf. Accessed December 27, 2011.

The responsibility for ensuring that MIS data are collected uniformly could reside with awardees, which could monitor the data from all locations in which their PROMISE interventions operate and from all providers involved in delivering services. The awardees could also be responsible for maintaining a consolidated MIS data file.¹⁰ The evaluator could be responsible for reviewing awardees' MIS designs, making recommendations for revisions to or acceptance of those designs to ED or SSA, and conducting periodic checks on the adequacy of the data in those systems. Awardees are likely to require technical assistance in modifying their systems, inputting data, and using the data to manage their projects. ED or SSA could contract for such assistance. To encourage MIS data quality, ED could predicate continued funding for a project, or incentive payments, on the adequacy and timeliness of the MIS data.

C. What Consideration Should Be Given in the Process Study to the Partnership Aspect of PROMISE?

Because partnerships may be important aspects of the PROMISE projects, the TAP members agreed that the process study should assess how much state and local agencies and providers actually work together as opposed to simply voicing support for improved cooperation. PROMISE projects might entail systems change activities that would enhance collaboration by agencies and providers that already serve SSI youth and their families. Beyond this, the projects could seek to include entities that have not typically worked with this population. Because the awardees will operate in unique environments and propose different interventions, the TAP members suggested that ED be cautious in requiring projects to include specific partners. Further, dictating the specific agencies, programs, or providers involved could result in suboptimal interventions if those entities were not fully committed to PROMISE or were unable to agree on a shared vision of how best to serve SSI youth and their families.

If cross-agency cooperation is integral to the design for a PROMISE project, then the project's MIS should reflect it. The MIS could include measures such as referrals among cooperating agencies and providers, coordination and alignment of service plans for participants, co-enrollment of multiple family members to receive concurrent services from a provider, and coordinated handoffs from one provider to another. The MIS could also include systems-level measures of involvement, such as blended funding from multiple agencies, shared staff, and project trainings involving staff from multiple agencies.

Site visits, including discussions with project staff and with staff of the relevant agencies and providers, could also yield valuable data with which to assess the effectiveness of partnerships. These data could address how much these organizations worked together to provide project services, how much the staff of the various providers coordinated their work with participants, and how agencies and providers changed as a result of their involvement in the PROMISE project.

¹⁰ A web-based MIS could be efficient for collecting data from the various organizations involved in a PROMISE project while also providing them with real-time access to the data; however, TAP members did not strongly advocate such a system. Similar data could be collected through other means, such as monthly reports by participating providers, and then entered into an awardee's MIS. It could be up to the awardees to design or modify systems that would meet their needs and those of the other organizations involved in the projects, and that would meet the needs of the PROMISE funding agencies and the evaluator.

D. What Elements of the Process Study Should Be Incorporated in the RFA for PROMISE Projects?

The TAP members noted that the PROMISE RFA process provides an opportunity to orient applicants to the demonstration objectives and to assess their capacity to design and run successful projects. The RFA could describe the ultimate objectives of PROMISE, illustrate the range of possible approaches, and provide details on the funders' expectations about implementation of the projects. The RFA could also ground applicants in all major aspects of the evaluation (not only the process study), including the expectations for them in terms of MIS and other data collection activities that might either be their direct responsibility or require their cooperation with the evaluator. With such elements incorporated in the RFA, the funding agencies would be better able to assess the capacity and willingness of the applicants to perform the activities required for the evaluation in addition to their capacity to design and implement successful PROMISE projects.

The TAP members suggested that the RFA specify, as concretely as possible, the long-term outcomes expected for participating youth and their families, while inviting applicants to propose other outcomes that are precursors to the ultimate outcomes. These precursor outcomes would likely reflect what the applicants believe they could achieve given their proposed interventions and target populations. Strong applications would include research-based evidence of linkages between their proposed outcomes and the long-term objectives of PROMISE.

One suggestion that arose during the TAP meeting was that the RFA could require applicants to present logic models showing how their proposed interventions would generate the intended long-term outcomes and the related intermediate outcomes. A logic model can be an important part of a project proposal because it requires the applicant to specify the components of the proposed intervention and the mechanisms or pathways between them and the various outcomes. A logic model can also be the foundation for evaluation activities, providing clear sets of services and outcomes to be measured.

TAP members also noted that the RFA could require applicants to describe their current use of management and reporting systems as an indicator of their ability to incorporate an MIS in their proposed projects. Applicants with limited experience using electronic data systems to monitor and report on services and outcomes might be less qualified to implement and manage a PROMISE project, as this would require intensive MIS use, particularly to establish eligibility for incentive payments. Applicants with more experience in these areas that are able to document their current or past use of MIS data reports for tracking participant services might be more capable awardees for the purposes of the PROMISE demonstration. In addition to requesting such documentation, the RFA could specify that applicants outline how they would adapt their existing MIS or design a new one to capture key project inputs and outcomes based on their logic models.

IV. EVALUATION DESIGN

The overarching goal of a project evaluation is to determine whether and how the project worked, by producing research findings that can be attributed directly to the project as opposed to the characteristics of the participants or the environment in which the project operated. Proper attention to evaluation design issues for PROMISE could help SSA and the other involved agencies learn about how the demonstration projects were implemented, what effects they had on participating minors and their families, the effectiveness of the mechanism for incentive payments, and other issues relevant for extending PROMISE beyond the demonstration. The federal government can establish the basic parameters for the PROMISE evaluation design as it determines the objectives and features of the interventions. A major challenge for the evaluation is producing comparable research findings across the anticipated two demonstration projects, when those projects might entail very different interventions and environments.

The TAP made the following recommendations for the PROMISE evaluation design:

- The final development of the evaluation design should occur in conjunction with the final development of the interventions by the PROMISE awardees, with the evaluator being selected at the same time as or before awardees are selected. This would ensure that the designs for the interventions and the evaluation are compatible, that the evaluation design addresses the legitimate concerns of the awardees, and that the evaluation can realize its objectives.
- The TAP members recommended three component studies for the evaluation—an impact study, a process study, and a cost-benefit study—and suggested a wide range of research questions across these studies.
- Given the many questions to be answered by the impact study, specifying a very limited number of primary outcomes and a broader set of secondary outcomes could reduce the risk of drawing incorrect inferences from the findings. (This is a response to the “multiple comparisons” problem in evaluation research.)
- For the impact study, TAP members overwhelmingly recommended an experimental design (that is, a design involving the random assignment of PROMISE-eligible youth to either a treatment group that would be offered project services or to a control group that would not have access to those services) over a nonexperimental design. This is because an experimental design typically carries less risk that the estimated impacts might reflect factors other than the intervention itself.
- Under an experimental design, randomly assigning only eligible youth and families who volunteer for a PROMISE project would likely result in a lower-cost evaluation and more-reliable estimates of impacts than randomly assigning all eligible youth in a project’s catchment area.
- Having certain consistent outcome measures across the research sites, with additional site-specific outcomes measures, would benefit the evaluation.¹¹

¹¹ In this chapter, *site* refers to the totality of areas in which a given PROMISE project would operate and in which youth would be randomly assigned for the impact study.

- The family as a unit of analysis represents an overarching data collection concern for the PROMISE evaluation because of issues surrounding informed consent and data access.

In addition to the above recommendations of the TAP, our analysis of TAP member comments and consideration of PROMISE design issues led us to the following conclusions:

- Based on what is currently known about the objectives for PROMISE and on the range of research questions proposed by the TAP, the evaluation's three component studies might address the following primary research questions:
 - The impact study could answer the following questions: What are the impacts of the interventions on the principal short- and intermediate-term outcomes specific to each project site? What are the impacts of the interventions on the long-term receipt of disability benefits by the target youth and their families?
 - The process study could answer the following question: What services did the projects actually deliver to participating youth and their families?
 - The cost-benefit study could answer the following question: What are the long-term savings in SSI benefits attributable to the projects and are they sufficient to sustain the ongoing operation of the projects?
- The TAP members recognized that a potential complication with an experimental design is the threat of spillover effects on control group members if the interventions were to involve agency-level reforms or systems change components. The issues and risks inherent in an evaluation facing such interventions warrant careful consideration in the case of PROMISE.
- Potential data collection challenges associated with surveys of youth and families include the need to capture information on services and outcomes that are likely to vary across sites, delays in obtaining data, and the need to hold down data collection costs, given the evaluation's limited budget.
- Potential data collection challenges associated with administrative information systems include accessing the data in a timely manner, ensuring data security, and collecting consistent data across sites.

This chapter provides details on the TAP's recommendations regarding the PROMISE evaluation and on the additional considerations that we have identified. We begin by identifying the key research questions for the evaluation, followed by a discussion of experimental and nonexperimental design issues. The chapter concludes with a discussion of the data collection challenges facing the evaluation.

A. What Are the Key Research Questions for the PROMISE Evaluation?

The final development of the evaluation design should occur in conjunction with the final development of the interventions by the PROMISE awardees, with the evaluator being selected at the same time as or before awardees are selected. This would ensure that the designs for the interventions and the evaluation are compatible, that the evaluation design addresses the legitimate concerns of the awardees, and that the evaluation can realize its objectives. The TAP members recommended that the PROMISE evaluation have three component studies—an impact study, a

process study, and a cost-benefit study—and suggested a broad range of research questions across these studies (Table IV.1). Here we discuss the various questions that could be answered by and the issues associated with each of the component studies, along with suggestions for additional aspects of the evaluation.

The two key questions that could be answered by the impact study are (1) What are the impacts of the interventions on the principal short- and intermediate- term outcomes specific to each project site? and (2) What are the impacts of the interventions on the long-term receipt of disability benefits by the target youth and their families? The first question could be answered by an evaluator (which could be a contractor, a team of SSA researchers, or a combination of the two) during the anticipated period of demonstration services. The evaluator would focus on benefit, employment, education, health, and other outcome measures that are either relevant across PROMISE projects or can be tailored to the design for a specific project (for example, the high school graduation rate for a project that provides supports for high school completion or parental earnings for a project that offers employment services to family members). Findings regarding impacts on these interim outcomes can be important on their own, and they can also suggest refinements to project designs that might improve impacts on the critical long-term outcome, SSI benefits. Depending on the evaluation design, this aspect of the impact study could rely on data from youth/parent surveys and state and federal administrative files—sources that could provide information on outcomes for members of both the treatment and control groups. The second key question for the impact study could be answered by an evaluator after the period of PROMISE demonstration services. This means that the impact study would have to be designed to support this approach to estimating long-term impacts on disability benefits. The primary data source for this part of the evaluation would likely be SSA administrative files, which would contain SSI and SSDI benefit histories for the youth participating in the evaluation and their family members. Additional sources of data on other long-term outcomes of potential interest include SSA’s master earnings file and administrative files maintained by ED (such as the RSA-911 data file) and DHHS (such as Medicaid and Medicare records files).

Using multiple outcome measures to answer the questions to be addressed by the impact study would increase the risk of finding statistically significant estimated impacts of a PROMISE project when in fact the project was ineffective and those results were obtained simply by chance.¹² Recognizing this issue, some TAP members suggested specifying a very limited number of primary outcomes and a broader set of secondary outcomes for the impact study to reduce this risk. The principal conclusions on the effectiveness of the interventions would be based on estimated impacts on the primary outcomes, whereas estimates based on the secondary outcomes would be used to improve understanding of those conclusions. For example, employment was an important outcome domain in the YTD evaluation. The primary outcome in this domain was having paid employment following random assignment. The secondary outcomes to help understand the impact

¹² Schochet, Peter. “Guidelines for Multiple Testing in Impact Evaluations of Educational Interventions.” Princeton, NJ: Mathematica Policy Research, May 2008. Available at <http://www.mathematica-mpr.com/publications/PDFs/EducationalInterventions.pdf>. Accessed February 8, 2012.

Table IV.1. Potential Research Questions for the PROMISE Impact, Process, and Cost-Benefit Studies

Impact Study
<p>Key Questions</p> <ul style="list-style-type: none"> • What are the impacts of the interventions on the principal short- and intermediate-term outcomes specific to each project site? • What are the impacts of the interventions on the long-term receipt of SSI (and SSDI) benefits by the target youth and their families? <p>Additional Questions</p> <ul style="list-style-type: none"> • What are the short-term impacts of the intervention on benefit receipt, employment, education, health, and other outcomes for SSI youth and family members? • What are the intermediate-term impacts of the intervention on benefit receipt, employment, education, health, and other outcomes for SSI youth and family members? • What are the long-term impacts of the intervention on employment, education, and other federal programs (such as Medicaid and Medicare) for SSI youth and family members?
Process Study
<p>Key Question</p> <ul style="list-style-type: none"> • What services did the projects actually deliver to participating youth and their families? <p>Additional Questions</p> <ul style="list-style-type: none"> • What was the local context (economic, employment, education, and social service environments) for each project? • What was the nature and quality of the proposed interventions? • Were the projects implemented with fidelity to the proposed intervention? • What outreach efforts did project staff make to enroll youth and families? • What was the participation rate among those eligible for the interventions? • How did participants differ from nonparticipants? • What types of services did participants (youth and families) receive in the short and intermediate terms? What was the intensity of those services? • What was the short-term impact of the intervention on the service use of SSI youth and family members compared with the counterfactual? • How satisfied were participants with the services they received? • What waivers were implemented for the projects and how did participants use them? • What kinds of employment and education experiences did youth and family members have? • How were incentive payments distributed? • How did recipients use their incentive payments? • What challenges and barriers did awardees face in implementing the projects? • How were different agencies and providers involved in the projects? • How did PROMISE change the relationships among participating agencies and providers (that is, PROMISE partnerships)? • How did PROMISE influence the policies and behaviors of participating agencies?
Cost-Benefit Study
<p>Key Question</p> <ul style="list-style-type: none"> • What are the long-term savings in disability benefits and are they sufficient to sustain the ongoing operation of PROMISE projects? <p>Additional Questions</p> <ul style="list-style-type: none"> • What is the steady-state cost of providing services? • What is the cost of providing incentive payments? • What are the costs and benefits from the perspective of the participants? • What are the costs and benefits from the perspective of the awardees? • What are the costs and benefits from the perspective of the state government? • What are the costs and benefits from the perspective of the federal government? • What are the costs and benefits from the perspective of society?

on paid employment included the percentage of weeks employed and the number of jobs held following random assignment and employment status at the time of the follow-up survey.¹³

The process study could answer the following primary question: What services did the PROMISE projects actually deliver to participating youth and their families? The awardees may provide or arrange for a diverse set of services for project participants to promote their independence from SSI benefits, and those services may vary across projects. The previous chapter addressed the critical components and issues for the process study. In addition to describing the services delivered to participants, the process study could be designed to answer questions about fidelity to the intervention design, the state and local service environments, and project partnerships, among others. Table IV.1 provides a more complete list of potential questions for the process study.

The cost-benefit study could answer the following primary question: What are the long-term savings in SSI benefits attributable to the PROMISE projects and are they sufficient to sustain the ongoing operation of the projects? The answer to this question would address a central issue with PROMISE—whether the projects can improve the independence of SSI youth and their families such that the savings in benefits can fund the services provided.¹⁴ This issue is particularly salient, given the large potential long-term savings that a project could generate by helping a youth exit the SSI rolls.

The discussion of incentive payments in Chapter II considered only the costs of those payments to SSA and the benefits to the agency in terms of reduced disability benefits; however, the cost-benefit study should be expanded to incorporate other perspectives. Consideration of these perspectives, such as those of PROMISE awardees, project participants, the federal government (defined more broadly than only SSA), state governments, and society as a whole, would entail analyzing additional savings and revenue enhancements, such as reduced Medicaid and SNAP expenditures, increased income and payroll taxes, increased value of economic output, and decreased state SSI payments. Measures of these financial benefits would have to be available through existing administrative data sets for the evaluation to estimate long-term impacts on them. The design for the cost-benefit study should consider the possibility that few financial benefits may materialize during the period of PROMISE demonstration services. Thus, the evaluator would be responsible for measuring project costs per participant and for designing an analytic framework that would enable a subsequent evaluator (again, this might be a team of SSA researchers) to incorporate estimates of the long-term financial benefits in the actual cost-benefit study. A key challenge for the cost-benefit study will be to obtain a relevant measure of the steady-state project cost per participant. The steady-state cost would ideally be covered by SSA incentive payments based on benefit savings attributable to the PROMISE projects (that is, based on estimated long-term impacts on benefits).

¹³ See, for example, Fraker, Thomas, Alison Black, Joseph Broadus, Arif Mamun, Michelle Manno, John Martinez, Reanin McRoberts, Anu Rangarajan, and Debbie Reed. “The Social Security Administration’s Youth Transition Demonstration Projects: Interim Report on the City University of New York’s Project.” Washington, DC: Mathematica Policy Research, April 2011. Available at <http://www.ssa.gov/disabilityresearch/documents/CUNY%20Final%20Interim%20Report%204-5-11.pdf>. Accessed December 21, 2011.

¹⁴ We have noted that this approach to financing the PROMISE projects would be similar to the use of social impact bonds as the basis for paying providers to obtain specific outcomes for the recipients of their services.

B. Should an Experimental or Nonexperimental Design Be Used?

TAP members overwhelmingly recommended an experimental design over a nonexperimental design for the PROMISE impact study. An experimental design—in which SSI youth and families meeting certain inclusion criteria (for example, youth ages 14 through 17 living in a specified geographic area and currently receiving SSI benefits) would be randomly selected to be offered the intervention—has many positive attributes. Most notably, with an experimental design, any estimated impacts can be attributed with confidence to the intervention, rather than to differences in characteristics between eligible individuals who did or did not have access to the intervention. This particular attribute would not be present to such a degree with a nonexperimental design, which could entail selecting a comparison group of SSI youth and families that resides outside the catchment area for a PROMISE project or using a regression discontinuity model. If a nonexperimental design obtained statistically significant impact estimates, there could be a lingering question as to whether those were actually due to differences in characteristics and environments between research cases that did or did not have access to PROMISE services.

Within the context of an experimental design for the impact study, there is an important question of who should be randomly assigned—all youth eligible for project services or only the subset of eligible youth who demonstrate interest in those services (volunteers). The TAP members preferred random assignment of volunteers because this approach often entails lower evaluation costs and stronger impact estimates. To see why this is so, consider the alternative approach of randomly assigning all eligible youth to either be offered project services (the treatment group) or not (the control group). If the PROMISE experience were to resemble those of other employment-focused interventions for people with disabilities, then the rate at which these treatment group members would take up project services would be low.¹⁵ This would imply small impacts of the interventions, based on a comparison of outcomes between treatment and control group members. To have a reasonable chance of detecting those small impacts, it would be necessary to enroll a large number of youth in the evaluation and collect follow-up data on them. This would entail greater data collection and processing costs than under the preferred approach of randomly assigning only volunteers. If the data collection were to include follow-up surveys, then the cost difference between the two approaches could be quite large. Furthermore, with random assignment of all eligible youth, PROMISE awardees would face a more challenging and costly task of recruiting treatment group members into their projects than they would if the focus of their recruiting efforts were youth who had already demonstrated an interest in the services being offered. However, there is a potential downside to randomly assigning only PROMISE volunteers: the impact estimates could not legitimately be generalized to the full PROMISE target population of minors on SSI. But, if PROMISE were ever to be rolled out nationwide, it would probably be as a voluntary program. The

¹⁵ The TTW participation rate for disability beneficiaries through December 2006 was 1.6 percent, with higher rates observed for younger beneficiaries. See Stapleton, David, Cindy Gruman, and Sarah Prenovitz. “Participation in Ticket to Work Continues to Grow but Assignments Under the Traditional Payment System Still Dominate.” Work Activity and Use of Employment Supports Under the Original Ticket to Work Regulations. Washington, DC: Mathematica Policy Research and Center for Studying Disability Policy, September 2009. Available at http://www.mathematica-mpr.com/publications/PDFs/disability/TTW_participation.pdf. Accessed January 4, 2012. The participation rate for Project NetWork was 8.0 percent for beneficiaries ages 18 to 29. See Rupp, K., and S. H. Bell. “Provider Incentives and Access in the Ticket-to-Work Program: Implications of Simulations Based on the Project NetWork Field Experiment.” In *Paying for Results in Vocational Rehabilitation: Will Provider Incentives Work for Ticket to Work?*, edited by K. Rupp and S. H. Bell. Washington, DC: Urban Institute Press, 2003.

TAP members thus felt that it would be appropriate to base the impact study on an experimental design that entails the random assignment of volunteers.

A PROMISE intervention that included significant agency-level reforms or systems change components would necessitate an alternative evaluation design that would likely include a nonexperimental component. With an experimental evaluation design, such an intervention could have spillover effects on control group members, compromising the impact estimates. For example, imagine that an awardee (such as a VR agency) implemented an intervention in one county that involved all county secondary schools, the community college, and all major private service providers. If the intervention included both special services for PROMISE participants and improved coordination of services among the participating organizations, then the members of the control group in an experimental design might benefit from the improved interagency relationships. The PROMISE funding agencies and the evaluator will have to think about the possible spillover effects on control group members of a proposed intervention, and they may need to consider alternatives or modifications to a standard experimental design.

TAP members recognized the complications that could be introduced by systems change interventions and mentioned several options for nonexperimental evaluation designs, such as a comparison group, difference-in-differences comparisons, and regression discontinuity. A feature common to these designs is a nonrandomly selected comparison group. In its simplest form, a comparison group design would require comparing outcomes for a treatment group with those for a comparison group. The comparison group might be drawn from youth and families that reside in a different geographic area than where the intervention is being implemented (such as another part of the county where the intervention is being implemented, another county in the same state, or another state) but otherwise meet the eligibility criteria for the PROMISE project. The validity of this design hinges on the identification of a comparison group that is truly similar to the treatment group. However, this can be very challenging because the two groups might differ on characteristics that are not observed in any available data set and because it can be hard to find two areas that are truly similar in key socioeconomic characteristics. Such differences could introduce bias to the impact estimates and limit the ability to attribute them to the intervention itself.

A difference-in-differences design is an extension of the basic comparison group design in which changes in outcomes for treatment group members between the baseline period and the follow-up period are compared with changes for comparison group members. This approach has the potential to control for dissimilarities between the two groups. This potential can be most fully realized when the outcomes are ones for which baseline differences may exist between the treatment and comparison groups, such as family earnings. However, a number of key outcomes for the PROMISE evaluation are ones for which baseline differences between the two groups are likely to be minimal or nonexistent, such as youth receipt of SSI benefits and youth employment and earnings. A regression discontinuity design may not be appropriate for PROMISE for the same reason as for a random assignment design. This approach requires a comparison group of individuals not offered the intervention, such as youth scoring below a threshold score. If a PROMISE intervention were to involve systems change, then youth and families not offered the intervention could still be influenced by the systemwide changes in the service environment.

A basic comparison group design or one of its extensions might have to be based on eligible youth rather than volunteers. This is because it would not be possible to identify volunteers in the comparison site, since there would be no intervention for which youth could volunteer. Alternatively, it might be possible to base the evaluation on intervention participants in the

treatment site and a matched subset of youth in the comparison site who satisfy all of the intervention eligibility criteria except for residency in the service delivery area. A score for the propensity of eligible youth to enroll in the intervention, or another mechanism, could be used to match youth receiving services in the intervention site with similar youth drawn from among all eligibles in the comparison site. This approach would provide an increased power to detect possible impacts with a given sample size, as virtually all sample members in the intervention site (volunteers) would receive some PROMISE services, whereas possibly only a small fraction of them would if the sample included the broader group of eligible youth. Furthermore, a survey would be a more realistic source of outcomes data for the impact analysis under a matched comparison group design (as opposed to a basic comparison group design), because the data collection costs would be lower, given the smaller number of sample members. However, this approach would carry a considerable limitation: the participant group and matched comparison group might end up being different, on average, for characteristics not used in the matching process. If any of those characteristics, such as motivation, were ones that influence the outcomes of interest, then differences in outcomes between the two groups might be incorrectly attributed to the intervention, when in fact they were due to preexisting differences between the groups.

The selection of awardees for PROMISE interventions could give formal consideration to the likely strength of the research findings from the evaluations of those interventions. Central to such consideration would be the types of evaluation designs that could be applied to the proposed interventions and the willingness and capacity of the applicants to support the execution of those designs. For proposed interventions that are conducive to experimental evaluation designs, selection criteria might include the ability and willingness of the applicants to be involved in a random assignment demonstration evaluation, as well as the counterfactual environments that control group members would face. For proposed interventions that could only be evaluated with nonexperimental designs, more weight might be given to those that could have stronger evaluation designs, such as those not involving statewide systems change (which might require a comparison group from another state, potentially reducing confidence that the impact estimates would reflect the true effects of the intervention).

C. What Data Collection Challenges Is the Evaluation Likely to Face?

When asked whether outcome measures for the evaluation should be permitted to vary across the research sites, the TAP members suggested that certain outcome measures could be consistent across the sites while others could be site specific. They supported the adoption of a consistent principal outcome measure for the impact study—the amount of SSI benefits (or SSI and SSDI benefits combined) received by youth (and possibly by their families). Many of the other outcome measures could also be consistent across the sites, because even though the interventions might differ, they are likely to have many goals in common. Furthermore, having standardized outcome measures could reduce evaluation costs and facilitate cross-site comparisons of findings. However, the proposed interventions are likely to reflect local service and economic environments, target populations, and assumptions regarding how best to promote the independence of youth with disabilities. Such variation will likely necessitate some site-specific outcome measures, particularly pertaining to the intermediate term. Capturing these might require follow-up (postenrollment) surveys, because administrative files tend to provide data on a limited number of standardized measures.

Potential data collection challenges with surveys of youth and families include capturing information on services and outcomes that are likely to vary across sites, delays in obtaining data,

and the need to hold down data collection costs, given the evaluation's limited budget. The projected funds available for the evaluation (approximately \$5 million if, as was mentioned during the TAP meeting, half of SSA's possible \$10 million in funding for PROMISE were allocated for incentive payments to awardees) might support only one formal survey of youth and their families. Therefore, rather than the evaluator conducting a baseline survey, project staff could administer a short information form along with a contact sheet and informed consent form as part of the process for enrolling youth in the evaluation (should an experimental design be used).¹⁶ The information form could be foregone if adequate baseline data were available from administrative files. A follow-up survey could be critical for measuring certain intermediate outcomes for the impact study, especially those that are site specific. Careful consideration should be given to the timing of this survey. On the one hand, the survey could capture services received, most of which are likely to occur early in the postenrollment period. Waiting too long could exacerbate problems of respondent recall of services. On the other hand, this survey may also be the principal source of data on intermediate outcomes. Conducting it too soon after enrollment could reduce the evaluation's capacity to estimate impacts on these outcomes, because the interventions would not have had sufficient time to serve participating youth and their families and make a difference in their lives. Surveys typically entail lags in collecting data, which can limit applications of the data. For example, basing intermediate incentive payments on outcomes measured through surveys could substantially delay the issuing of those payments. A final point regarding a follow-up survey is that it may be the only source of comparable data for treatment and control youth on services received. The TAP members underscored the importance of having such comparable data. However, as noted in the previous section, survey data collection costs could be prohibitively high under certain nonexperimental evaluation designs that target all eligible youth.

Potential data collection challenges with administrative information systems include accessing the data in a timely manner, ensuring data security, and collecting consistent data across sites. Administrative systems are often good and inexpensive sources of data on outcomes, as they provide information on both treatment and control group (or comparison group) members and they can track outcomes in the long term. They can also be good sources of baseline control variables for regression impact models, especially when the outcome measures contain postenrollment values of the same variables. Furthermore, some administrative systems, such as those maintained by school districts and state VR agencies, provide data on services. However, there can be long lags between the collection of administrative data and when the data are made available to researchers (for example, there is a lag of approximately 15 months between the end of a calendar year and the release of SSA data on annual earnings to researchers). These lags could reduce the value of such data in the estimation of impacts and the determination of intermediate incentive payments. Another factor that could restrict the use of administrative data in the PROMISE evaluation is that there are often issues concerning data confidentiality and/or data security in obtaining access to those data for research purposes. Finally, for state and local administrative databases, the available data elements might differ across the evaluation sites, leading to research results that are not comparable. Also, especially for these databases, the funding agencies and the evaluator may want to weigh the cost of obtaining access to and processing the administrative data against the cost of collecting similar measures through a follow-up survey.

¹⁶ If a comparison group design were used, then it would probably not be feasible for project staff to administer a baseline information form in the comparison sites. This task would either fall upon the evaluator or it would be necessary to rely exclusively on baseline data from administrative files.

The family as the unit of analysis represents an overarching data collection concern for the PROMISE evaluation because of issues surrounding informed consent and data access. Family involvement with respect to both services and outcomes is a unique aspect of PROMISE; the interventions could involve organizations that provide adult services. As such, the evaluation may include an assessment of the services and outcomes at both the family and youth levels. Involving family members in the evaluation could significantly complicate data collection efforts. For example, such involvement might necessitate including survey questions to assess the earnings, employment, and health of all family members. This would add to the respondent burden, make it more challenging to obtain high response rates, and increase the cost of the survey. An additional concern for the funding agencies and the evaluator is whether informed consent will be required of parents as well as youth during the evaluation enrollment process. The consent of parents and other household members may be necessary for the evaluation to obtain access to certain administrative data, such as Unemployment Insurance earnings records and state education data. If such consent were needed, it could dramatically increase the burden on PROMISE projects of enrolling youth (and their families) in a random assignment evaluation and substantially reduce the enrollment rates ultimately achieved.

V. EXTERNAL VALIDITY

External validity issues for the PROMISE demonstration and evaluation concern the specification of the target population, the state and local characteristics of demonstration sites that could influence the evaluation findings, and the needs of SSA and its federal partners in their use of the findings. Ultimately, the funding agencies would like to know the implications of implementing PROMISE, including SSI-based incentive payments, on a broader scale as a national program. Even if the evaluation findings are internally valid (that is, the impact estimates are not biased for the effects of PROMISE on those who receive demonstration services), there are a variety of reasons why they might not apply, on average, to all youth who are SSI beneficiaries. The selection of awardees by the funding agencies, the selection of subjects by the awardees, the nature of the intervention at each site, the roles of collaborating agencies and service providers, and many other features of the PROMISE demonstrations will affect the appropriate uses of the evaluation findings by the funding agencies and others involved in supporting the efforts of youth with disabilities and their families to achieve independence.

The TAP made the following recommendations on the external validity of findings from the PROMISE evaluation:

- Most TAP members agreed that demonstration projects designed to serve the full PROMISE target population—youth ages 17 and under receiving SSI benefits and their families—would provide the maximum generalizability for the evaluation findings and would maximize the number of locations that could be considered for demonstration projects.
- As an alternative to serving the full PROMISE target population, some TAP members suggested that allowing applicants to propose serving segments of that population might increase the probability that high-quality projects would be proposed—that is, projects that, if implemented, would be more likely to achieve the intended outcomes.
- Regardless of how the target populations for specific PROMISE projects are defined, the presence of adequate pools of eligible youth in the catchment areas for the projects will be critical to the success of both the projects and the evaluation.
- An area's existing resources and characteristics could affect the ability of youth and families to increase their independence, but the TAP members were uncertain whether or how these should be factored into award decisions for PROMISE projects. However, they agreed that the RFA process should identify applicants with the vision and capability to move their states or localities to substantially better services for SSI youth and their families, as opposed to making marginal improvements to existing services.
- The evaluation should inform SSA and its partner agencies as to whether projects of the size and complexity of those envisioned under PROMISE can be cost-effective.
- The evaluation should document how policy changes at the federal and state levels affect the interactions of state and local agencies and providers in delivering services to SSI youth and their families.

- The evaluation should provide information to guide local, state, and national policies and practices for serving youth with disabilities and their families, particularly information on the effectiveness of specific intervention approaches.

Based on the above recommendations by the TAP and our consideration of what is currently known about the design for PROMISE, we make the following recommendations:

- As noted by one TAP member, if intermediate-term reductions in dependency by youth beneficiaries as they transition into adulthood are a priority for PROMISE, then it would be advantageous for one or more of the projects to target youth at the upper end of the eligible age range.
- The RFA for PROMISE projects should require each applicant to identify and justify the specific population it proposes to recruit and serve: the full PROMISE target population or a subset of it.
- While some TAP members suggested that a strong record of leadership by awardees and demonstrated collaboration with other agencies and with service providers might be important state and local determinants of the successful implementation of PROMISE, these might not be reliable predictors of project impacts.

We begin this chapter with a discussion of the target population for PROMISE. The next section considers features of the state and local environments that could influence a project's success. The final section reviews the essential information that SSA and its partner agencies ultimately need from the PROMISE evaluation.

A. Who Should Participate in PROMISE?

The target population for PROMISE has been defined broadly to consist of minors (youth ages 17 and under) receiving SSI and their families. No lower age bound has been specified. From a policy perspective, the main advantage of this population is that there might be greater potential to address barriers to independence for children and adolescents than for adults on SSI. Furthermore, the potential savings to SSA in terms of lifetime benefits avoided is especially large for this population. Finally, although there is a substantial and growing literature on youth with disabilities, there is little in the way of conclusive findings on what works to promote independence for SSI youth. However, this population also poses risks for the demonstration. While there is little solid evidence on how to facilitate independence for SSI youth, even less is known about how to effectively work with their families. Also, the young age of the target population suggests that the principal way that the demonstration could have an impact on the key outcome—receipt of disability benefits—in the short and intermediate terms would be through improving the employment of other family members, rather than that of the youth. Most of the target youth are likely to remain in school during much of the demonstration; consequently, determination of the impact of the demonstration on their benefits as they transition into adulthood will have to be based on long-term analysis of administrative data, to be conducted after the demonstration is over. Interventions that include youth less than 14 years old might need an especially long follow-up period to observe such outcomes. If the reduction of future adult SSI benefits for current youth beneficiaries is a critical outcome for PROMISE, then, as recommended by one TAP member, the projects might target older youth (ages 14 to 17) with interventions that include specific employment services for youth in the upper end of that age range.

Abstracting from concerns about time lags in the manifestation of impacts on benefits, most TAP members agreed that projects designed to serve the full PROMISE target population—minors receiving SSI benefits and their families—would provide the maximum generalizability for the evaluation findings and would maximize the number of locations that could be considered for demonstration projects. From the funders’ perspectives, it is likely that no single subgroup of youth is more or less important to serve, so their interest is in improving outcomes for the entire target population. This approach would yield evaluation findings with the broadest generalizability. However, other features of the demonstration might limit the generalizability of the findings, such as the anticipated small number of projects and the specificity of their geographic settings and services. Given this, a more compelling reason for the projects to serve the full target population is that this strategy would likely maximize the projects’ ability to recruit and enroll enough youth in the evaluation for the findings to have adequate statistical reliability. We provide more details about this issue in the next section.

As an alternative to requiring projects to serve the full target population, some TAP members suggested allowing applicants to propose serving segments of the population. This might increase the probability that high-quality projects would be proposed—that is, projects that, if implemented, would be more likely to achieve the intended outcomes. Restricting eligibility, such as by age or disabling condition, could allow for finer targeting of project services. For example, the City University of New York’s YTD project focused on an age-based segment of the broader YTD target population.¹⁷ This enabled the managers and staff of the project to design and deliver services that they considered to be most appropriate for that age group. This project was well implemented and produced early impacts on paid employment. In a similar fashion, permitting applicants for PROMISE projects to propose serving segments of the broader target population might lead to more creative and effective interventions. Furthermore, some otherwise well-qualified potential applicants might have experience working only with specific subpopulations. Such restrictions, though, could have downsides in terms of greater recruitment challenges and reduced external validity of evaluation findings. The latter could be mitigated by projects specifying well-defined, quantitative eligibility criteria.

One approach to the issue of who should participate in PROMISE that was favored by some TAP members would be to direct applicants to identify and justify in their proposals the specific populations they would recruit and serve—the full PROMISE target population or subsets of it. The funding agencies could then assess the appropriateness of the proposed interventions for the specified populations. We note that, as an extension of this approach, the RFA could encourage applicants to specify any proposed subpopulations on the basis of data in SSA files. This would enable the evaluator to systematically assess the implications of the inclusionary criteria for the adequacy of the size of the subpopulation in the project’s catchment area to meet the statistical precision requirements of the evaluation. This could also help efficiently identify eligible youth (as opposed to, for example, screening youth by using a short telephone script).

¹⁷ Fraker, Thomas, Alison Black, Joseph Broadus, Arif Mamun, Michelle Manno, John Martinez, Reanin McRoberts, Anu Rangarajan, and Debbie Reed. “The Social Security Administration’s Youth Transition Demonstration Projects: Interim Report on the City University of New York’s Project.” Washington, DC: Mathematica Policy Research, April 2011. Available at <http://www.ssa.gov/disabilityresearch/documents/CUNY%20Final%20Interim%20Report%204-5-11.pdf>. Accessed December 21, 2011.

B. What Features of the State and Local Environments Will Help Projects Achieve Their Specific Objectives and Reduce SSI Dependency?

The TAP recognized that adequate pools of eligible SSI youth in the catchment areas for the PROMISE projects will be critical to the success of both the projects and the evaluation. The YTD experience may be enlightening in this regard. The six random assignment YTD projects, assisted by Mathematica, recruited extremely aggressively and achieved rates of enrollment in the evaluation ranging from 15 to 30 percent. If each PROMISE site required 800 evaluation enrollees (some of whom would comprise a treatment group and others a control or comparison group) to achieve acceptable statistical power for impact estimates, then a 15 percent enrollment rate would imply a minimum pool of 5,300 eligible youth who meet the selection requirements during the recruitment time frame.¹⁸ Recruitment for PROMISE would be more challenging than for YTD if parental services were a key component of the interventions and/or parental informed consent for release of their own administrative data were required, both of which are likely given current plans for PROMISE. This would suggest the need for an even larger minimum pool of eligible youth.

While some TAP members suggested that a strong record of leadership by awardees and demonstrated collaboration with other agencies and with service providers might be important state and local determinants of the successful implementation of PROMISE, these might not be reliable predictors of project impacts. Awardees that have provided leadership on youth and disability issues in their states and have spearheaded other projects involving multiple agencies could be well positioned to successfully implement PROMISE. However, successful implementation would not ensure significant impacts. To produce impacts, an intervention must be distinct from the counterfactual service environment, and the distinction must be such that it effectively promotes improved short-, intermediate-, and long-term outcomes. Another perspective on this issue is that although demonstrated success at cross-agency collaboration and program innovation might be predictive of successful PROMISE implementation, those attributes could have already resulted in a strong counterfactual environment, thus limiting the potential impacts of the intervention. Therefore, rather than summarily rejecting applications from agencies with weak histories of cross-agency cooperation and innovation, the funders should evaluate proposals based on the merits of the proposed interventions.

An area's resources and characteristics could affect the ability of youth and families to increase their independence, but the TAP members were uncertain whether or how these should be factored into award decisions for PROMISE projects. Important resources and characteristics cited by TAP members include economic factors (such as the unemployment rate, per capita income, poverty rate, industry and occupation mix, and the cost of living), state factors (such as state funding for services and transportation infrastructure), and socioeconomic factors (such as the number of youth with disabilities and local attitudes about work). The TAP members and meeting attendees suggested that the funding agencies should not necessarily favor applicants from resource-poor areas in the hope that they can improve outcomes for their SSI youth and families to national average levels. They also suggested that the funding agencies not discount applicants from resource-rich areas that are likely

¹⁸ Absent substantial support for PROMISE project enrollment activities, such as the support Mathematica provided for the YTD projects, a prudent approach for SSA and its federal partner agencies would be to assume a lower range of enrollment rates than was achieved on the YTD evaluation. This would imply a need for a larger pool of eligible youth than the 5,300 cited in the text.

to have strong counterfactual services (that is, the presence of a wide array of services for youth with disabilities before the implementation of PROMISE). Rather, selecting applicants with the best chances of changing the way state and local agencies and service providers operate to promote success for youth with disabilities would be most consistent with the vision for PROMISE by the funding agencies. The TAP members broadly agreed that the RFA process should identify applicants with the vision and capability to move their states and localities to substantially better services for SSI youth and their families, regardless of the current resources and characteristics, as opposed to making marginal improvements to existing services.

C. What Information Should the Evaluation Provide to SSA and Its Partner Agencies?

The TAP members, along with the agency staff who participated in the TAP meeting, identified three key domains about which PROMISE should provide information.

First, the evaluation should inform SSA and its partner agencies as to whether projects of the size and complexity of those envisioned under PROMISE can be cost-effective. The primary indicator of success will not be whether the projects result in long-term SSI benefit savings, but whether those savings are sufficient to fund successful youth interventions. If the evaluation were to show such success, then PROMISE could be a compelling model for funding programs for youth with disabilities and their families. For instance, the evaluation might provide evidence that cost-effective programs for SSI youth could be funded through anticipated collective benefit savings over a 7- to 10-year horizon. This would be a very different model from those currently in place, under which organizations are paid either based on the services they provided, without regard to outcomes, or based on their achieved success with individual beneficiaries over more limited horizons (the TTW outcomes model).

Second, the evaluation should document how PROMISE-related policy changes at the federal and state levels affect the interactions of state and local agencies and providers in delivering services to SSI youth and their families. The four federal agencies that will be actively involved in PROMISE might have the potential to develop new waivers or expand existing ones to promote cooperation among state and local agencies in serving SSI youth. The evaluation should document successes in cooperation and barriers to cooperation among agencies at the state level. It should also document the relationships among agencies and service providers at the local level, the areas in which they improved their interactions (such as referrals), and the lessons that they learned in working with each other. States and localities other than those participating in the PROMISE demonstration could use these findings to adapt and improve their own systems.

Third, the evaluation should provide information to guide future local, state, and national policies and practices for serving youth with disabilities and their families, particularly information on the effectiveness of specific intervention approaches. Although the evaluation will likely be unable to show impacts for specific component services within an intervention, it will be able to provide information on the overall effectiveness of the intervention, on the types and intensity of services delivered, and on which aspects of the intervention were implemented especially well or poorly. The evaluation should also document the extent to which fidelity to the proposed intervention design was achieved and components that project staff believed were more or less successful. An important set of lessons from the evaluation will pertain to incentive payments—whether they were implemented as planned, their number and amount, how they were used by recipients to promote services among youth, and administrative difficulties in their application.

Finally, the PROMISE evaluation should identify strategies that projects and individual staff members used to engage youth and families in services to promote independence—strategies that could be useful to others working in the field.

VI. OTHER CONSIDERATIONS FOR PROMISE

The TAP's assignment was to provide guidance to SSA and its partner agencies on incentives in PROMISE and on the evaluation design. However, in the process of advising on these matters, the TAP members provided comments on areas outside the panel's scope but nevertheless pertinent to PROMISE. These comments—on the role of the family in PROMISE, aspects of the interventions other than the incentives, the age-18 redetermination for SSI youth, and incentive payments provided by ED—are summarized here for consideration by SSA and its partner agencies.

A. The Family Is a Prominent Aspect of PROMISE, but the Role of the Family in the Intervention Has Not Been Fully Developed and Is Potentially Problematic

The TAP agreed with the funding agencies that the PROMISE intervention should involve parents and potentially other family members, given the importance of family on outcomes for children well into adulthood and the potential for intergenerational transmission of SSI dependency. However, some TAP members expressed concern that improving parental and family outcomes could be more challenging than improving youth outcomes alone. Youth who receive SSI benefits not only have disabilities, they also have families with incomes and assets that are sufficiently low that the youth can qualify for benefits. The parents of these youth could have characteristics—such as limited educations, disabling conditions, and their own SSI receipt—that tend to suppress their paid employment. PROMISE interventions might include services for parents to promote their employment and improve their incomes, with the expectation that this would help them model lives of independence for their children (leading to less long-term dependency by the children) and reduce the benefits received by the children in the short term (due to higher family incomes). This raises the question of whether the involvement of parents in PROMISE should necessarily entail services to improve parental outcomes, as opposed to services to bolster parental expectations and support for their children. The answer to this question might come only when ED decides on the awardees and their interventions.

PROMISE interventions providing considerable services to improve parental outcomes could result in smaller family SSI benefits, but that might leave the interventions with fewer resources to provide services directly to youth to improve their human capital and their prospects for long-term independence. The positive-parental-role-model effect might dominate the reduced-services-for-youth effect, resulting in less long-term dependency among youth than would be the case with an equivalently funded intervention that focused exclusively on youth outcomes. However, the opposite might also occur, resulting in an intervention that would be less effective at reducing youth dependency than one more narrowly focused on youth outcomes. Some TAP members expressed concern that interventions with a focus on parental outcomes might not be efficient in reducing long-term dependency among youth. A strength of the PROMISE demonstration is its potential to test the effectiveness of fundamentally different strategies for promoting the independence of youth. As with other elements of their planned interventions, applicants proposing significant services to improve parental outcomes could provide arguments based on existing research and their own experiences for how those services would ultimately reduce youth dependency. The funding agencies could then consider the strength of those arguments in making award decisions.

Youth and their families may be at financial risk in this intervention, given their dependence on SSI and other benefits, so TAP members suggested safeguards for protecting them against the loss of those benefits. Many families depend on monthly SSI benefits, and their concerns about the

possible loss of those benefits might be great enough that some of them would refuse to participate in an intervention that put their benefits at risk. Families might need initial counseling to explain how benefits operate, particularly with respect to earned income. SSA waivers of certain rules regarding SSI benefits could mitigate such concerns by ensuring continued benefit receipt, a simplified process for returning to the rolls, or expanding the earned income exclusion. Such waivers might serve more as protective elements to encourage participation in an intervention, rather than as incentives to promote independence.¹⁹ If families are central to PROMISE, then additional issues may arise regarding privacy, the option for family members not to participate, and confidentiality. These issues could increase the challenges of implementing PROMISE projects and evaluating their impacts.

B. TAP Recommendations About Aspects of the PROMISE Interventions

While TAP members viewed certain services as essential for any PROMISE project, they also agreed that applicants for projects should be encouraged to be creative in their proposed approaches. Overprescribing the facets of the intervention—whether the services to be provided, the population to be served, or the agencies to be involved—would risk stifling innovation on the part of applicants, or worse, having applicants incorporate components into their interventions that they do not have the capacity or interest to execute well. Several TAP members suggested that a possible approach might be to specify the desired outcomes and ask applicants to design interventions to achieve those outcomes. However, ED might want to encourage the inclusion of specific components in the interventions that it regards as critical. For instance, benefits counseling is widely acknowledged as a necessary component of services to promote progress toward independence by youth beneficiaries. ED could incorporate such components into the RFA by identifying them as promising approaches for applicant consideration.

Additional levers for guiding the interventions discussed by the TAP include setting forth requirements in the RFA, selecting proposals that the funders believe would result in the best outcomes, and making continued funding contingent on successful annual reviews. Rather than dictating a specific approach or menu of service items, SSA and its partner agencies could adapt these implementation activities to help shape the interventions and increase the likelihood of positive results, while providing the applicants/awardees the flexibility to design and implement interventions to meet the fundamental PROMISE objectives. Using cooperative agreements rather than grants as the funding mechanism would provide important ongoing leverage for ED to influence the interventions. In addition, close monitoring of services provided and outcomes attained could help the funders and awardees identify deviations from intervention designs and opportunities for improvements to the interventions as a formative implementation process. These levers would require more involvement by the funding agencies in the interventions than simply making the awards and anticipating the impact findings, but they could potentially guide the sites to better outcomes.

Some TAP members noted that trying to do too much with the interventions—having diverse goals and involving too many agencies and programs—could be problematic. PROMISE risks being

¹⁹ YTD included waivers that were designed to promote independence in the long run (such as allowing youth to retain more of their benefits with increased earnings than current SSI rules allow); however, the YTD evaluation has yet to produce findings that such waivers can ultimately lead to reduced dependency. Anecdotal evidence does suggest that the existence of the waivers facilitated the enrollment of youth in the YTD evaluation.

too complex an initiative, with applicants having to propose interventions that address youth needs, family independence, and inefficiencies in local and state service systems, any one of which alone would be difficult to tackle. This could result in diffused services and failure to achieve key outcomes. Furthermore, it might be hard to disentangle what worked from what did not and difficult to determine how the interventions did or did not lead to the intended outcomes.

Suggesting cross-agency partnerships in the RFA might point applicants in a useful direction, but mandating specific partnerships might be counterproductive. TAP members pointed out that requiring the involvement of specific agencies could result in consortia of agencies that have no or poor work histories, that are competing for the same scarce resources, and/or whose missions are not well aligned to achieve PROMISE outcomes. Partnerships work best when the involved agencies agree on project approaches and objectives, and when those are congruent with the agencies' own missions. Requiring potentially misaligned agencies to work together would require the awardees to invest time and resources in building effective partnerships, with no assurance of success, and could detract from implementing the interventions. The more measured approach of specifying in the RFA the outcomes expected through PROMISE and suggesting that partnerships could be helpful in achieving them would enable applicants to organically identify the linkages they would need with other agencies to effectively implement their proposed interventions.

C. The Effect of the Age-18 Redetermination Process Should Be Fully Considered for PROMISE

At age 18, all SSI youth must go through a redetermination process to assess their status relative to adult disability criteria. A substantial proportion of youth lose their benefits through this process, which the designs for estimating PROMISE impacts and calculating incentive payments should reflect. TAP members noted that those designs should avoid attributing reductions in benefits due to the age-18 redetermination to the PROMISE interventions. The selection of a control group would provide a mechanism for appropriately accounting for this change in benefits. The critical benefit reduction goal of PROMISE would be promoted if the projects were to continue serving youth who lose their benefits during the redetermination process, as those services might reduce the likelihood that they would reenter the disability rolls at a later date. To address the TAP's concerns, SSA and ED will need to consider whether or not PROMISE services should end at age 18, upon termination of eligibility for SSI, or in other situations that would result in participants no longer meeting the definition of the target population.

If the PROMISE evaluation is based on youth who volunteer for the opportunity to receive services, as discussed in Section IV.B, then the evaluation's take-up rate might depend on youths' subjective assessments of their probabilities of receiving negative age-18 redetermination decisions. Some youth and their parents who fear a negative redetermination decision might avoid enrolling in a project intended to improve human capital and reduce SSI dependence. If this were the case, youth who enrolled in the evaluation would be less likely to have negative redetermination decisions than those who did not enroll. Alternatively, it is possible that youth who anticipate the loss of benefits at age 18 would embrace the concept of PROMISE and be more likely to enroll in the evaluation. These scenarios illustrate the need to be aware of the potential influence of the age-18 redetermination process on the self-selection of youth into the evaluation.

D. Incorporating Incentive Payments by ED in PROMISE

Another issue that was beyond the scope of the TAP but that the panel members commented on was the potential for ED to provide incentive payments to awardees, providers, and/or families based on more-immediate outcomes achieved by individual participants. The mechanism for incentive payments provided by ED would be different from that envisioned for the incentive payments to be provided by SSA, in that the payments would not necessarily be based on reductions in SSI benefits nor would they be based on group-level (aggregate) outcomes. However, the mechanism could be similar in many ways to the one SSA uses to make incentive payments under its TTW program, in that a PROMISE awardee or its designee could document a participant's attainment of a specific goal and request the associated incentive payment. The TTW experience has shown that these payments have to be paid quickly and efficiently, with a minimum of administrative burden on the claimants, to be effective incentives.²⁰

As with the intermediate incentive payments by SSA, discussed in Chapter II, extant research findings of positive relationships between the underlying outcome measures and the goals associated with them and long-term reductions in SSI dependency would have to support payments by ED. The TAP members recommended that these payments not be based on the completion of processes (for example, interagency coordination of service plans for family members) or on outcomes for which there is no research basis for expecting that they would lead to reductions in SSI benefits. Incentive payments by ED could decrease the likelihood that project resources would be spent on activities that are unlikely to lead to lower disability benefits in the long term.

ED could allocate a proportion of its anticipated \$30 million in funding for PROMISE specifically for its incentive payments. For example, ED could allocate \$10 million each to two awardees (an average of \$2 million per awardee per year for five years). It could then use the remaining \$10 million for incentive payments. This structure could promote intensive services for participants soon after they enroll in a PROMISE project, as the awardees would strive to help them achieve the individual outcome targets that would trigger the incentive payments. However, it would carry two risks: (1) a risk of lower service provision at the end of the project if the awardees were to fail to achieve the individual-level targets and consequently have fewer funds than anticipated to serve current and future participants and (2) a risk of unproductive distortion of behaviors if the awardees and/or providers were to seek to achieve the targets in ways that meet the incentive payment criteria but do not ultimately lead to lower benefits.

As mentioned in Chapter II, TAP members were divided as to who should receive incentive payments from ED: awardees, service providers, or youth and their families. Regardless of the recipients, the payments would be based on goals achieved by specific individual participants, as follows:

- Financial incentives for *awardees* would encourage them to develop creative approaches to their partnerships with state and local agencies and service providers, perhaps

²⁰ For an assessment of provider experiences with TTW and its outcome payments, see Altshuler, Norma, Sarah Prenovitz, Bonnie O'Day, and Gina Livermore. "Provider Experiences Under the Revised Ticket to Work Regulations." Washington, DC: Center for Studying Disability Policy, September 2011. Available at http://www.mathematica-mpr.com/publications/PDFs/disability/TTW_process.pdf. Accessed January 3, 2012.

incorporating braided funding and waivers to pool resources. Awardees could choose to pass along some or all of the incentive payments to providers or youth and families, even creating their own incentive payment system within their PROMISE projects. However, there is a risk that providing incentive payments to awardees might result in less funding for providers to deliver actual services, depending on a number of factors, including awardee administrative costs and the influence of state and local politics.

- The advantage of offering incentive payments to *service providers* rather than awardees is that providers, being closer to the PROMISE participants, might be better positioned to design and deliver services to individual participants so as to maximize their chances of achieving the individually based targets. However, the payments would have to be structured to be attractive to providers (that is, with targets that are achievable and amounts that are large enough to constitute meaningful financial incentives). The potential disadvantage of ED offering incentive payments to providers is that the incentives might lead providers to focus intensely on those services that most directly lead to payments to the exclusion of other productive services.
- *Youth and families* might also be responsive to financial incentives, working toward goals that would result in immediate payments (for example, a payment for graduating from high school). However, recent research on the effectiveness of such incentives is not encouraging.²¹ Furthermore, some TAP members advised that a PROMISE project might want to avoid creating external motivations based on programmatic incentives as opposed to helping participants develop their own internal motivations. The direct payment of incentives to participants also could be detrimental to service provision, as it would reduce funding and incentives for providers to work with youth and families to achieve independence. The decision to offer incentive payments to participants could be left to awardees and providers; if they believe that participants need such incentives to achieve project goals, then they could finance the incentives out of their own PROMISE funding.

Outcomes on which to base incentive payments from ED could include a mix of education and employment measures that reflect participant and project characteristics, with applicants encouraged to identify measures they feel are important indicators of progress toward independence. As discussed in Chapter II for intermediate outcomes, the TAP members agreed that any objectives proposed as the basis for incentive payments should be ones that have been documented in the literature to result in long-term reductions in SSI dependency, and those objectives should be included in applicants' logic models. Measures that are the basis for incentive payments should be appropriate for the PROMISE target population—SSI recipients age 17 and under and their family members. Employment-related measures might not be appropriate for younger youth. For older youth, appropriate employment-related measures might include preparation of a resume,

²¹ An emerging area of policy interest is providing youth and parents with financial incentives based on school performance, such as Opportunity-NYC, which provided cash incentives to families for their children's improved school performance to break the cycle of poverty. The initial quantitative results from that evaluation have been mixed on the project's educational and employment outcomes. See Riccio, James, Nadine Dechausay, David Greenberg, Cynthia Miller, Zawadi Rucks, and Nandita Verma. "Toward Reduced Poverty Across Generations: Early Findings from New York City's Conditional Cash Transfer Program." New York: MDRC, March 2010. Available at <http://www.mdrc.org/publications/549/full.pdf>. Accessed January 4, 2012.

participation in work site tours, completion of an internship, and part-time or summer employment, whereas those for parents might include maintaining employment for specified durations, increasing the number of hours worked, and obtaining employment in certain occupations or with fringe benefits. Disproportionate emphasis on short-term employment objectives for youth could hinder their achievement of education objectives, to the detriment of their long-term independence. Education-related measures that might be appropriate for youth in the upper age range of the target population include graduation from high school, enrollment in vocational training and postsecondary education programs, and receipt of a higher education degree. These and other measures that are potential triggers for ED incentive payments might have to vary across PROMISE projects, depending on the nature of the interventions and the populations served.

In contrast to the designs for SSA incentive payments presented in Chapter II, it may be preferable to avoid linking payments from ED to counterfactual outcomes. Payments from ED that are timely and flexible would allow resources to more rapidly accrue to awardees, providers, and families, and thus provide them with stronger incentives to achieve outcomes associated with reduced dependence on benefits. Linking those payments to a counterfactual would be administratively difficult, requiring surveys or administrative records to obtain information about a comparison group. Such approaches would necessarily entail lags in gathering and analyzing the data. By comparison, not linking payments to a counterfactual would allow awardees to use an MIS to monitor and report the designated individually based outcomes for PROMISE participants, and ED could implement an administrative system to make payments based on the reports. A verification system could be developed to confirm that outcomes reported for payment were actually achieved.

Establishing separate payment mechanisms—from SSA and ED to awardees, providers, and individuals—could enable each of the federal agencies to promote the specific outcomes that it would like to achieve under PROMISE, but those mechanisms must be designed to avoid providing inconsistent incentives to the key actors. There is a risk that SSA or ED might provide incentives for a set of outcomes that would decrease the likelihood of achieving another set of outcomes for which the other agency is providing incentives. There is also a risk that awardees, providers, and individuals could become confused by multiple incentives and could lose track of some outcomes in pursuit of others, resulting in poorer long-term results than might be achieved with fewer incentives or a more integrated set of incentives.

MATHEMATICA **Policy Research**

www.mathematica-mpr.com

Improving public well-being by conducting high-quality, objective research and surveys

Princeton, NJ ■ Ann Arbor, MI ■ Cambridge, MA ■ Chicago, IL ■ Oakland, CA ■ Washington, DC

Mathematica® is a registered trademark of Mathematica Policy Research