

THE COST - BENEFIT ANALYSIS OF THE PRESCHOOL CURRICULUM COMPARISON STUDY

A Final Report to the John D. and Catherine T. MacArthur Foundation

Power of Measuring Social Benefits Research Initiative

Lawrence J. Schweinhart, Ph.D.
HighScope Educational
Research Foundation
Lena Malofeeva, Ph.D.
HighScope and LeCroy & Milligan
Associates, Inc.

James J Heckman, Ph.D.
Rodrigo Pinto, M.S.
Seong Moon, M.S.
Adam Yavitz, B.S.
University of Chicago
Economics Research Center

September 30, 2010

Acknowledgements

The research team for this project includes James Heckman, the Henry Schultz Distinguished Service Professor of Economics at the University of Chicago, Professor of Science and Society, University College Dublin, Alfred Cowles Distinguished Visiting Professor, Cowles Foundation, Yale University and Senior Fellow, American Bar Foundation and Lawrence Schweinhart, HighScope's President. Lena Malofeeva is a Senior Evaluation Associate at LeCroy & Milligan Associates, Inc.. Rodrigo Pinto and Seong Moon are graduate students at the University of Chicago. Adam Yavitz is a research professional at the University of Chicago.

We want to thank Marijata Daniel-Echols, HighScope's Research Chair, and Alison Baulos for their support and assistance.

Supplementary materials may be retrieved from http://jenni.uchicago.edu/mac_arthur.

Suggested Citation:

HighScope (2010). *The Cost-Benefit Analysis of the Preschool Curriculum Comparison Study Through Age 23*. Ypsilanti, MI: HighScope.

Table of Contents

Acknowledgements.....	2
Table of Contents.....	3
List of Tables	4
Executive Summary.....	5
Introduction and Background	7
Noncognitive Skills.....	9
Planning.....	9
Decision-Making and Control.....	9
Direct Instruction Did Not Improve Outcomes.....	10
Cost-Benefit Analyses.....	11
Curriculum Comparison Studies.....	14
The Research Question	15
Hypotheses.....	16
Methods.....	16
Sample	16
Random Assignment Procedures	20
Curriculum Models	21
Findings.....	22
Cost-Benefit Analyses.....	23
Initial Cost.....	24
Benefits	25
Education.....	25
K-12 Education.....	25
GED and Special Education.....	25
2- and 4-Year Colleges.....	26
Vocational Training.....	26
Employment and Earnings.....	26
Crime.....	28
Deadweight Cost of Taxation.....	32
Differences in Benefit-to-Cost Ratios across Curriculum Groups	33
Conclusions.....	37
References.....	40

List of Tables

Table 1. Cost-Benefit Analyses of Three Long-Term Studies	12
Table 2. Demographic Info	19
Table 3. Four Theoretical Approaches to Preschool Education	21

THE COST-BENEFIT ANALYSIS OF THE PRESCHOOL CURRICULUM COMPARISON STUDY THROUGH AGE 23

The Final Report to the John D. and Catherine T. MacArthur Foundation

Executive Summary

The economic case for investing in early childhood programs came first from the evidence from the Perry Preschool Program. The current economic literature does not examine early childhood curriculum effectiveness. In this study, a cost-benefit methodology compared the initial cost of investment in different types of early childhood curriculum to benefits to individuals and society later in life. We identified three types of interventions that were widely used around the world - Direct Instruction (DI), HighScope (HS), and a traditional theme-based child-centered approach called Nursery School (NS). We calculated the cost-benefit ratios for each model based on the longitudinal data from the experimentally designed Preschool Curriculum Comparison Study. This has never been attempted before in economics, education, or psychology. We could test this research question with confidence due to the random assignment design used in the study, the longitudinal nature of the data (with the latest wave collected at age 23), and very strict adherence to curriculum models (e.g., all three programs had similar funding, similar salaries, same class schedules, etc.) so that differences in child outcomes could be related to differences in intervention approaches. Statistical procedures involving restricted permutations to correct for initial

randomization swaps as well as statistical adjustments for familywise error rates were utilized. Sensitivity analyses were run to see how the amount of cost-savings varied depending on various assumptions made in calculations. The cost-benefit estimates between HS and DI and between HS and NS are presented in this report. Even though the HS group seemed to outperform DI and NS participants on some outcomes, the effects were variable and generally not statistically significantly different from zero. While high quality early childhood interventions pay off, our preliminary estimates indicated that participation in HS or NS did not lead to statistically significantly higher savings at age 23 than participation in DI. The cost-benefit estimates from the age 23 Curriculum Comparison Study are, however, higher than those observed for the HighScope Perry Preschool Study at the same age interval. Given the observed trend in the Curriculum Comparison Study, we expect a positive increase in cost-benefit ratios once participants age into midlife.

Introduction and Background

The findings of the HighScope Perry Preschool study, one of the most famous educational experiments of the 20th century, have challenged our understanding of how children develop and prosper in life. Young people born in poverty have greater educational and economic success and reduced crime rates if they attend a high-quality preschool program than if they do not do so; such programs return seven to ten dollars to taxpayers for every dollar invested (Heckman, Moon, Pinto, Savelyev, & Yavitz, 2010; Schweinhart, Barnes, & Weikart, 1993). The question then becomes whether these benefits result equally from all curriculum approaches or from some more than from others. The HighScope Preschool Curriculum Comparison study was designed to answer this question (Schweinhart & Weikart, 1997a). It assessed which of three preschool curriculum models were the most effective for young children living in poverty. It was hypothesized that young people born in poverty would achieve greater success and manifest greater social responsibility if they had attended a HighScope (HS) or traditional Nursery School preschool program (NS) than if they had attended a Direct Instruction preschool program (DI).

The Curriculum Comparison Study differences between the three groups were negligible the first decade. At age 23, however, certain patterns emerged (Schweinhart & Weikart, 1997a). The interpretation of the findings revolves around two comparisons: DI versus the other two curriculum models and the HS model versus the NS model. While statistically significant curriculum group differences were not found for most variables,

ten such differences were previously reported even after accounting for gender. All of these differences favored the HS group, the NS group, or both over the DI group:

- The HS group had advantages over the DI group in highest year of schooling planned, higher percent living with spouse, fewer sources of irritation, less self-reported misconduct at age 15, fewer felony arrests, and fewer arrests for property crimes.
- The NS group had advantages over the DI group in fewer suspensions from work and fewer felony arrests at age 22 and over.
- Both HS and NS groups had advantages over the DI group in experiencing fewer years of identified emotional impairment or disturbance and in ever doing volunteer work (Schweinhart & Weikart, 1997a).

To summarize previous findings, across all outcomes at age 23 the DI group did not have any advantages over either the HS group or the NS group. On no variable did the HS group have an advantage over the NS group, and on no variable did the NS group have an advantage over the HS group (Schweinhart & Weikart, 1997a).

What processes make HS or NS preschool curricula improve adult outcomes better than preschool curriculum based on DI? Curriculum comparisons suggest that planning and social reasoning are important ingredients of executive functioning and social intelligence that helped shape young children's minds, thus contributing to their success later in life.

Noncognitive Skills

The early childhood years are important for the development of noncognitive skills and self-regulation. These early skills provide the foundation for later success in life.

Planning

Both the ability to plan and the ability to initiate may mediate between the HS and NS curricula and later success and the development of social responsibility. In the HS curriculum and, to a lesser extent, in the NS curriculum, adults encourage children to take initiative, to select and plan their own activities, and to be decision makers to the extent that their ages and abilities permit. Early childhood education can help children develop their ability to take initiative and make plans (Bronson, 1994), and this ability has now been found to be a factor in school achievement (Cohen, Bronson, & Casey, 1995). Linking planning ability to social responsibility, correctional programs that place an emphasis on participants' critical thinking and reasoning about their behavior have been found to effectively prevent crime (Andrews et al., 1990; Antonowicz & Ross, 1994).

Decision-Making and Control

Children in both HS and NS groups had more control over their preschool classroom environments and were able to exercise their choices there in a much more systematic fashion. Both HS and NS based their programs on what we now call developmentally appropriate practices. There were opportunities in the classroom to put children in charge of parts of their day, to learn from them, and to share in their interests. DI group instruction was based on workbooks and preset objectives and exercises that led to

those objectives. Systematic fidelity of curriculum implementation observations confirmed these differences. Children in the DI classrooms spent about 72% of their time in teacher-directed activities as opposed to 48% for children in the HS group and 56% for children in the NS group (Schweinhart & Weikart, 1997a).

Direct Instruction Did Not Improve Outcomes

With about half of the participants in the DI group having received treatment for emotional disturbance and with three times as many felony arrests, the DI group, especially males, was not as successful in the long run as the other two curriculum groups. Curriculum comparisons point out that unlike the HS and NS approaches, DI focused on academic objectives, not on planning or social objectives. This strategy, while successful short-term, did not lead to long-term improved outcomes for children served.

This report explains the methodology and findings of the cost-benefit analysis of the 3 curriculum programs described above. We first briefly review cost-benefit analyses results coming from the short- and long-term findings of longitudinal studies and curriculum comparison studies to date and identify ways of linking these independent research methodologies. We then provide further details for the Curriculum Comparison Study and describe the results of the cost-benefit analysis conducted with these data. Finally, we conclude with recommendations and policy implications.

Cost-Benefit Analyses

Cost-benefit analysis is a way of measuring costs and benefits of a program. It allows one to estimate the net gain to society and individual society members. If the return on investments are known, better decisions about investments could be achieved. The cost-benefit analysis of the Curriculum Comparison study is of interest because policymakers, stakeholders, politicians, educators and other parties involved need to know on how investing in different types of curriculum affects children long-term.

Cost-benefit analysis is a well-known technique, but is still infrequently applied to educational data (Schweinhart et al., 2005). The cost of a given program is compared to benefits both to participants and the public. Given that benefits usually occur beyond early childhood, these benefits need to be discounted (valued at a lower rate). Usually, such advantages to interventions as gains in earnings, crime reduction, differences in retention rates and special education services are compared against the costs of the programs to derive their net present value. Most previous cost-benefit analyses compared returns on investments between those participating in a targeted intervention and those who did not participate in an intervention (e.g., HighScope's Perry Preschool Study). This study went a step further to calculate returns on investment produced by different curriculum models. This has never been attempted before in economics, education, or psychology.

Longitudinal studies show that high quality early childhood interventions pay off (Schweinhart et al., 2005). **Error! Reference source not found.** summarizes three longitudinal studies- HighScope Perry Preschool Study- Carolina Abecedarian Study and Chicago Child-Parent Centers Study-- and the cost-benefit ratios they report.

Table 1. Cost-Benefit Analyses of Three Long-Term Studies

	HighScope	Carolina	Chicago Child-
Economic Statistic	Perry Preschool	Abecedarian	Parent Centers
Program Cost	\$15,166	\$34,476	\$6,956
Program Cost per Year	8,540	13,362	4,637
Public Return Total	195,621	--	26,637
Public Return per Dollar Invested	12.90	--	3.83
Societal Return Total	244,812	130,300	49,364
Societal Return per Dollar Invested	16.14	3.78	7.10

One drawback of these cost-benefit analyses is that the economic benefits they produce have not been compared across different curriculum models. Policymakers not only need to know that early investments have long-lasting effects, but also which specific types of interventions produce higher, longer-lasting outcomes.

Separately, the field of cost-benefit analysis continues to struggle to establish its own standards of quality, and methodologies tend to vary from researcher to researcher. It is not uncommon that some of the methodological aspects are often overlooked. Randomized Controlled Trial studies, for example, are considered to be the gold standard for assessing the effectiveness of a treatment; however, such randomization procedures are often compromised, and the statistical consequences of these compromises are very often neglected. Statistical procedures involving restricted permutations to correct for nonrandom swaps as well as statistical adjustments for familywise error rates are utilized in the cost-benefit analysis presented in this report to account for the presence of multiple outcomes and imperfections in the randomized assignment of treatments.

When randomization imperfections happen - participants who did not comply with the initial assignment may be precisely the ones who would respond adversely to treatment. As a result, differences in outcomes between the treated and untreated groups may simply be due to the imperfectness in the randomization instead of the treatment itself. Failing to account for the imperfections of the randomization can produce biased inferences, casting doubt on the validity of traditional inference. We strengthen methods to measure social benefits by proposing a set of standards and cost-benefit analysis techniques that address critical theoretical, methodological, and data limitations, by dealing with issues related to randomization, multiple hypothesis, and small sample size, accounting for locally determined costs, missing data, the deadweight costs of taxation, and the value of non-market benefits and costs.

Specifically, this paper presents estimates of the rate of return and the benefit-cost ratios for the Curriculum Comparison Study groups.

- We account for compromised randomization in evaluating the three programs. The randomization implemented in this study is somewhat problematic because of reassignment of group status *after* random assignment (Heckman, Moon, Pinto, Savelyev, & Yavitz, 2009).
- For the remaining components of costs and benefits to which meaningful standard errors cannot be determined, we examine the sensitivity of estimates of rates of return to plausible ranges of assumptions.
- We present estimates that adjust for the deadweight costs of taxation. Previous estimates ignore the costs of raising taxes in financing programs.
- We use a much wider variety of methods to impute within-sample missing earnings than have been used in the previous literature, and examine the sensitivity of our estimates to the application of alternative imputation procedures that draw on standard methods in the literature on panel data.¹
- We use state-of-the-art methods to extrapolate missing future earnings. We examine the sensitivity of our estimates to plausible alternative assumptions about out-of-sample earnings.
- We use local data on costs of education, crime, and welfare participation whenever possible, instead of following earlier studies in using national data for these components of the rate of return.

Curriculum Comparison Studies

The position that the use of a defined curriculum model is an essential part of the definition of quality in early childhood education is not new. There have been a number of curriculum comparison studies, both short and long-term, that point to the effectiveness of using a curriculum as well as the greater effectiveness of some curriculum models than others. The

¹See, e.g., MaCurdy (2007) for a survey of these methods.

Louisville Study of Head Start (Miller & Bizzell, 1983), the University of Illinois Study (Karnes *et al.*, 1983), and the HighScope Preschool Curriculum Comparison study (Schweinhart & Weikart, 1997b) all compared Direct Instruction and the child-centered Nursery School model. These studies were experimental in nature. Several recent preschool curriculum studies (e.g., Burts *et al.*, 1990; Marcon, 1992; Stipek *et al.*, 1992) employed quasi-experimental designs and did not track participants up to high school graduation. None of these studies, however, has attempted cost-benefit analysis before. Therefore, while teachers, administrators and stakeholders have choices in curriculum utilization, it is not clear what impact these models have on return on investment.

The Research Question

The cost-benefit analysis of the Preschool Curriculum Comparison Study is a logical extension of the idea of linking investments in human capital to later productivity (Heckman, 2006; Knudsen *et al.*, 2006). We are able to test this research question due to the random assignment design utilized in the study, longitudinal nature of the data (with the latest wave collected at age 23), and very strict adherence to curriculum models (e.g., all three programs had similar funding, similar salaries, same class schedules, etc.), so that differences in child outcomes could be related to differences in intervention approaches.

We calculated returns both to the individuals and society (in present dollar terms) that each of the three curriculum models has produced. The question is whether these benefits result equally from all curriculum approaches or more from some than from others. This research uses the data from Preschool Curriculum Comparison study participants from age 3 through age 23.

Hypotheses

We hypothesize that curriculum models that are based on child-initiated learning activities (the HS model and the theme-based NS model) will produce higher benefits 20 years later both to individuals and the public.

Methods

Sample

The study sample consisted of 68 children (65% African-American, 35% White) who lived in families of low socioeconomic status; and had low scores on the Stanford-Binet Intelligence Scale (Terman & Merrill, 1960) at age 3 (see Tables 1-2). Before random assignment to one of the three conditions, children's parents provided information on (a) parental occupation (father's occupation in two-parent families or the only parent's occupation in single-parent families was coded from 1 to 5 - unemployed [1], unskilled [2], semiskilled [3], skilled [4], or professional [5]); (b) parents' highest year of schooling (this was the highest year completed by the mother in single-parent families or the mean of the highest years completed by both parents); and (c) number of rooms per person in the household. The scores on each of the three measures were divided by the sample's standard deviation on that measure to give the three factors equal weighting. The transformed rooms-per-person score was then divided by 2 to give it half the weight of the other two factors. Families scoring a total of 11 or less on the occupation, schooling, and rooms-per-person measures were judged to be living in poverty and eligible for the study sample. These eligible 3-year-olds also completed the Stanford-Binet Intelligence Scale (Form L-M; Terman & Merrill, 1960) with those scoring between 60 and 90

and with no evidence of physical disability being the study sample. The mean IQ for the sample at age 3 was 78 (Schweinhart & Weikart, 1997a).

Several children who were invited to participate in the study declined due to attending other early childhood programs. Few children dropped out of the study because they moved out of the district. Two children, after initial selection, were judged ineligible for the study because their parents' socioeconomic level had risen past the poverty level.

The children who met the entry criteria and completed the preschool programs---a total of 68---served as the original study sample. The oldest class (first cohort) had 27 children in it; the next-oldest (second cohort) had 19; and the youngest (third cohort) had 22. The original study sample had the following characteristics:

- 65% were African-American, and 35% were White.
- 54% were females, and 46% were males.
- 32% of the children lived in single-parent families, and 68% lived in two-parent families.
- 98% of the children in the two-parent families had fathers who were employed outside the home.
- 33% of the children had mothers who were employed outside the home.
- Fathers on average had completed grade 9.
- Mothers on average had completed grade 10.

- Households averaged 6.7 residents, 1.1 per room (Schweinhart & Weikart, 1997a).

The 52 participants who were found and interviewed at age 23 had characteristics that were quite similar to those of the original study sample.

Table 2. Demographic Info

		n	Blacks	Whites	SES (entry)	IQ (entry)	Sex Ratio (M:F)	
Wave 5 (1967)	High/Scope (HS)	11	11	0	8.64 (1.53)	74.55 (6.09)	4:7	At age 3, these children were at the treatment group of Perry Program.
	Direct Instruction (DI)	8	8	0	8.10 (0.73)	73.88 (5.77)	3:5	At age 3, these children were at the control group of Perry Program.
	Nursery School (NS)	8	4	4	8.30 (0.69)	76.50 (4.87)	2:6	Children newly selected at age 4 in 1967.
Wave 6 (1967-1968)	High/Scope (HS)	5	1	4	10.05 (0.70)	83.20 (5.36)	3:2	Children newly selected at age 3 in 1967.
	Direct Instruction (DI)	7	3	4	8.84 (0.60)	85.29 (2.43)	5:2	Children newly selected at age 3 in 1967.
	Nursery School (NS)	7	3	4	8.89 (1.15)	74.57 (7.46)	5:2	Children newly selected at age 3 in 1967.
Wave 7 (1968-1969)	High/Scope (HS)	6	5	1	8.68 (0.77)	78.33 (7.47)	2:4	Children newly selected at age 3 in 1968.
	Direct Instruction (DI)	8	4	4	9.10 (1.25)	78.00 (6.72)	3:5	Children newly selected at age 3 in 1968.
	Nursery School (NS)	8	5	3	9.33 (0.92)	84.25 (3.81)	4:4	Children newly selected at age 3 in 1967.

Random Assignment Procedures

All eligible participants in each of three annual cohorts were first randomly assigned to one of the three groups. After initial assignment, they were reassigned from one group to another until each group had about the same percentages of Blacks and Whites, percentages of boys and girls, and mean Stanford-Binet IQ. Finally, each of the three groups was randomly assigned to one of the curriculum models---DI, HS, or NS. To avoid confounding the effects of different curriculum models within any one family, 9 of the 68 children were reassigned so they would experience the same curriculum model as an older sibling; this in effect, changed the sampling units from children to families (Schweinhart & Weikart, 1997a).

Wave 5 children attended the HS program at ages 3 and 4, while the other subgroups attended the DI and NS programs at age 4 only (see Table 2). Wave 6 and Wave 7 children attended the three programs at both age 3 and age 4. In each program each school year, 3-year-olds from one cohort attended with 4-year-olds from the previous cohort. Wave 8 children were not included in the study sample as they attended the three different programs at age 3 but all attended the HS program at age 4.

Curriculum Models

Once children were **randomly** assigned to one of the three curriculum models, the models were implemented independently and to high standards, in 2½-hour classes 5 days a week and biweekly 1½-hour home visits. This study, therefore, operated three high-quality preschool programs that differed only in the curriculum models employed. The HighScope Preschool Curriculum Comparison study collected data annually from ages 3 to 8 and at ages 10, 15, and 23. Findings are presented for curriculum-group differences through age 23 in education, household and income, community activities, and criminal arrests (Schweinhart & Weikart, 1997a).

The curriculum models used in this study represent three distinct theoretical approaches to early childhood education, three visions of what early childhood education ought to be (Table 3). These approaches differ with respect to the degree of initiative expected of the child and the degree of initiative expected of the teacher---whether the child is primarily initiator or respondent and whether the teacher is primarily initiator or respondent (Weikart, 1972). At the same time, each of these models have been implemented across multiple cultures and in various contexts (Montie *et al.*, 2006).

Table 3. Four Theoretical Approaches to Preschool Education

		Teacher Initiates Much		
Child Initiates Little		Direct Instruction	HighScope	Child Initiates Much
		Custodial Care	Nursery School	
		Teacher Initiates Little		

The Direct Instruction curriculum model, developed by Bereiter and Engelmann (Bereiter, 1986), taught academic skills---specifically, the skills and content assessed by intelligence and achievement tests. Teachers led small groups of children in precisely planned, 20-minute, question-and-answer lessons in language, mathematics, and reading. In the **theme-based Nursery School curriculum model** (Sears & Dowley, 1963), the teachers organized class activities, discussions, and field trips around broad units or themes, such as community helpers, circus animals, and holidays. Children had freedom to choose activities, move from one activity to another, and interact with adults and peers. The emphasis was on developing social skills rather than intellectual skills. In the **HighScope curriculum model**, developed by Weikart and his associates (Hohmann *et al.*, 1979; Hohmann & Weikart, 1995, 2002), adults engaged children as active learners and arranged their classrooms in discrete, well-equipped interest areas. Each day, children planned, carried out, and reviewed their own activities; engaged in small- and large-group activities; and spent time outdoors. Teachers facilitated intellectual, social, and physical key experiences (the domains of children's initiative, social relations, creative representation, music and movement, language and literacy, and the logical and mathematical operations).

Findings

For a decade, virtually no curriculum group differences in intellectual and academic performance were found. Through age 10, the only curriculum group difference on intellectual tests was that the Direct Instruction group had a significantly higher mean IQ on the Stanford-Binet Intelligence Test than did the Nursery School group at the end of the preschool program at age 5 (Weikart *et al.*, 1978). In many areas, no statistically significant differences were found at age 15 or at age 23; however, a pattern of group differences in community behavior did

emerge at age 15 and became more pronounced at age 23. At age 15, for example, the Direct Instruction group reported committing 2½ times as many acts of misconduct as the High/Scope group. At age 23, compared to the other curriculum groups, the Direct Instruction group had three times as many felony arrests per person, especially those involving property crimes; 47% of the Direct Instruction group was treated for emotional impairment or disturbance during their schooling, as compared to only 6% of either of the other curriculum groups.

Cost-Benefit Analyses

Any computation of the lifetime rate of return to the Curriculum Comparison Study must address the following challenges: (a) the randomization protocol was compromised, (b) the lack of any data past age 23 and the need to extrapolate out-of-sample to obtain earnings profiles past that age to estimate lifetime impacts of the program, (c) missing data for participants prior to age 23, and (d) the difficulty in assigning reliable values to non-market outcomes such as crime.

- We account for compromised randomization in evaluating the three programs.
- We develop standard errors for all of our estimates and for the benefit-to-cost ratios accounting for components of the model by which standard errors can be reliably determined. We also examine the sensitivity of estimates of rates of return to plausible ranges of assumptions.
- We present estimates that adjust for the deadweight costs of taxation. Previous estimates ignore the costs of raising taxes in financing programs.
- We use a much wider variety of methods to impute within-sample missing earnings than have been used in the previous literature, and examine the

sensitivity of our estimates to the application of alternative imputation procedures that draw on standard methods in the literature on panel data.

- We use local data on costs of education, crime, and welfare participation whenever possible, instead of following earlier studies in using national data for these components of the rate of return (see Table 15 for an example).

Lifetime benefits and costs through age 23 are directly measured using follow-up interviews. Extrapolation can be used to extend these profiles through age 65. Alternatively, we also compute rates of return through age 23. The scope of our evaluation is confined to the costs and benefits of education, earnings, criminal behavior, tax payments, and reliance on public welfare programs. There are no reliable data on health outcomes, marital and parental outcomes, the quality of social life, and the like. It is likely that our estimated rate of return understates the true rate of return. We present separate estimates of rates of return for private benefits and social benefits.

Initial Cost

We use estimates of the initial program costs for the Perry Preschool Program presented in Barnett (1996) and Heckman, Moon, Pinto, Savelyev, and Yavitz (2010), which preceded the Curriculum Comparison Study. These include both operating costs (teacher salaries and administrative costs) and capital costs (classrooms and facilities). In undiscounted year 2006 dollars, costs of the program per child were \$17,759. We assumed that all three curriculum programs incurred the same initial cost.

Benefits

Education. Early childhood education programs promote educational attainment mainly in two ways: by increasing the total years of education attained and by speeding up how quickly one progresses to a given level of education. In this section, we report estimates of tuition and other pecuniary costs paid by individuals to regular K-12 educational institutions, colleges, and vocational training institutions, and the additional social costs incurred to educate them.² The amount of educational expenditure that the general public spends are larger if a person attains more schooling or if they progress through school less efficiently or by help of some extra special services.

K-12 Education. To calculate the cost of K-12 education, we assumed that all participants went to public school at the cost of the annual per pupil expenditure in the state of Michigan for 1968-1987, \$6,645.³

GED and Special Education. Our estimates of the private costs of K-12 education include the cost of getting a GED.⁴ Special services required spending extra resources. To calculate this cost, we used estimates from Chambers, Parrish and Harr (2004), who

² All monetary values are in year-2006 dollars unless otherwise specified. Social costs include the additional funds beyond tuition paid required to educate students.

³ “Total expenditure per pupil in public elementary and secondary education” for years 1968-1987, as reported by the Digest of Education Statistics (1968-1987, each year in year-2006 dollars). We assume that public K-12 education entails no private cost for individuals. Detailed per pupil expenditures for Ypsilanti schools are not available for the relevant years.

⁴For detailed statistics about the GED, see Heckman and LaFontaine (2008).

provided a historical trend of the ratio of per-pupil costs for special and regular education.⁵

2- and 4-Year Colleges. To calculate the cost of college education, we used each individual's record of post-secondary school attendance collected at the age-23 interview. We computed tuition costs paid by individuals and the expenditure paid by the institutes, taking into account the type of schools attended.⁶

Vocational Training. Some participants also attended vocational training programs. We assumed that all costs were paid by the general public. Estimates by Tsang (1997) suggest per-trainee costs which are 1.8 times the per-pupil costs of regular high school education. Using the number of months each participant attended a vocational training institute, we calculated individual costs. Table 16 summarizes the components of our estimated educational costs through age 23. Table 17 presents educational attainment at age 23 interview by the type of curriculum group.

Employment and Earnings. At the age 23 interview, study participants were asked to provide information about their employment history and earnings at each job for 5 previous jobs. Three problems arose. First, for people with high job mobility, some past jobs were unreported.

⁵ In 1968-69, this ratio was about 1.92; in 1977-78, it was 2.17. Since subjects of curriculum comparison study attended K-12 education mostly in between these two periods, we simplify the ratio to 2 and apply it to all K-12 schooling years.

⁶ The total cost is the sum of private tuition and public expenditure. For student-paid tuition costs at a 2-year college, we used the 1985 tuition per credit hour for Washtenaw Community College (\$29); and for a 4-year college, that of Michigan State University for the same year (\$42). To calculate public institutional expenditure per credit hour, we divided the national mean of total per-student annual expenditure estimated by the National Center for Education Statistics (1991) by 30, a typical credit-hour requirement for full-time students at U.S. colleges. This calculation yielded \$590 per credit hour for 2-year colleges and \$1,765 for 4-year colleges.

Second, for people who were interviewed in the middle of a job spell, it might have not been possible to precisely specify the end point of the job spell. Third, even when the dates for each job spell were precisely specified, it was not possible to identify how earning profiles evolved within each job spell because an interviewee reported only one value of earnings for each job. In order to generate complete earnings profiles, it was necessary to impute missing values. We used a simple piece-wise linear interpolation, based on weighted averages of the nearest observed data points around a missing value. This approach was also used by Belfield et al. (2006) and by Heckman, Moon, Pinto, Savelyev, and Yavitz (2010) for sensitivity analysis. Table 18 summarizes the estimated earnings for each curriculum group by age and gender through age 23. For truncated spells,⁷ we first imputed missing employment status with the mean of the corresponding gender-treatment data from the available sample at the relevant time period, and then we interpolated.

It appears that HS males are the ones that benefit the most in terms of gross earnings. In the case of Females, the NS group is the one that benefit the most (see Table 18). Observe that NS females are also the ones that spend the most in schooling according to Table 16. Even though the earnings pattern is very striking, we have to acknowledge that the age of participants is still low to assess major earning patterns. If the earning series maintain this tendency for later ages, we are likely to find statistically significant results for NS females and HS males, when compared to other curricula. The small sample of data and short earnings spell harms the statically significance of these earnings results.

⁷As noted in the previous footnote, there are job spells in progress at the time of interview.

The earnings presented include all types of fringe benefits listed in the Employer Costs for Employee Compensation (ECEC), a Bureau of Labor Statistics (BLS) compensation measure. Even though the share of fringe benefits in total employee compensation varied across industries, we assumed it to be constant at its economy-wide average regardless of industry, given the data limitations.⁸

Higher earnings translate into higher absolute amounts of income tax payments (and consumption tax payments) that are beneficial to the general public. Since U.S. individual income tax rates and the corresponding brackets have changed over time, in principle we should apply relevant tax rates according to period, income bracket, and filing status. In addition, most wage earners must pay the employee's share of the Federal Insurance Contribution Act (FICA) tax, such as Social Security tax and Medicare tax. In 1978, employees' marginal and average FICA tax rate for a four-person family at a half of US median income was 6.05 percent of taxable earnings. It gradually increased over time, reaching 7.65 percent in 1990, and has remained at that level ever since.⁹ We simplified calculations by applying a 15% individual tax rate and 7.5% FICA tax rate to each participant's taxable earnings. Table 18 shows how individual gross earnings were decomposed into net earnings and tax payments under this assumption.

Crime. For each participant, the Curriculum Comparison Study data provided a full record of arrests, convictions, charges and incarcerations for the years 1978-1987

⁸ The share of fringe benefit has fluctuated over time with the historical average of about 30%. We applied the economy-wide average share at the corresponding year to each person's earnings assuming all fringe benefits tax-free, given data limitation.

obtained from administrative data sources. When we considered the curriculum effect on crime cost reduction through age 23, the main empirical challenge was how to assign values to the impact of criminal activity. To address this challenge, we used national crime statistics published in the Uniform Crime Report (UCR), which were collected by the Federal Bureau of Investigation (FBI) from state and local agencies nationwide. The UCR provided arrest rates by gender, race, and age for each year. We applied population rates to estimate missing crime. Estimating the impact of the program on crime required estimating true levels of crime at each age and obtaining reasonable estimates of the social cost per crime.

Suppose that V_t^c is the total social cost of type c crime at time t . It can be calculated as a product of the social cost per unit of crime C_t^c and the incidence I_t^c :

$$V_t^c = C_t^c \times I_t^c$$

We did not directly observe the true incidence level I_t^c . Instead, we only observed each participant's arrest record at age t for crime c , A_t^c . If we knew the incidence-to-arrest ratio I_t^c/A_t^c from other data sources, we could estimate V_t^c by multiplying the three terms in the following expression:

$$V_t^c = C_t^c \times \frac{I_t^c}{A_t^c} \times A_t^c$$

To obtain the incidence-to-arrest ratio I_t^c/A_t^c for each crime of type c at time t , we used two national crime datasets: the Uniform Crime Report (UCR) and the National Crime Victimization

⁹ See Tax Policy Center (2007).

Survey (NCVS).¹⁰ The UCR provides comprehensive annual arrest data between 1977 and 2004 for state and local agencies across the U.S. The NCVS is a nationally representative household-level data set on criminal victimization, which provides information on unreported crime levels across the U.S. By combining these two sources, we calculated the incidence-to-arrest ratio for each crime of type c at time t . As noted by the Federal Bureau of Investigation (2002), however, the crime typologies derived from the UCR and those of the NCVS are “not strictly comparable.” To overcome this problem, we developed a unified categorization of crimes across the NCVS, UCR, and the Curriculum Comparison Study data sets for felonies and misdemeanors. Heckman and others (2010) used the same incidence/arrest ratios for these crimes.

To check the sensitivity of our results to the choice of a particular crime categorization, we used two sets of incidence/arrest ratios and compared the results. For the first set, we assumed that each crime type had a different incidence-to-arrest ratio. For the second set, we used two broad categories, violent versus property crime. Further, to account for local context, we calculated ratios using UCR/NCVS crime levels that were geographically specific to the Perry program: only crimes committed or arrests made in Metropolitan Sampling Areas (MSAs) of the Midwest.¹¹ Using a simplified version of a decomposition developed by Anderson (1999) and Cohen (2005), we divided crime

¹⁰ The Federal Bureau of Investigation website provides annual reports based on UCR (<http://www.fbi.gov/ucr/ucr.htm>). NCVS are available at Department of Justice website (<http://www.ojp.usdoj.gov/bjs/cvict.htm>).

¹¹ For this purpose, the Midwest is defined as Ohio, Michigan, Indiana, Illinois, Wisconsin, Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, and Kansas. The City of Ypsilanti, where the Perry Preschool Program was conducted, belongs to Detroit MSA.

costs into victim costs and Criminal Justice System (CJS) costs, which consist of police, court, and correctional costs. To obtain total costs from victimization levels, we used unit costs from Cohen (2005). Different types of crime were associated with different victimization unit costs. Some crimes are not associated with any victimization costs. Heckman et. al.(2010) summarize the unit cost estimates that we used for different types of crime. Police, court, and other administrative costs were based on Michigan-specific cost estimates per arrest calculated from UCR and Expenditure and Employment Data for the Criminal Justice System (CJEE) micro datasets.¹² Since we only had data on observed arrests, and did not know whether and to what extent the courts were involved (for example, whether there was a trial ending in acquittal), we assumed that each arrest incurred an average level of all possible police and court costs. This unit cost was applied to all observed arrests regardless of crime types. Estimating correctional costs was a more straightforward task since the data included a record of incarceration and parole/probation for each participant. To estimate the unit cost of incarceration, we used expenditures on correctional institutions by Michigan state and local governments divided by the total institution population. To estimate the unit cost of parole/probation, we performed a similar calculation. Table 19 summarizes our estimated social costs of crime.

¹² From Bureau of Justice Statistics (2003), we obtain total expenditures on police and judicial-legal activities by federal, state and local governments. We divide the expenditures from Michigan state and local governments by the total arrests in this area obtained from UCR. To account for federal agencies' involvement, we add another per-arrest police/court cost which is calculated by dividing the total expenditure of federal government with the total arrests at

In the Curriculum Comparison Study data through age 23, only one murder was reported, which was committed by a male in Direct Instruction group. To avoid having this single event dominate the whole computation, we used two different values of the victim cost of murder: an estimate of “the statistical value of life” (\$4.1 million) and an estimate of assault victim cost (\$13,000).¹³ We reported separate benefit-to-cost ratio estimates for each value, and compared the results. In addition, we assumed that there were no victim costs associated with “driving misdemeanors” and “drug-related crimes”, which we considered to be “victimless”. Although such crimes could be the proximal cause of victimization, such victimizations would be more directly associated with other crimes for which we already explicitly accounted.¹⁴ Table 19 presents our estimates of social costs of crime through participants’ age 23.

Deadweight Cost of Taxation

Finally, we adjusted all benefit streams for deadweight cost of taxation, which is dollars of welfare loss per tax dollar. Since different studies suggest various estimates of the size of the deadweight cost of taxation, in this paper we show the results for various assumptions for the size of deadweight loss associated with \$1 of taxation: 0, 50, and 100 percent.¹⁵ There are many items affected by this consideration: initial program cost, school expenditure paid by the

the national level. This calculation is done for years 1982, 1987, 1992, 1997 and 2002. For periods between selected years, we use interpolated values.

¹³ See Cohen (2005) and, for a literature review, see Viscusi and Aldy (2003), who provides a range of \$2-9 million for the value of a statistical life.

¹⁴ “Driving misdemeanors” include driving without a license; suspended license; driving under the influence of alcohol or drugs; other driving misdemeanors; failure to stop at accident; improper license plate. “Drug-related crimes” include drug abuse, sale, possession, or trafficking.

general public, and all kinds of criminal justice system costs such as police, court, and correctional costs.¹⁶ The early programs incurred the deadweight costs associated with initial funding but saved the deadweight costs associated with taxes used to fund transfer recipients. All benefit-to-cost ratios reported in the next section were adjusted for this consideration.¹⁷

Differences in Benefit-to-Cost Ratios across Curriculum Groups

The differences between groups are reported in Tables 3-14. Our preliminary estimates suggest that at age 23 the differences in means (see Tables 3-14) were not significant across the curriculum groups. Below are statistically significant differences between curriculum groups on variables studied at age 23 (with males and females combined). These findings are adjusted for the alpha level using the stepdown procedure (they are therefore different than what has been reported in Age 23 monograph (Schweinhart & Weikart, 1997a).

HS vs DI (Tables 3-6)	HS vs NS (Tables 7-10)	NS vs DI (Tables 11-14)
Work hard	Work hard	
Vote for president	Vote for president	
Car accident	Car accident	
Family acceptance		Family acceptance
Vote for president	Vote for president	
Car accident		

¹⁵ See Browning (1987); Feldstein (1999); Heckman and Smith (1998); Heckman, LaLonde and Smith (1999); Heckman, Moon, Pinto, Savelyev, and Yavitz (2010) for discussion of estimates of deadweight costs.

¹⁶ We do not apply this adjustment to income tax paid by subjects to avoid double counting. Observed earnings are already adjusted for taxes.

¹⁷ In principle, since different types of taxes are levied by different jurisdictions that create different deadweight losses, we should account for this feature of the welfare system. In practice, this is not possible.

HS vs DI (Tables 3-6)	HS vs NS (Tables 7-10)	NS vs DI (Tables 11-14)
Alcohol use	Alcohol use	
Books	Books	
Book frequency	Book frequency	
Self-esteem		
Hospital visits last year	Hospital visits last year	
Smoked marijuana		Smoked marijuana
Drug use		Drug use
	Have you ever saved any money?	
	ITPA scores at end of 1st year	
	verbal expression at kindergarten	
		Length at welfare
		Follows plans
		WISC Coding at grade 4

At least 30 such significant differences were found and remained after controlling for multiple comparisons was taken into account at age 23. Accounting for multiple hypothesis testing before conducting cost-benefit analysis is an important first step in understanding study findings.

Table 20 presents a summary of benefit streams for each gender-curriculum groups through age 23 under a preferred set of assumptions. In the absence of the “ideal” control group, it was not

possible to compute benefit-to-cost ratio or rate of return for each group separately. Instead, in Tables 21, 23, 24, 25 we show the estimated differences between benefit-to-cost ratios for each pair of curriculum groups. The linear feature of benefit-to-cost ratio enables this comparison.

Suppose that we had an “ideal” control group that received no preschool program so that we could compute the benefit-to-cost ratio, R^j , for a curriculum group, $j \in HS, DI, NS$, by

$$R^j = \sum_{t=1}^T \frac{Y_t^j - Y_t^{Control}}{(1 + \bar{r})^t} / \bar{C}_0$$

where a fixed amount \bar{C}_0 denotes the initial program cost, \bar{r} is a given discount rate, and Y_t^j denotes the net benefit stream at the group j subject’s age t . By the linear feature of this formula, we can easily compare two different curriculum groups simply by subtracting R^j and $R^{j'}$ for any $j \neq j'$ because

$$R^j - R^{j'} = \sum_{t=1}^T \frac{Y_t^j - Y_t^{j'}}{(1 + \bar{r})^t} / \bar{C}_0$$

which captures the extra return to the group j in addition to that to group j' .¹⁸ In Tables 24-27, we present these differences between benefit-to-cost ratios and the associated standard errors for each pair of curriculum groups under various assumptions on the discount rates and the deadweight cost of taxation. Standard errors are computed through a bootstrapping method.¹⁹

¹⁸ This comparison is not possible for internal rates of returns given its non-linear feature. The linearity of benefit-to-cost ratio is the key for this comparison in the absence of an “ideal” control group.

¹⁹ We use 100 replications in this computation, where each replication had the exactly same composition of gender-curriculum groups as the original data.

In most of cases, we could not find statistically significant differences between benefit-to-cost ratios because of huge standard errors. This *discouraging* result could be due to the fact that age 23 is *too early* to see significant differences between curricula. As noted by Heckman et. al. (2010), major effects of early childhood education occur in participants' later life in the case of Perry Preschool Program, and difference in early life trajectories are offset when we convert them to a pecuniary metric.²⁰

While Tables 21-24 already show some sensitivity analyses for the deadweight loss of taxation, the discount rate, and the crime cost estimates, the results from additional sensitivity analyses are presented in Table 25 as well, which compares the results obtained from local cost estimates versus those at the national level. Educational expenditure and all types of Criminal Justice System costs were affected by this consideration. While we observed some difference in the presented numbers, the basic pattern did not change. We did not see any statistically significant differences between benefit-to-cost ratios regardless of gender and curriculum groups compared.

The cost-benefit analysis estimates for the HighScope Perry Preschool Longitudinal Study (see Table 22) at age 23 were lower than at age 40. They ranged from .58 to .84 for typical discount rates (3-5%) at 50% deadweight loss of taxation. Yet at age 40 estimated social rates of return were 7-10% (Heckman, et. al., 2010). The cost-benefit estimates from the age 23 Curriculum Comparison Study are already higher than those observed for the HighScope Perry Study. It

²⁰ For example, pecuniary gains from more efficient progress of female schooling are offset by more educational expenditure due to their greater educational attainment whose effects are realized in other forms in their later life. For more detailed discussion, see Heckman et. al. (2010).

could be the case that in young adulthood crime effects are not as pronounced as later in life. Data collected at midlife are needed to see how cost-benefit analysis ratios changed from age 23 to later life outcomes.

Conclusions

The MacArthur policy initiative “The Power of Measuring Social Benefits” sparked this research agenda to compare economic effects of different teaching approaches. A cost-benefit methodology compared the initial cost of investment in different types of early childhood curriculum to benefits to individuals and the public later in life. We identified three types of interventions that were widely used around the world (i.e., Direct Instruction, HighScope, and traditional theme-based Nursery School approach) and calculated the cost-benefit ratios for each model based on the longitudinal data from the experimentally designed Preschool Curriculum Comparison Study data. This has never been attempted before in economics, education, or psychology. We were able to test this research question due to the random assignment design used in the study, the longitudinal nature of the data (with the latest wave collected at age 23), and very strict adherence to curriculum models (e.g., all three programs had similar funding, similar salaries, same class schedules, etc.) so that differences in child outcomes could be related to differences in intervention approaches. At least 30 such significant differences were found and remained after controlling for multiple comparisons was taken into account at age 23. All of them favored the HighScope group and the Nursery School group over the Direct Instruction group. While high quality early childhood interventions paid off by early adulthood, our preliminary estimates at age 23 the type of curriculum used did not lead to higher savings.

We examined the cost-benefit analysis estimates for the HighScope Perry Preschool Longitudinal Study (see Tables 21-25). The cost-benefit estimates from the age 23 Curriculum Comparison Study are already higher than those observed for the HighScope Perry Study at age 23. It could be the case that in young adulthood crime effects are not as pronounced as later in life. Given the pattern of findings, we expect identifying specific types of interventions that produce higher, longer-lasting outcomes as participants age. Data collected at midlife are needed to see how cost-benefit analysis ratios changed from age 23 to later life outcomes.

This line of work is innovative and impacts multiple fields (i.e., economy, psychology, education) for four reasons:

- 1) Investing in human capital is crucial. What to invest in and how high the return of such investments is an empirical question we need to answer. We calculated economic returns both to the program participants and to the public in present dollar terms that targeted interventions have produced.

- 2) The dataset used for this project is unique. Preschool curriculum comparison studies are rare enough. They are difficult to conduct well, with tight experimental design. No other study has randomly assigned children to curriculum models and then tracked them to age 23 as the Preschool curriculum Comparison Study has. These features place it in a unique position to examine the economic effects of the three types of the curriculum that appear later in life. We are now seeking financial support to track participants into their midlife.

- 3) Most previous cost-benefit analyses compared returns on investments between those participating in a targeted intervention and those who did not participate in an intervention (e.g., HighScope's Perry study). This study went a step further to calculate returns on

investment produced by different curriculum models. This has never been attempted before in economics, education, or psychology.

4) We developed a methodology that allowed us to develop techniques to correct for random assignment procedures, small sample size, and multiple hypothesis testing before conducting cost-benefit analysis. Statistically, we simultaneously resolved issues related to compromised randomization and uncertainty of treatment assignments (an often overlooked phenomenon that a lot of high quality randomized experiments face due to practical considerations) and cost-benefit analysis methodology (by accounting for locally determined costs, missing data, the deadweight costs of taxation, and the value of non-market benefits and costs.) In addition, we strengthened our cost-benefit analysis findings by presenting standard errors of the estimates and by exploring the sensitivity of estimates to alternative assumptions about missing data and the value of non-market benefits.

Most cost-benefit analyses we reviewed do not address issues related to randomization procedures or sample sizes. We developed methodology that outlines ways to conduct cost-benefit analysis accounting for randomization procedures, small sample size, and multiple hypothesis testing. We hope other researchers will be able to utilize this methodology in their studies. The return on investment was calculated for three major types of early childhood educational programs that are widely used around the world (i.e., Direct Instruction, HighScope, and the traditional Nursery School approach). We considered how the results would uphold if the programs were to be implemented under current economic conditions under various assumptions about interpolation, extrapolation, and deadweight losses.

References

- Anderson, D. A. (1999). The Aggregate Burden of Crime. *Journal of Law and Economics*, 42(2), 611-42.
- Anderson, M. Forthcoming. Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool and Early Training Projects. *Journal of the American Statistical Association*
- Barnett, W. S. (1996). *Lives in the Balance: Age 27 Benefit-Cost Analysis of the High/Scope Perry Preschool Program*. Ypsilanti, MI: High/Scope Press.
- Belfield, C.R., Nores, M., Barnett, W. S., & Schweinhart, L. (2006). The High/Scope Perry Preschool Program: Cost-Benefit Analysis using Data from the Age-40 Follow up. *Journal of Human Resources*, 41(1), 162-90.
- Bereiter, C. (1986). Does direct instruction cause delinquency? *Early Childhood Research Quarterly*, 1(3), 289.
- Bureau of Justice Statistics (2003). Expenditure and Employment Data for the Criminal Justice System Series (Annual). United States Department of Justice.
- Burts, D. C., Hart, C. H., Charlesworth, R., & Kirk, L. (1990). A comparison of frequencies of stress behaviours observed in kindergarten children in classrooms with developmentally appropriate versus developmentally inappropriate instructional practices. *Early Childhood Research Quarterly*, 5(3), 407.
- Chambers, J. G., Parrish, T. B., & Harr, J. H. (2004). What Are We Spending on Social Education Services in the United States, 1999-2000?" Special Education Expenditure Project(SEEP) Report 1. Washington, DC: United States Department of Education, Office of Special Education Programs.
- Cohen, Mark A. (2005). *The Costs of Crime and Justice*. New York: Routledge.
- Federal Bureau of Investigation. 2002. "Crime in the United States." Uniform Crime Reports
- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900-1902.

- Heckman, J. J., Moon, S. H., & Pinto, R. R., Savelyev, P.A., & Yavitz, A. Q. (2010). The Rate of Return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94, 114-128.
- Heckman, J. J., Moon, S. H., Pinto, R. R., Savelyev, P. A., & Yavitz, A. Q. (2008). *A reanalysis of the HighScope Perry Preschool Program*. Unpublished manuscript, University of Chicago, Department of Economics.
- Heckman, J. J., R. P. A. Pinto, A. M. Shaikh, and A. Yavitz (2008). *Compromised randomization and uncertainty of treatment assignments in social experiments: The case of Perry Preschool Program*. Unpublished manuscript, University of Chicago, Department of Economics.
- Heckman, James J. and Jeffrey A. Smith. 1998. "Evaluating the Welfare State." In *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, ed. Steinar Strøm, 241-318. New York: Cambridge University Press.
- Hohmann, M., Banet, B., & Weikart, D. P. (1979). *Young children in action. A manual for preschool educators*. Ypsilanti, MI: High/Scope Press.
- Hohmann, M., & Weikart, D. P. (1995). *Educating young children: Active learning practices for preschool and child care programs (1st ed.)*. Ypsilanti, MI: High/Scope Press.
- Hohmann, M., & Weikart, D. P. (2002). *Educating young children: Active learning practices for preschool and child care programs (2nd ed.)*. Ypsilanti, MI: High/Scope Press.
- Karnes, M. B., Schwedel, A. M., & Williams, M. B. (1983). A comparison of five approaches for education young children from low-income homes. In Consortium for Longitudinal Studies, *As the twig is bent...lasting effects of preschool programs* (pp. 133-170). Hillsdale, NY: Erlbaum.
- Knudsen, E., Cameron, J., Heckman, J. J., & Shonkoff, J. (2006). *Economic, neurobiological, and behavioral perspectives on building America's future workforce*. Paper presented at the Proceedings of the National Academy of Sciences.
- Lehmann, E. L. & Romano, J. P. (2005). *Testing Statistical Hypotheses*, 3rd Ed. New York: Springer Science and Business Media.
- Marcon, R. A. (1992). Differential effects of three preschool models on inner-city 4-year-olds. *Early Childhood Research Quarterly*, 7(4), 517.
- Miller, L. B., & Bizzell, R. P. (1983). The Louisville experiment: A comparison of four programs. In Consortium for Longitudinal Studies, *As the twig is bent...lasting effects of preschool programs* (pp. 171-199). Hillsdale, NY: Erlbaum.

- Montie, J., Schweinhart, L. J., & Xiang, Z. (2006). Preschool experience in 10 countries: Cognitive and language performance at age 7. *Early Childhood Research Quarterly*, 213, 313-331.
- National Center for Education Statistics (1991). Digest of Education Statistics, 1990. Washington, D.C.: U.S. Department of Education.
- Romano, J. P. & Wolf, M. (2005). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, 100(469), 94-108.
- Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, S. W., Belfield, C. R., & Nores, M. (2005). Lifetime effects: The High/Scope Perry preschool study through age 40 (Vol. 14). Ypsilanti, MI: High/Scope Press.
- Schweinhart, L. J., & Weikart, D. P. (1997a). The High/Scope preschool curriculum comparison study through age 23. Ypsilanti, MI: HighScope Press.
- Schweinhart, L. J., & Weikart, D. P. (1997b). The High/Scope preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly*, 12(2), 117.
- Sears, P. S., & Dowley, E. M. (1963). Research on teaching in the nursery school. In Gage, N. L. (Ed.), *Handbook of research on teaching*. Chicago, IL: Rand McNally.
- Stipek, D., Daniels, D., Galluzzo, D., & Milburn, S. (1992). Characterizing early childhood education programs for poor and middle-class children. *Early Childhood Research Quarterly*, 7(1), 1.
- Tax Policy Center (2007). Individual Income Tax Brackets, 1945-2008.
- Terman, L. M., & Merrill, M. A. (1960). *Stanford-Binet Intelligence Scale form L-M: Manual for the third revision*. Boston, MA: Houghton-Mifflin.
- Tsang, M. C. (1997). The Cost of Vocational Training. *International Journal of Manpower* 18(1/2):63-89.
- Viscusi, W. K., & Aldy, J. E (2003). The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World. *Journal of Risk and Uncertainty* 27(1):5-76.
- Weikart, D. P. (1972). Relationship of curriculum, teaching, and learning in preschool education. In Stanley, J. C. (Ed.), *Preschool programs for the disadvantaged*. Baltimore, MD: Johns Hopkins University Press.

Weikart, D. P., Bond, J. T., & McNeil, J. T. (1978). *The Ypsilanti Perry Preschool project: Preschool years and longitudinal results through fourth grade (Monographs of the High/Scope Educational Research Foundation, 4)*. Ypsilanti, MI: High/Scope Press.

Weikart, D. P., Epstein, A. S., Schweinhart, L., & Bond, J. T. (1978). *The Ypsilanti Preschool Curriculum Demonstration Project: Preschool Years and Longitudinal Results*. Ypsilanti, MI: High/Scope Press.

Weikart, D. P., Rogers, L., Adcock, C., & McClelland, D. (1971). *The Cognitively Oriented Curriculum: A framework for preschool teachers*. Urbana, IL: University of Illinois.

[REDACTED]

[REDACTED]



