



# Outcomes Measurement and Pricing

Pay for Success Early Childhood Education  
Toolkit Report #3



## What Is Pay for Success?

Pay for success (PFS) offers an alternative approach to investing in the future, including early childhood education. This innovative financing mechanism shifts financial risk from a traditional funder—usually government—to a new investor, who provides up-front capital to scale an evidence-based social program to improve outcomes for a vulnerable population. If an independent evaluation shows that the program achieved agreed-upon outcomes, then the investment is repaid by the traditional funder. If not, the investor takes the loss.

For more information on pay for success, please visit [pfs.urban.org](http://pfs.urban.org).



## About the Early Childhood Education Toolkit

This toolkit is designed to guide jurisdictions and their partners through the core elements of a PFS project in early childhood education: the existing evidence for early childhood interventions, the role of data, the measurement and pricing of outcomes, program funding and financing, implementation, and evaluation design. The toolkit includes a series of helpful features, including checklists, charts, and questions for consideration, to help direct and clarify thinking around the feasibility of pay for success to scale what works in early childhood education. Together, these briefs can help jurisdictions decide if pay for success is the right approach for them—and if so, how to get started.



## Acknowledgments

The Pay for Success Early Childhood Education Toolkit was conceived by a working group led by the Urban Institute and its partners at the fifth annual Clinton Global Initiative America meeting in 2015 in Denver, Colorado.

The Urban Institute is incredibly grateful to our working group partners, who lent their time and talents to the development, research, and writing of this toolkit: Accenture, Bank of America, Enterprise Community Partners, Goldman Sachs, the Institute for Child Success, Nonprofit Finance Fund, Reinvestment Fund, Salt Lake County, Social Finance US, and Third Sector Capital Partners.

Support for the Pay for Success Initiative at the Urban Institute is provided by the Laura and John Arnold Foundation. We thank our funders, who make it possible for Urban to advance its mission.

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at [www.urban.org/support](http://www.urban.org/support).

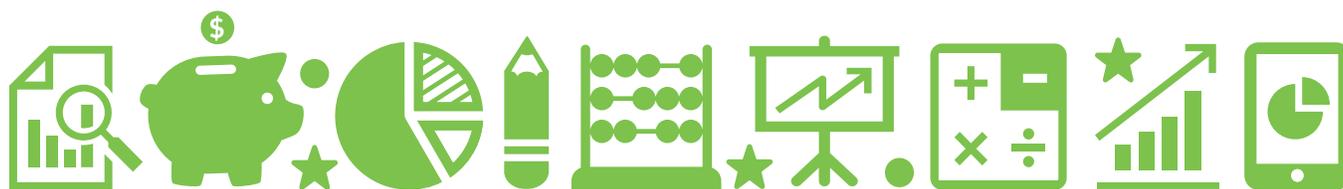


## Outcomes Measurement and Pricing

For this outcomes measurement and pricing report, we are particularly grateful to Anna Fogel, Casey Littlefield, and Lara Metcalf (Social Finance US); Emily Fabiaschi and Caroline Whistler (Third Sector Capital Partners); and Ben Nichols (Enterprise Community Partners).

Setting and pricing outcomes are important steps of every pay for success (PFS) project because they allow PFS partners to quantify the benefits of successful programs. Outcome measurement and pricing are integral to planning for repayments to funders, and they help stakeholders evaluate whether the total expected repayment amount and schedule are acceptable to both funders and the agencies that issue outcome payments. What is acceptable to funders will vary; most will want to recoup their principal and receive some level of return, but some philanthropic funders do not require returns. Likewise, what is acceptable to governments or other outcome payers will vary from project to project; in general, they will want the social and economic value of the outcomes to exceed the total repayment amount.

In the first part of this report, we describe the types of early childhood education (ECE) outcomes that might be used to establish success payments. In the second and third parts, we describe the steps and key questions that stakeholders might ask when selecting and pricing outcomes. Specifically, we provide guidance on how the partners in a PFS project might select an outcome, a measure, and a definition of success that is appropriate for the project. After that, we outline approaches to assigning dollar values to selected outcome measures. Examples are provided throughout, drawing primarily from the first two ECE PFS projects in the United States.



## What Outcomes Might Form the Basis for Payment in ECE PFS Projects?

Outcome measures play multiple roles in PFS projects and contracts. One critical role, and the focus of this report, is to establish a plan for repayments. For example, if the goal is to reduce the school readiness achievement gap, funders might be repaid a certain dollar amount for each child who is ready for school—as defined by agreed-upon metrics—as a result of the program. Outcome measures also have a role in broader evaluation and implementation monitoring, helping programs further build their evidence base and improve program quality.

## Outcomes Used in Past Evaluations

Evaluations of ECE interventions and cost-benefit analyses are two key sources of information for stakeholders working to value outcomes in a PFS project. A summary of findings from ECE evaluations and cost-benefit analyses is provided in toolkit report #1, *The State of the Science on Early Childhood Interventions*. Table 1 lists some of the outcomes researchers have used to quantify the benefits of ECE interventions.

**TABLE 1**

**Common Outcomes in ECE Evaluation or Cost-Benefit Research**

OUTCOME	TIMING
School and kindergarten readiness	Start of kindergarten
Developmental status, knowledge, skills, and abilities <sup>a</sup>	Prekindergarten–12th grade
On-time grade advancement	Kindergarten–12th grade
Need for special education services	Kindergarten–12th grade
High school completion	End of 12th grade
Overall educational attainment	Adulthood
Involvement with child welfare system	Kindergarten–adulthood
Involvement with criminal justice system	Childhood–adulthood
Teenage child bearing	Ages 13–19
Labor market participation and earnings	Adulthood

**Sources:** Urban Institute, 2016.

<sup>a</sup>These outcomes are measured through developmental assessments that cover one or more of the five domains of child development (approaches toward learning, cognition and general knowledge, language development, physical well-being and motor development, and social-emotional development) or through standardized achievement tests.



The economic benefits associated with the outcomes listed in table 1 accrue to different stakeholders—federal, state, and local governments; individual participants; and society as a whole—in increased earnings and tax contributions and avoided costs for school districts, government agencies, taxpayers funding those programs and agencies, and victims of crime. The outcomes listed in table 1 also cover a wide timespan, from the preschool years to adulthood. Building on research that follows children into adulthood, and thus provides evidence of the relationship between outcomes over time, some cost-benefit analyses of ECE use outcomes from earlier years as proxy measures for later outcomes. Similarly, PFS ECE projects may base success payments on the value of early outcomes chosen as proxy measures for teenage or adult outcomes.

### **Other Possible Outcomes for Repayment Structures**

The most feasible outcomes for repayment in a PFS project have a relatively short time horizon, yield a substantial total economic benefit, and represent clear cost savings or avoidance for the agencies that will be making the success payments. Given those considerations, reductions in K–12 special education and grade repetition are among the outcomes that might be easiest to implement in a PFS ECE project. However, ECE programs are generally more concerned with ensuring children’s overall healthy development or long-term success rather than specifically reducing the incidence of grade repetition or special education use.

Therefore, PFS projects may also want to consider one or more outcomes more directly related to children’s developmental progress or long-term success. In some cases, when evidence exists to associate measures of children’s developmental progress with longer-term outcomes that have concrete economic benefits, interim developmental outcomes can act as proxies for those long-term economic benefits. For example, Bartik, Gormley, and Adelstein (2012) estimated returns to investments in preschool programming in Tulsa, OK, by using third-grade achievement test scores to estimate adult labor market earnings. Other analysts, including those at the Washington State Institute for Public Policy (WSIPP 2015b), estimate returns from higher achievement test scores through their association with high school graduation and high school graduation’s association with higher labor market earnings. Beyond test scores, research indicates that grades, attendance, on-time grade promotion, and disciplinary actions can also predict or act as a proxy for high school graduation (Allensworth and Easton 2007; Balfanz, Herzog, and Mac Iver 2007; Rumberger 2011).<sup>1</sup>

---

<sup>1</sup>See appendix B for other resources related to monetizing ECE outcomes.

# What Should Be Considered when Selecting Outcomes and Metrics of Success?

Careful selection of appropriate outcomes—and metrics—for evaluating success is important and involves several related steps. The first major step is choosing a general outcome, such as third-grade literacy. The next step is selecting a specific measure, and measurement tool, such as student scores on the state’s standardized reading achievement test. Third is defining success, such as “scoring at or above the 25th percentile on the reading test,” and a way to evaluate that success. Outcomes, measures, and the strategy for defining success must all be clearly specified to establish when payments will be made in a PFS project.

## Steps and Key Questions for Selecting & Pricing Pay for Success Outcomes

### **SELECTING OUTCOME(S) FOR SUCCESS PAYMENTS**

- Will the intervention directly affect the target outcome?
- Is the outcome relevant and important to the community?
- Is the outcome measurable?
- Can the outcome be measured in a sufficiently short time frame?
- Can a dollar value be assigned to the outcome?
- Are expected effects from the intervention large enough to motivate a deal?

### **SELECTING THE INSTRUMENT TO MEASURE OUTCOME(S)**

- Does the instrument have a known track record?
- Is the instrument feasible to implement given budget constraints?
- Does the instrument have high validity and reliability?

### **DEFINING AND EVALUATING SUCCESS**

- What is the benchmark for success? How likely is success?
- Is success absolute (each child is achieving benchmark) or relative (each child who would not otherwise have achieved benchmark is now achieving it)?
- Are perverse incentives possible?

### **PRICING OUTCOMES FOR SUCCESS PAYMENTS**

- What is the value of intangible economic benefits?
- What is the value of tangible costs avoided?
- What is the value of other tangible economic benefits?
- How large a total repayment is the government or others willing or able to pay?
- Is the total expected repayment high enough to motivate funders?

## Features of a Good Outcome for a PFS Project



**The outcome is relevant to the community.** Each community faces different challenges and has different goals for improving civic and individual circumstances. Public support for a PFS project will be higher if success payments are made for outcomes valued in the community, either because of the potential for cost avoidance (e.g., costs for special education) or the clear benefits (e.g., third graders meeting literacy objectives). Outcomes that are relevant in one locale are not necessarily relevant in others. For example, communities may vary in how they value reducing grade repetition, and such valuations can also vary if state laws or local district criteria restrict social promotion.



**The outcome can be verified and measured.** A valid and reliable measurement tool and evaluation approach must be available that can be used to quantify success at an acceptable cost.



**The outcome can be measured in a sufficiently short time frame.** Lengthy follow-up evaluations of early demonstration ECE programs such as Perry Preschool have found differences in adult outcomes (e.g., earnings, criminal activity) observed at age 21 and even age 40. A tighter time frame, such as five years, is generally recommended for PFS projects (Social Finance UK 2013). However, the ECE PFS projects in Utah and Chicago provide annual payments that begin within a year of intervention and continue annually until children reach 6th and 12th grades, respectively. One way to shorten the time frame is to use measures of success during the early school years as proxies for later success.<sup>2</sup>



**The outcome can be assigned a dollar value.** Various strategies can place monetary values on outcomes, whether the outcomes are government costs avoided or social benefits valued by the community.



**The planned intervention is likely to move the outcome enough to motivate a deal.** Some interventions may affect a given outcome without having a large enough impact to motivate public interest in the intervention or to support repayments sizable enough to attract funders. If most students already score 90 percent on a certain test, for example, there is not much opportunity for improvement. Similarly, if an intervention is

---

<sup>2</sup>The case for using measures of success during the early school years as proxies for later success depends on the strength of evidence regarding the relationship between measures observed during early school years and longer-term outcomes such as high school completion, improved adult earnings, or avoidance of criminal activity.

only expected to reduce the number of preterm births by 4 preterm births per 1,000 pregnant women, it may be difficult to observe any change if the treated population is substantially smaller than 1,000. Previous research on the outcomes and effect size of an intervention can help determine whether the intervention will have a sufficient impact.

## Considerations for Selecting Measures

Often a specific outcome can be measured in multiple ways. For example, language development at the end of preschool can be measured by the Peabody Picture Vocabulary Test III, the Woodcock Johnson IV, the Early Learning Scale, Teaching Strategies GOLD, or any number of other less-known assessments. Each assessment has strengths and weaknesses, and each tends to capture a slightly different subdomain of development—which means the best possible assessment will vary based on the specifics of the PFS project (see toolkit report #1, *The State of the Science on Early Childhood Interventions*). Consulting with local or national experts in child development and education can help in the selection of appropriate assessments. Written materials are available that provide summary reviews describing the attributes of various assessments (see appendix A). Measures should be extended to children receiving the intervention, as well as to those who are not. The following principles may also help guide the selection process of measurement tools, including developmental assessments.<sup>3</sup>



**The assessment tool or other measure has a known track record.** There are obvious benefits to using an assessment tool that has been used in multiple studies and has known properties. It is also important to know whether the assessment tool has been used for children of the same age range, race, and ethnicity as the local population. Of course, the assessment field is constantly evolving, and there can be pros and cons to using a newer and possibly improved measurement tool rather than one with well-known strengths and weaknesses.



**It is feasible to verify success and collect data at an acceptable cost.** One common approach is to use a measure that is already used in the preschool or elementary school system in the locality or state. For example, the Chicago PFS project relies on school readiness and third-grade standardized assessments already in use (or planned for use) throughout the Chicago Public Schools. Although using existing administrative data can reduce costs, sometimes it may be necessary to collect additional data. The costs of doing so will vary

---

<sup>3</sup>Appendix table A.1 in toolkit report #1, *The State of the Science on Early Childhood Interventions*, lists examples of measurement tools for other key outcomes.

depending on the nature of the assessment, the sample of children needed for statistically significant results, whether data are collected for children in a control group, whether the data are collected by an independent evaluator or by teachers, the frequency of data collection, and other elements related to the evaluation of the intervention (see toolkit report #2, *Using Data to Inform Decisionmaking*, and toolkit report #6, *Evaluation Design*).



**The measure has high validity and reliability.** No measure is perfect, but some are more accurate than others at capturing underlying conditions across populations with different characteristics. A developmental assessment has high validity if it actually measures what it claims to measure. In addition, a measure should have high reliability, which means it produces consistent results on repeated trials under similar conditions. The validity and reliability of a measure can be affected by how the data are collected, including who collects the data, when they are collected, and over how many children (see toolkit report #2, *Using Data to Inform Decisionmaking*). They can also vary with the characteristics of the population being measured. For example, a test that is reliable and valid among an English-speaking population may not be reliable for non-English speakers. The measures should also have the appropriate sensitivity to measure likely effects.

## Defining and Evaluating Success

Selecting an outcome and a measure is not enough to define success. It also is important to set the level of success at which payments will be made and to develop a way to evaluate how many children meet those criteria. A crucial consideration is whether payments are made for all children who achieve a certain benchmark, or only if children achieve a higher level of success than a comparison or control group of untreated children. Defining success also depends to some degree on the type of evaluation design used to evaluate the program (see toolkit report #6, *Evaluation Design*). Questions to consider include the following.



**Will success be defined relative to outcomes for “untreated” children—that is, children in a control group?** Some PFS projects define success for the treated group relative to success for a comparison or control group of children who do not participate in the program. Such an evaluation design focuses on the intervention’s true impact by setting the baseline at what would be expected in the absence of the intervention. An alternative approach is to make success payments for all children who meet a selected benchmark, even though some children might meet the benchmark in the absence of the PFS project (box 1). The Chicago PFS project takes both approaches: payments for the avoidance of special education services are based on the rate of special education among children attending the Chicago Child-Parent

Center preschool program relative to the rate of special education use among a control group of children who do not attend Chicago public preschools or Head Start programs. Payments for kindergarten readiness and third-grade literacy are made for each program participant who meets a given benchmark.

## BOX 1

### **PAYMENT STREAMS UNDER TWO DEFINITIONS OF SUCCESS**

*How success is defined can affect the payment stream. Consider two scenarios for success payments. In both, the success payment is set at \$1,000 for each child who is “school ready,” and 500 of 1,000 children (50 percent) in the program meet this benchmark.*

*In scenario one, payments are made based on all children meeting the benchmark. Success payments in this scenario total \$500,000 (500 children × \$1,000).*

*In scenario two, payments are made based on the difference between the program participants and a control group in which 35 percent scored as school ready. Success payments in this scenario total \$150,000 [(0.50 - 0.35) × 1,000 children × \$1,000].*



**If there is a control group, what steps will be taken to maximize the comparability of children in the treatment and control groups?** Random assignment is the gold standard for ensuring the two groups are comparable. Many PFS projects—such as an employment reentry program for former prisoners in New York State—rely on a randomized controlled trial. Other projects have explored alternative evaluation designs to measure success. The Chicago PFS project, for example, will develop a comparison group from a districtwide pool of eligible children who do not attend the Chicago Child-Parent preschool program or Head Start by using a statistical technique called propensity score matching. This method improves the comparability between two nonrandomized groups, but it is subject to biases not present in a randomized controlled trial design (see toolkit report #6, *Evaluation Design*).



**What will be the benchmark for success?** The benchmark might be based on national (or state) levels of performance or standards set by the test developer or an expert panel. The Chicago PFS project uses national data to set benchmarks for two of its three outcomes: its school readiness outcomes (children must be at or above national averages in at least five of six domains) and third-grade literacy outcomes (children must be at or above the 25th percentile of children tested across the country). Another option is to make success

payments if a rate of success is higher than historical rates of success in the local area.



### **What are the dangers of perverse incentives associated with the success metric?**

It is important to be aware that focusing on a particular outcome as “successful” can lead to unintended consequences. Avoiding the need for special education and foster care are both positive results—but not if truly needy children are denied essential services. Both Utah and Chicago have provisions to ensure children who need special education receive it. For example, certain children are excluded from the study population (e.g., those identified as needing special education in preschool and/or those with severe disabilities). In the Utah project, the schools do not know which children received preschool services through the PFS project; no one but the evaluator knows which children are in the study population.



### **How likely is success?** Past evaluations of the planned intervention and/or past administrative data on program outcomes can provide useful information to predict the likelihood of success. It also is helpful to have baseline data on likely outcomes in the absence of the program.

One challenge of using the PFS model for ECE programs is the difficulty of predicting outcomes of ECE services with sufficient certainty, even when compelling evidence is available. Choosing outcomes with short time frames and outcome measures or assessment tools with high reliability and validity reduces the level of risk. Even so, results for the children in one community may differ from observed results in another community because of differences in the children and the community; variation in how the intervention is implemented; normal random variation in small groups of children; and other varying factors, such as economic downturns or adverse conditions (e.g., severe weather, high rates of illness and absenteeism) during assessment periods (see toolkit report #6, *Evaluation Design*).

## **Outcomes and Measures for Repayment in the Chicago and Utah PFS Projects**

As summarized in table 2, the ECE PFS project in Utah uses one outcome (avoided special education). The lead researcher for Utah explains that they viewed this as a useful measure, and one that was easier to value than kindergarten readiness.

To make a social impact bond worthwhile for investors, a success metric must be identified that pays back within a short time period. . . . But . . . “kindergarten readiness” . . . is hard to identify, and it is difficult to define a monetary cost. Thus, given the known cost of special education, research showing the strongest predictor of special education placement is academic achievement in kindergarten and the [selected] preschool program’s success in helping children avoid special education, it was chosen to define a reduction in special education for a group of children at the highest risk for academic failure as the social impact bond’s success metric. (Innocenti 2015, 2)

Special education was also selected as a key outcome in Chicago along with two other outcomes: school readiness and third-grade literacy. As noted, the particular tools measuring these two outcomes are assessments that were already used by the Chicago Public Schools (or, in the case of the third-grade assessment, scheduled to be put in place).

**TABLE 2**

**Outcomes, Measures, Definition of Success, and Pricing in Utah and Chicago**

	UTAH (1)	CHICAGO (1 OF 3)	CHICAGO (2 OF 3)	CHICAGO (3 OF 3)
Outcome	Avoiding need for special education	Avoiding need for special education	Kindergarten readiness	Third-grade literacy
Measure	Whether children receive special education services	Whether child has an IEP issued by Chicago Public Schools	Teaching Strategies GOLD Assessment, which is already used in Chicago Public Schools kindergartens and many other settings	English Language Arts/Literacy portion of the PARCC standardized exam, which is used in Illinois and seven other states
Definition of success	A binary (yes/no) indicator of children being assigned to special education ----- Measured annually in grades K–6; children receiving a very low score on the Peabody Picture Vocabulary Test (two standard deviations below the mean, a score below 70) at the beginning of preschool are counted as “at risk” of needing special education, and repayment is provided for all such children not receiving special education (no comparison group) <sup>a</sup>	A binary indicator (yes/no) of children having an IEP ----- Measured annually in grades K–6, based on the difference in IEP rates among the treatment and comparison groups; comparison group is quasi-experimental, drawn from eligible “no pre-K” children districtwide by using propensity score matching techniques; <sup>b</sup> average effect size measured in grades 4–6 is locked in for grades 7–12 <sup>c,d</sup>	Children scoring at or above national average on five or more of six domains of development: literacy, language, math, cognitive, socioemotional, and physical health ----- Measured at end of preschool, among all children in treatment group (no comparison group) <sup>e</sup>	Children reading at grade level, defined as scoring at or above the 25th percentile on reading portions of the PARCC standardized exam, based on results for students across the country. ----- Measured in spring of 3rd grade, among all children in treatment group (no comparison group) <sup>e</sup>
Pricing <sup>e</sup>	\$2,600 per student, paid each year (for up to 7 years) the student avoids special education	\$9,100 per student, paid each year (for up to 13 years) the student avoids special education; outyear payments increase 1.0% annually	\$2,900 per student, paid once for each student who is kindergarten ready	\$750 per student, paid once for each student reading at grade level in 3rd grade

**Sources:** City of Chicago (2014); Innocenti (2015).

**Notes:** IEP = individualized education plan; K = kindergarten; PARCC = Partnership for Assessment of Readiness for College and Careers.

<sup>a</sup>Children identified as needing special education in preschool are removed from the sample

<sup>b</sup>“No pre-k” means children were not enrolled in the Chicago Public Schools Child-Parent Centers or in City of Chicago Head Start.

<sup>c</sup>Children identified with severe disabilities in preschool or any year of study are removed from sample.

<sup>d</sup>Calculation of success payment includes an adjustment for children moving out of the school district.

<sup>e</sup>PFS contract terms may result in lower success payments than described here, such as when payments are based on only part of the price assigned to the outcome, or when a maximum cap on payments has been reached.

## Strategies for Using Social and Economic Benefit to Price Outcomes in ECE?

Stakeholders in a PFS project must agree about the value of the outcomes defining success. Each outcome underlying success payments must be monetized or have a dollar value assigned. Pricing outcomes is part science and grounded in evidence, but it is also part art. In the most straightforward scenario, an outcome is related to specific costs saved or avoided that accrue within a relatively short time horizon to the agency that will make success payments. In many cases, however, ECE outcomes are not associated with clear-cut cost savings or avoidance, so agencies must determine the social value of those outcomes and how much of that social value they are willing or able to pay for if the program is successful. Although each jurisdiction must determine its own appropriate price for a given outcome, evidence on the total economic benefits of an outcome can help inform those decisions.

When setting prices for selected outcomes, it may be helpful to tackle the problem from different angles. On the one hand, it is useful to think of the unit value of the outcome (which can be based on expected cost savings to schools, such as an avoided need for grade repetition), cost savings to other government agencies, or the broader social and economic benefit. On the other hand, it is important to think about the total amount that is likely to be paid, given the product of the unit price and the likely number of children meeting the definition of success. At least two questions need to be answered about this total repayment amount.

- How large a total repayment is the government willing or able to pay?
- Is the total expected repayment high enough to motivate funders?

For a project to be viable, the total repayment amount must be acceptable to funders, to governments or other outcome payers, and to other key stakeholders. Decisionmakers should weigh this consideration in tandem with total expected cost savings and social and economic benefits.

### Potential Methods for Pricing ECE Outcomes

Most ECE outcomes have both tangible and intangible economic benefits. Tangible economic benefits can be readily monetized, either directly or because they serve as a proxy for other outcomes that can be. Intangible economic benefits have value, but they are not easily monetized. Both types of benefits may be considered by stakeholders planning a PFS project. The following questions frame the specific approaches PFS stakeholders can use when pricing

ECE outcomes (see appendix B for additional resources related to assigning dollar values to ECE outcomes).



**How much is the agency willing to pay?** Agencies making success payments may simply price an outcome by making a qualitative judgment about the worth of some or all of its intangible benefits. That is, how much does the outcome contribute to individual, community, or societal well-being? For example, improved achievement test scores may be meaningful, in and of themselves, to school districts charged with successfully providing high-quality education to all students. Communities may also see it as a moral obligation to ensure that all children start school ready to learn. Other outcomes of ECE PFS projects that could be priced based on intangible benefits to schools include students' grades and the incidence of disciplinary issues such as suspensions (Kohli et al. 2015).

This contingent valuation approach—in which intangible benefits are priced according to how much someone would be willing to pay to get a positive outcome—has rarely been used to value outcomes in cost-benefit analyses for ECE interventions. However, contingent valuation is common in cost-benefit analyses of health improvement and environmental projects.<sup>4</sup>



**What is the dollar value of costs avoided?** Government or other outcome payers may also factor tangible economic benefits into the estimation of what they are willing to pay. In fact, cost-benefit analysts often avoid the tricky issue of monetizing intangible benefits by only monetizing tangible benefits. This practice tends to underestimate the total value of an intervention, but the tangible economic benefits of ECE are often large enough on their own to demonstrate a good return on investment. Some desired ECE outcomes that increase costs in the short term may nonetheless retain wide acceptance as a social good. When an intervention prevents negative outcomes, those outcomes are usually valued according to costs avoided, and their price is based on what is actually being spent as a result of the negative outcome.<sup>5</sup> Evaluations of demonstration ECE programs have measured outcomes that are associated with expected cost savings for school districts (e.g., through reduced need for special education services and grade repetition) as well as other government agencies and citizens (e.g., through reduced rates of juvenile and adult criminal activity or child abuse and neglect). Both the Utah and Chicago PFS projects tie success payments to special education avoidance and have assigned prices to that outcome that reflect the cost of a year of special education in their area.

---

<sup>4</sup> Contingent valuation can also be based on how much someone would need to be paid to accept a negative outcome.

<sup>5</sup> Actual spending on negative outcomes is usually viewed as a lower-bound estimate on the tangible economic value of avoiding those outcomes. This view reflects that public spending on these types of programs generally does not represent an equilibrium price in a competitive market.



**What is the dollar value of other tangible economic benefits?** Savings that accrue when negative outcomes are avoided is one type of tangible economic benefit.

Another is the economic benefit that is realized when there are positive outcomes.

As discussed above, many outcomes of ECE, especially those measured in the short term, have primarily intangible economic benefits. Earnings from increased labor market participation among program participants once they become adults is the main positive tangible economic benefit that is captured in cost-benefit analyses of ECE programs. Most of the associated value accrues to the participants themselves, but some also accrues to society as a whole through increased contributions to the tax system and overall economic growth.

Analysts have used different approaches to price ECE outcomes according to their association with adult labor market earnings. Some have used national survey or Internal Revenue Service data to directly estimate the effect higher achievement test scores in elementary school have on adult labor market earnings. For example, Bartik, Gormley, and Adelstein (2012) build on work by Chetty and colleagues (2010) that indicates a one percentile increase in test scores is associated with a one percentile increase in earnings. Those estimates indicate that, on average, a year of the Tulsa prekindergarten program yielded a present value benefit of between approximately \$13,000 and \$30,000 in lifetime earnings.<sup>6</sup> The Washington State Institute of Public Policy takes a slightly different approach in its cost-benefit analysis tool. It estimates the value of higher achievement test scores by first evaluating the impact of test scores on the probability of high school graduation and then calculating the value of increased earnings associated with being a high school graduate (WSIPP 2015b). Thus, achievement test scores serve as a proxy for high school graduation, which serves as a proxy for adult earnings.

Other school performance measures are also candidates to serve as proxies for high school graduation and thus are feasible to price as success payment outcomes. For example, some evidence suggests that grades and on-time promotion better predict high school completion than test scores (Allensworth and Easton 2007; Neild and Balfanz 2006).<sup>7</sup> Overall, economic benefits estimated on the basis of lifetime earnings can be large. Even if PFS projects don't set outcome prices that fully cover these tangible economic benefits, knowing the amounts can help decisionmaking about prices for outcomes and providing information about the overall benefits that might accrue from a given program.

---

<sup>6</sup> The economic benefit varied with children's family income and whether children were enrolled in a part-day or full-day classroom.

<sup>7</sup> Although these outcomes are promising, there is less evidence linking ECE interventions to these outcomes than there is linking them to achievement test scores. PFS projects interested in targeting outcomes such as grades or on-time promotion may need to do additional research to ensure it is sufficiently likely that the proposed intervention will significantly affect those outcomes.

## Other Considerations in Pricing Outcomes

Beyond estimating the actual total value of a given outcome, PFS projects may consider other factors when pricing outcomes. Examples of these considerations include the following.



**Pricing outcomes based only on benefits for certain stakeholders.** Because the total benefit of ECE programs can be large, one option in pricing outcomes is to estimate their value based only on the economic benefits that accrue or are most important to the government or other agencies making the success payments.



**Adjusting prices over time to account for inflation.** When success payments are made on outcomes that occur over several years, projects must consider whether the price assigned will remain constant or will change over time. In the Chicago ECE project, the price assigned to special education avoidance increases annually to account for the fact that the schools expect the cost of special education will go up each year. The Utah ECE project took a different approach and used a constant price for special education avoidance over a period of years.



**Translating outcome prices into success payment amounts.** The price or value of an outcome is not necessarily equivalent to a PFS repayment amount. In the Utah ECE project, the outcome payer (school district or state) retains some of the annual economic benefit from avoided special education. Specifically, annual success payments represent only 90 percent of avoided cost for each child who does not need special education services. After funders are repaid principal with 5 percent interest, success payments will be reduced to 40 percent of the price assigned to a year of special education. This approach is useful in that funders and government agencies or other outcome payers are explicitly sharing the tangible economic benefits as they accrue. In addition, Utah's pricing approach may be designed to help account for the fact that the model is not using a comparison group in its definition of success. Without a comparison group, it is challenging to determine how many children would have avoided needing special education without the program. Given that the school district and state are not realizing cost savings for that subset of children, paying only a portion of the per child price of special education helps bring the success payment amount closer to Utah's actual expected savings.

---

<sup>8</sup> In these examples, payments are being made for outcomes in roughly the same time frame in which the economic benefit (special education cost avoidance) occurs. For cases in which payments are tied to economic benefits that will accrue in the future, the price assigned to those benefits actually needs to be discounted to reflect the fact that a dollar is worth more today than it will be in the future.

## Conclusion

Pricing outcomes is an important part of the PFS model; it provides a concrete measure of whether the value of an intervention outweighs its costs. Outcome pricing also factors into the calculation of success payments. Because the analytic approach to pricing outcomes can be complicated, jurisdictions may want to consult with experts in evaluation, performance management, and/or cost-benefit analysis as they develop or justify a proposed outcome-pricing plan. They may also want to consider the monetary values other analysts have assigned to outcomes of ECE programs (see appendix B for some resources). Ultimately, outcome-pricing decisions in a given PFS project are specific to the project and its stakeholders. Those decisions depend heavily on the cost of implementing the proposed project, the relative feasibility of different data collection and evaluation strategies, and, most important, what is of greatest value to the larger community and the project's stakeholders.



# Appendix A

## APPENDIX A.

### Selected Compendia on Instruments to Measure Children's Developmental Outcomes

REPORT TITLE	AGE	CONTENT	COMPENDIUM PURPOSE	DEVELOPMENT DOMAINS
Early Childhood Assessment: Why, What, and How	Birth–5 years	<p>Four-part report including</p> <p>Purposes of assessment, a brief history of early childhood standards</p> <p>What should be assessed and why</p> <p>How to assess</p> <p>Assessing systematically</p>	<p>Identify important outcomes for children from birth to age 5 and the quality and purposes of different techniques and instruments for developmental assessments</p> <p>Present guidelines for assessment related to purposes, domains and measures, implementation, and systems</p>	<p>Physical well-being and motor development</p> <p>Social and emotional development</p> <p>Approaches toward learning</p> <p>Language development (including emergent literacy)</p> <p>Cognition and general knowledge (including mathematics and science)</p>
A Guide to Assessment in Early Childhood: Infancy to Age Eight	Birth–8 years	<p>Summary table of 130 assessment instruments including age, purpose, domains, and languages</p> <p>Individual descriptions of 29 assessment tools including age, time, administration and training, cost, languages, source, scoring, subscales, norming samples, and reliability and validity</p>	<p>Serve as a resource for designing professional development activities for program administrators and direct service staff responsible for gathering and interpreting assessment information</p>	<p>Approaches to learning</p> <p>Cognitive</p> <p>Language</p> <p>Literacy</p> <p>Math, numeracy</p> <p>Motor</p> <p>Adaptive</p> <p>Social and emotional</p>
Early Childhood Developmental Screening: A Compendium of Measures for Children Ages Birth to Five	Birth–5 years	<p>One summary table with general information about 16 screeners; 1 summary table with information about the evidence of reliability and validity for 17 screeners; and 1 summary table with information about the evidence of reliability and validity for different languages and different populations for 17 screeners</p> <p>Individual profiles of 14 developmental screening tools including type of purpose, time, administration, languages, domains, age, number of items, settings, availability and cost, training, information reporting system, family and parent input, options for use with special and diverse populations, norming samples, and reliability and validity</p> <p>"Abbreviated" profiles of 2 developmental screening tools</p>	<p>Discuss the purpose of developmental screening and how it differs from child assessment</p> <p>Translate technical psychometric information about the reliability and validity of commonly used developmental screening tools into language that is easily understood by early childhood practitioners</p> <p>Highlight areas in which the early childhood field is lacking information on reliability and validity of available developmental screening tools</p>	<p>Varies—inclusion criteria require each screening tool included to cover multiple developmental domains</p>

REPORT TITLE	AGE	CONTENT	COMPENDIUM PURPOSE	DEVELOPMENT DOMAINS
Measures of Child Development: A Review	2–2½ years	<p>Table showing main characteristics of 35 measures of child development</p> <p>Detailed descriptions of 13 measures of child development including purpose, age, format and administration, time, scoring, training and materials, standardization and psychometrics, use, strengths, and limitations</p>	<p>Identify and review the range of existing standardized instruments and tools that could be used to measure children’s developmental progress at age 2–2 ½ years to enable selection of a population-level outcome measure that meets specific criteria</p> <p>Pay specific attention to tools that could be used as part of the 2–2 ½ years Healthy Child Programme (HCP) review in the United Kingdom</p>	<p>Physical</p> <p>Social and emotional</p> <p>Cognitive</p> <p>Speech and language</p>
Understanding and Choosing Assessments and Developmental Screeners for Young Children Ages 3–5: Profiles of Selected Measures	3–5 years	<p>Summarizes information from the assessment and developmental screening instruments most commonly used by Head Start programs for 3-, 4-, and 5-year-olds through</p> <p>3 summary tables with general information, information about the evidence of reliability and validity, and information about the evidence of reliability and validity for different languages and different populations for 8 assessments</p> <p>3 summary tables with the same information for 9 developmental screeners</p> <p>Individual profiles for the 8 assessments and 10 developmental screeners including type of purpose, time, administration, languages, domains, age, number of items, settings, availability and cost, training, information reporting system, family and parent input, options for use with special and diverse populations, norming samples, and reliability and validity</p>	<p>Help Head Start managers and other ECE administrators review information regarding the reliability and validity of commonly used assessment and developmental screening tools</p> <p>Increase awareness about reliability and validity and how to evaluate whether an instrument is reliable and valid</p> <p>Highlight areas in which the early childhood field is lacking information or reliability and validity of early childhood assessments and developmental screeners</p>	<p>Varies—inclusion criteria require all assessments and screening tools to cover more than two domains of the Head Start Child Outcomes Framework</p>

**Sources:** Bedford, Walton, and Ahn (2013); "Snow and Van Hemel (2008)"; Halle et al. (2011); Moodie et al. (2014); Slentz, Early, and McKenna (2008).

# Appendix B

## SELECTED RESOURCES RELATED TO MONETIZING EARLY CHILDHOOD EDUCATION OUTCOMES

### **Washington State Institute for Public Policy Benefit-Cost Model**

The WSIPP benefit-cost model (WSIPP 2015a) estimates the present-value economic benefit of a range of interventions, including ECE. Because the model is used primarily to help the state legislature make informed spending decisions, some of its parameters are specific to costs and benefits in Washington. However, the model serves as an important resource for analysts across the country because of its comprehensive and thorough methodology in looking across studies to estimate average program effects and applying those effects to a rigorous cost-benefit calculation. A report focusing on ECE findings is also available (Kay and Pennucci 2014).

### **Long-Term Longitudinal Evaluations of Specific Model Preschool Programs**

Two longitudinal studies are the primary data sources for all long-term estimates of the potential impact of preschool programs: the Perry Preschool Project Evaluation (Schweinhart et al. 2005) and the Chicago (Child-Parent Centers) Longitudinal Study (Reynolds et al. 2011). Both studies evaluated the impact of preschool on participant outcomes into the adult years and had accompanying cost-benefit analyses. (The Abecedarian Project is another key source of longitudinal and cost-benefit estimates, but because the intervention started in infancy, its findings are not as applicable to programs primarily targeting 3- or 4-year olds.) Evaluation reports from Perry Preschool and the Chicago Longitudinal Study offer examples of valuation methods, information about the size of actual economic benefits realized as a result of those programs, and information about how those benefits are distributed across different outcomes and stakeholders.

### **Estimates of Economic Benefits of Proposed Preschool Programs in Specific States**

Analysts in many states and localities have drawn on resources such as the WSIPP benefit-cost model and the Perry Preschool and Chicago Longitudinal studies to estimate the potential return on investment of increased spending on preschool programs. These analyses are useful examples of the process PFS projects might engage in: defining likely outcomes, estimating the size and economic benefit of those outcomes, and comparing program benefits to costs. However, it is important to note that the purpose of these types of reports is to more generally demonstrate that the economic benefits of a proposed program can outweigh the costs, rather than to prove they actually will. Because this purpose differs slightly from estimating the value of select, relatively short-term outcomes for the purpose of making success payments, it is unlikely that PFS projects will want to fully replicate the methods or that they will be able to directly pull a given outcome valuation from these sources. Two of the most useful reports in this category are *The Economics of Investing in Universal Preschool Education in California* (Karoly and Bigelow 2005) and *Cost Savings of School Readiness per Additional At-Risk Child in Detroit and Michigan* (Richard and Diaz 2011).

### **United Kingdom Department of Works and Pensions Innovation Fund Rate Cards**

Although related to an intervention for older children, and different in that it is run by a federal government, an additional resource worth mentioning is the UK's Department for Works and Pensions approach to pricing outcomes. The department sets a specific level of payment it will make to service providers participating in the agency's Innovation Fund project, which was designed to help disadvantaged youth succeed in education, training, and employment opportunities. Outcomes and their pricing include improved attitude toward school (£700), improved behavior (£1,300), improved school attendance (£1,400), and test scores (£900–5,100 depending on score) (Department for Work and Pensions Innovation Fund, n.d.).

## References

- Allensworth, Elaine M., and John Q. Easton. 2007. *What Matters for Staying On-Track and Graduating in Chicago Public High Schools: A Close Look at Course Grades, Failures, and Attendance in the Freshman Year*. Chicago, IL: Consortium on Chicago School Research at the University of Chicago. <https://consortium.uchicago.edu/sites/default/files/publications/07%20What%20Matters%20Final.pdf>.
- Balfanz, Robert, Liza Herzog, and Douglas J. Mac Iver. 2007. "Preventing Student Disengagement and Keeping Students on the Graduation Path in Urban Middle-Grades Schools: Early Identification and Effective Interventions." *Educational Psychologist* 42 (4): 223–35.
- Bartik, Timothy J., William Gormley, and Shirley Adelstein. 2012. "Earnings Benefits of Tulsa's Pre-K Program for Different Income Groups." *Economics of Education Review* 31 (6): 1143–61.
- Bassok, Daphna, RaeHyuck Lee, Sean F. Reardon, and Jane Waldfogel. 2015. "Socioeconomic Gaps in Early Childhood Experiences, 1998 to 2010." EdPolicyWorks Working Paper Series No. 38. Charlottesville, VA: Curry School of Education, Frank Batten School of Leadership and Public Policy, University of Virginia. [http://curry.virginia.edu/uploads/resourceLibrary/38\\_Gaps\\_In\\_Early\\_Experience.pdf](http://curry.virginia.edu/uploads/resourceLibrary/38_Gaps_In_Early_Experience.pdf).
- Bedford, Helen, Suzanne Walton, and Jane Ahn. 2013. "Measures of Child Development: A Review." London: UCL Institute of Child Health.
- Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2010. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence From Project STAR." Working Paper 16381. Cambridge, MA: National Bureau of Economic Research. <http://www.nber.org/papers/w16381.pdf>.
- City of Chicago. 2014. "Loan Agreement and Contract with IFF Pay For Success, LLC to Serve at-Risk Children to Increase School Readiness and Reduce Later Public School Spending." <http://www.payforsuccess.org/sites/default/files/o2014-8677.pdf>.
- Department for Work and Pensions Innovation Fund. n.d. "Innovation Fund: Key Facts." London: United Kingdom Department for Work and Pensions. [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/212328/hmg\\_g8\\_factsheet.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/212328/hmg_g8_factsheet.pdf).
- Halle, Tamara, Martha Zaslow, Julia Wessel, Shannon Moodie, and Kristen Darling-Churchill. 2011. "Understanding and Choosing Assessments and Developmental Screeners for Young Children Ages 3-5: Profiles of Selected Measures." Washington, DC: US Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation. [http://www.acf.hhs.gov/sites/default/files/opre/screeners\\_final.pdf](http://www.acf.hhs.gov/sites/default/files/opre/screeners_final.pdf).
- Innocenti, Mark. 2015. "The Facts: Behind Utah's Social Impact Bond for Early Childhood Education." New York: Nonprofit Finance Fund. <http://www.payforsuccess.org/sites/default/files/UTAH%20Facts.pdf>.
- Karoly, Lynn A., and James H. Bigelow. 2005. *The Economics of Investing in Universal Preschool Education in California*. Santa Monica, CA: RAND Corporation. [http://www.acf.hhs.gov/sites/default/files/opre/hs\\_quality\\_report\\_4\\_28\\_14\\_final.pdf](http://www.acf.hhs.gov/sites/default/files/opre/hs_quality_report_4_28_14_final.pdf).
- Kay, Noa, and Annie Pennucci. 2014. "Early Childhood Education for Low-Income Students: A Review of the Evidence and Benefit-Cost Analysis." Doc. No. 14-01-2201. Olympia, WA: WSIPP. [http://www.wsipp.wa.gov/ReportFile/1547/WSipp\\_Early-Childhood-Education-for-Low-Income-Students-A-Review-of-the-Evidence-and-Benefit-Cost-Analysis\\_Full-Report.pdf](http://www.wsipp.wa.gov/ReportFile/1547/WSipp_Early-Childhood-Education-for-Low-Income-Students-A-Review-of-the-Evidence-and-Benefit-Cost-Analysis_Full-Report.pdf).
- Kohli, Jitinder, Megan Golden, Joe Coletti, and Luke Bo'sher. 2015. "From Cashable Savings to Public Value: Pricing Program Outcomes in Pay for Success." Washington, DC: Center for American Progress. <https://cdn.americanprogress.org/wp-content/uploads/2015/08/31113147/PayForSuccess-brief.pdf>.
- Moodie, Shannon, Paula Daneri, Samantha Goldhagen, Tamara Halle, Katie Green, and Lauren LaMonte. 2014. "Early Childhood Developmental Screening: A Compendium of Measures for Children Ages Birth to Five." Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, US Department of Health and Human Services.
- Neild, Ruth Curran, and Robert Balfanz. 2006. "Unfulfilled Promise: The Dimensions and Characteristics of Philadelphia's Dropout Crisis, 2000-2005." Baltimore, MD, and Philadelphia, PA: Philadelphia Youth Network, The Johns Hopkins University, and University of Pennsylvania. <http://files.eric.ed.gov/fulltext/ED538341.pdf>.

- Reynolds, Arthur J., Judy A. Temple, Barry A. White, Suh-Ruu Ou, and Dylan L. Robertson. 2011. "Age-26 Cost-Benefit Analysis of the Child-Parent Center Early Education Program." *Child Development* 82 (1). <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3817956/>.
- Richard, Chase, and Jose Diaz. 2011. "Cost Savings of School Readiness per Additional at-Risk Child in Detroit and Michigan." Saint Paul, MN: Wilder Research. <https://www.wilder.org/Wilder-Research/Publications/Studies/Detroit's%20One%20Child%20School%20Readiness%20Dividend/Cost%20Savings%20of%20School%20Readiness%20per%20Additional%20At-Risk%20Child%20in%20Detroit%20and%20Michigan,%20Full%20Report.pdf>.
- Rumberger, Russell W. 2011. *Dropping Out: Why Students Drop Out of High School and What Can Be Done About It*. Cambridge, MA: Harvard University Press. <http://mina.education.ucsb.edu/rumberger/book/ch1.pdf>.
- Schweinhart, Lawrence J., Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, and Milagros Nores. 2005. "The High/Scope Perry Preschool Study Through Age 40: Summary, Conclusions, and Frequently Asked Questions." Ypsilanti, MI: High/Scope Press. [http://www.highscope.org/file/specialsummary\\_rev2015\\_01.pdf](http://www.highscope.org/file/specialsummary_rev2015_01.pdf).
- Slentz, Kristine L., Diane M. Early, and Margaret McKenna. 2008. "Guide to Assessment in Early Childhood: Infancy to Age Eight." Olympia: Washington State Office of Public Instruction. [http://www.k12.wa.us/EarlyLearning/pubdocs/assessment\\_print.pdf](http://www.k12.wa.us/EarlyLearning/pubdocs/assessment_print.pdf).
- Snow, Catherine E., and Susan B. Van Hemel, eds. 2008. *Early Childhood Assessment: Why, What, and How?* Washington, DC: National Academies Press.
- Social Finance UK. 2013. "A Technical Guide to Developing Social Impact Bonds." Guide to Social Impact Bond Development. London: Social Finance UK. <http://www.socialfinance.org.uk/wp-content/uploads/2014/05/Technical-Guide-to-Developing-Social-Impact-Bonds1.pdf>.
- WSIPP (Washington State Institute for Public Policy). 2015a. "Benefit-Cost Results." Olympia, WA: WSIPP. <http://www.wsipp.wa.gov/BenefitCost>.
- . 2015b. "Benefit-Cost Technical Documentation." Olympia, WA: WSIPP. <http://www.wsipp.wa.gov/TechnicalDocumentation/WsippBenefitCostTechnicalDocumentation.pdf>.