# Web Appendix: Early Childhood Education*

Sneha Elango
The University of Chicago

Jorge Luis García
The University of Chicago

James J. Heckman
American Bar Foundation
The University of Chicago

Andrés Hojman
The University of Chicago

First Draft: May 30, 2014
This Draft: July 29, 2016

*This research was supported in part by the American Bar Foundation, the Pritzker Children's Initiative, the Buffett Early Childhood Fund, NIH grants NICHD R37HD065072, NICHD R01HD54702, and NIA R24AG048081, an anonymous funder, Successful Pathways from School to Work, an initiative of the University of Chicago's Committee on Education funded by the Hymen Milgrom Supporting Organization, a grant from the Institute for New Economic Thinking (INET) supporting the Human Capital and Economic Opportunity Global Working Group (HCEO)—an initiative of the Center for the Economics of Human Development (CEHD) and Becker Friedman Institute for Research in Economics (BFI). We are very grateful to Marianne Haramoto, Fernando Hoces, Joshua Ka Chun Shea, Matthew C. Tauzer, and Anna Ziff for research assistance and useful comments. We thank Robert Moffitt, David Blau, the other authors of this volume, and Raquel Bernal, Avi Feller, Micheal Keane, Patrick Kline, Sylvi Kuperman, and Rich Neimand for valuable comments. The views expressed in this chapter are those of the authors and not necessarily those of the funders or persons named here or the official views of the National Institutes of Health.

# Contents



# List of Figures

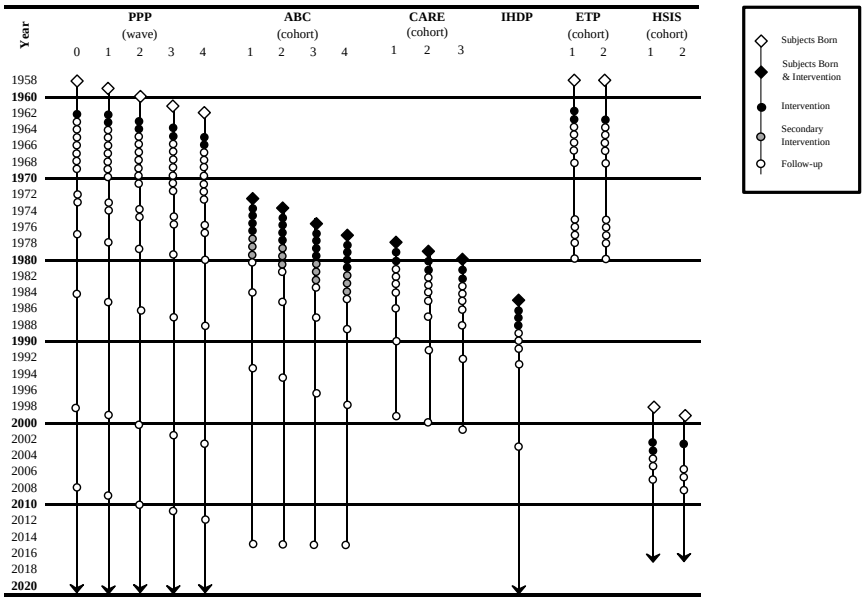


# List of Tables

# A Supplemental Information on Demonstration Programs

This appendix provides detailed descriptions of the program content and implementation for the randomized control trials of Head Start and the demonstration programs discussed in this chapter. Figure A.1 shows the time periods in which the demonstration programs were operating and the randomized control trial of Head Start, Head Start Impact Study (HSIS) (Puma et al., 2012). The demonstration programs discussed in this chapter are: (i) the Perry Preschool Project, or PPP; (ii) the Carolina Abecedarian Project, or ABC; (iii) Carolina Approach to Responsive Education, or CARE; (iv) the Infant Health and Development Project, or IHDP; and (v) the Early Training Project, or ETP.

Figure A.1: A Timeline of Cohorts and Follow-Ups of Early Childhood Education Programs




Note: This figure shows the timeline of birth years, intervention years, and follow-up years for each randomized control trial of an early childhood education program. Each arrow illustrates the data collection stream for each cohort of a program. Lines for PPP and IHDP end at the bottom of the chart, and this shows that follow-ups for these programs will continue after the year 2020. Arrows for HSIS show that the samples will be tracked at least through 2016. Follow-up data on ABC and PPP may continue to be collected. Lines that end with white dots indicate that the final follow-up was carried out and no more official follow-up is planned. The most recent follow-ups for ABC and ETP took place in 2015 and in 1980 respectively. The most recent follow-ups for CARE took place when participants in each cohort reached age 21, which are years 1999, 2000, and 2001. PPP, ABC, CARE, and ETP have follow-ups during intervention years as well. For IHDP, follow-ups took place at the following ages of subjects: 40 weeks, 4 months, 8 months, 1 year, 18 months, 2 years, 30 months, 3 years, 4 years, 5 years, 6.5 years, 8 years, and 18 years. For HSIS, although the chart shows that the first cohort had 2 years of treatment, only those in the age 3 cohort who reapplied to the program in the 2003-2004 program year received 2 years of treatment. "Secondary Intervention" refers to the re-randomization and school-age treatment of ABC, which is excluded from the chapter.

## A.1 Program Description

### A.1.1 Perry Preschool Project

#### Summary

The Perry Preschool Project was conducted in Ypsilanti, MI, with the selection of 123 African-American and low-income children age 3 or 4 (Schweinhart et al., 2005). These children were selected over five cohorts from 1962 through 1965. Of the 123 children that met the study eligibility criteria, 58 were selected to be in the treatment group and 65 in the control group (Schweinhart et al., 2005). The treatment group attended a 2.5-hour preschool session on weekdays during the school year, and received weekly home visits lasting 1.5 hours from their teachers. The duration of the program was two years except for children from the first wave, who were all 4 years of age upon entry and only received treatment for one year (Weikart et al., 1978). The curriculum for the preschool focused on active learning and child-teacher interaction (Weikart et al., 1978). Social skills were cultivated in group reviews of individual tasks. Data on the subjects were collected annually while they were ages 3 through 11, and again at ages 14, 15, 19, 27, and 40. At the age 40 follow-up, missing data was only 6% (Schweinhart et al., 2005), largely due to death.

#### Program Components

The program lasted for two years and included 3-hour weekday sessions and biweekly home visits (Schweinhart, 2003). Teachers acted as guides to child learning, and as the program progressed, the curriculum changed from an experimental phase to include a rigid schedule that gave the children the independence needed to take charge of their own skill development (Schweinhart, 2003, Appendix A). This structure in addition to a low student-teacher ratio (20-25 students for 4 teachers) created an environment that helped the children improve in the defined 10 developmental factors, such as "creative representation" (Schweinhart, 2003).

#### Eligibility Criteria

Children for the study were drawn from the area surrounding Perry Elementary School, and

found through census data, neighborhood-group referral, and canvassing in the neighborhoods (Schweinhart et al., 2005). To be eligible, children needed to be African-American, have an IQ between 70 and 85 (compared to the national mean of 100), and come from a disadvantaged family defined by parental employment, income, and education, as well as housing characteristics (Weikart, 1967).

**Randomization and Attrition**

The eligible children in the first cohort were matched based on IQ score and socioeconomic class, and then randomly placed in treatment or control groups based on the result of a coin toss (Schweinhart et al., 2005; Schweinhart, 2006). Subsequent cohorts were randomized by a more complex protocol. Siblings of children already in the study were placed in the same group as that of their families. The remaining children were ranked by IQ scores (but not socioeconomic class) with the even-ranked children and odd-ranked children separated into two groups. Mean demographic characteristics were balanced between the two groups, and then treatments and controls were assigned to the two groups with equal probability. Some switches between the two groups were made after the assignments based on maternal employment, as employed mothers were less available for home visits than unemployed mothers (Schweinhart et al., 2005).

Table A.1: Perry Attrition

| | **Perry** | Original Sample | Age 19 Data? | Age 27 Data? | Age 40 Data? | Age 50 Data? |
|---|---|---|---|---|---|---|
| Cause of Attrition | Not reached due to imprisonment | 0 | N/A | N/A | N/A | N/A |
| | Death | 0 | N/A | N/A | N/A | 15 |
| | Other attrition | 0 | N/A | N/A | N/A | 6 |
| | **Not interviewed** | **0** | **2** | **6** | **11** | **21** |
| | **Interviewed** | **123** | **121** | **117** | **112** | **102** |
| | Total | 123 | 123 | 123 | 123 | 123 |
| | Sample left | – | 98% | 95% | 91% | 83% |

Note: The question marks indicate numbers that HighScope needs to confirm. At age 50, the 5 in prison are in negotiation to be interviewed. The effective sample size is 102 with 97 already interviewed.

### A.1.2 Carolina Abecedarian Project (ABC)

**Summary**

The Carolina Abecedarian Project (ABC) was a study designed to investigate how intensive early childhood education affects the social and cognitive development of disadvantaged children. In order to do this, experimenters selected 122 children born between 1972 and 1977 who lived in or near Chapel Hill, North Carolina, and randomly placed them in either a center-based intervention group or a control group (Ramey et al., 1976). The treatment and control groups were redefined at age 5 at which point they moved from the preschool to school-age treatment (Campbell and Ramey, 1991). The preschool-age treatment children were given educational games to develop basic skills, while those in the school-age treatment group were introduced to math, science, and music (Ramey and Campbell, 1991). About 96% of the participating families were African-American and 4% were white (Ramey and Smith, 1977). The project led to a spinoff program, the Carolina Approach to Responsive Education (CARE) that is described in Appendix A.1.3.

**Program Components**

The preschool intervention was for children from birth to age 5, and included all-day child care 5 days a week, 50 weeks a year (Campbell and Ramey, 1991). The curriculum focused on language, social, perceptual-motor, and cognitive areas of development (see Ramey et al., 1977; Haskins, 1985; Ramey and Haskins, 1981; Ramey and Campbell, 1979; Ramey and Smith, 1977; Ramey et al., 1982; Sparling and Lewis, 1979, 1984). The program offered nutrition and medical treatment to participants. The treatment group received formula as infants and two meals and a snack daily after the age of 15 months; the control group received formula until the age of 15 months. During the first year of program implementation, treatment and control children both received medical care. The medical staff provided regularly scheduled well-child checkups, immunizations, parental counseling, and initial assessments of illnesses (Ramey et al., 1977). For the duration of the programs and for all following cohorts, only treatment children received medical care. When the children were toddlers

and preschoolers, a licensed practical nurse visited classrooms daily for up to two hours to monitor children's health status (Sanyal et al., 1980).

After age 5, the school-age intervention stage of the study began and the treatment and control groups were redefined after another round of randomization. During this stage, students were introduced to math, science, and music. The school-age phase (grades 1-3) followed, during which treatment-group students received full-day, year-round care with individualized curriculum packets with which parents were familiarized. Additional support was given to the treatment group including transportation and help with paperwork (Ramey and Campbell, 1991). Additionally, the treatment group received home visits conducted by a certified teacher to track the academic and socio-emotional progress of the children. The visit was biweekly in each child's own school and biweekly in his or her own homes. This treatment is not comparable to that of CARE or IHDP because the age does not coincide. This is significant, because in CARE and IHDP, the children were very young in during the visits, leading visitors to interact mostly with the parents.

**Eligibility Criteria**

Recruitment to ABC typically began in the last trimester of pregnancy. Potential families were referred by local social service agencies and local hospitals. Eligibility for inclusion was determined by a score of 11 or more on a weighted 13-factor High Risk Index (Ramey et al., 2000). Table A.2 lists the factors of the index and their respective weights. Of the 122 families that were eligible, 121 agreed to participate. The final sample consisted of 120 families with 122 participating children.

Table A.2: ABC High-Risk Index and Weights

| Factor | Weight |
|---|---|
| Mothers Educational Level (last grade completed) | |
| 6 | 8 |
| 7 | 7 |
| 8 | 6 |
| 9 | 3 |
| 10 | 2 |
| 11 | 1 |
| 12 | 0 |
| Fathers Educational Level (last grade completed) | |
| 6 | 8 |
| 7 | 7 |
| 8 | 6 |
| 9 | 3 |
| 10 | 2 |
| 11 | 1 |
| 12 | 0 |
| Family Income ($ per year) | |
| 1,000 | 8 |
| 1,001-2,000 | 7 |
| 2,001-3,000 | 6 |
| 3,001-4,000 | 5 |
| 4,001-5,000 | 4 |
| 5,001-6,000 | 0 |
| Father absent for reasons other than health or death | 3 |
| Absence of maternal relatives in local area | 3 |
| Siblings of school age one or more grades behind age-appropriate level or with equivalently low scores on school-administered achievement tests | 3 |
| Payments received from welfare agencies within past 3 years | 3 |
| Record of fathers work indicates unstable or unskilled and semi-skilled labor | 3 |
| Records of mothers or fathers IQ indicate scores of 90 or below | 3 |
| Records of sibling's IQ indicates scores of 90 or below | 3 |
| Relevant social agencies in the community indicate the family is in need of assistance | 3 |
| One or more members of the family has sought counseling or professional help in the past 3 years | 1 |
| Special circumstances not included in any of the above that are likely contributors to cultural or social disadvantage | 1 |

Source: Replicated from Ramey et al. (2000). Note: Criterion for inclusion in high-risk sample was a score of more than 11. Base years of the family income criteria change for every cohort; for every year of recruitment, program implementors used nominal-valued income cutoffs.

### Randomization and Attrition

The original sample included 109 families with a total of 111 children, including one set of twins and one sibling pair. The sample was divided into 57 treatment and 54 control children. Of these, 59 were female and 52 were male (Campbell and Ramey, 1994). ABC's attrition rate is reported as 18.9% over the 13-year span from entry of the first cohort until the youngest child reached age 8 years and completed the secondary-phase treatment (Campbell and Ramey, 1989; Ramey and Campbell, 1991; Campbell and Ramey, 1994; Clarke and Campbell, 1998; Campbell et al., 2014). After assignment, seven experimental families and one control family declined to participate in the study. Two control children were dropped

from the control group and added to the treatment group when local authorities requested that they be permitted to attend the day-care center, and two children were dropped due to diagnosis of organic developmental delay. Three families moved from the area, and four children died before age 5 years (Ramey et al., 1984).[1] Prior to 1979, new children were admitted to the study to replace three children who either died or moved away before 6 months of age and have participated in data collection (Ramey and Campbell, 1979).

Table A.3: ABC Attrition

| | **ABC** | Original Sample | Baseline Data? | Age 21 Data? | Age 30 Data? | Health Data?* |
|---|---|---|---|---|---|---|
| Cause of Attrition | Withdrawn from study | 0 | 4 | 5 | 5 | 5 |
| | Not followed; compromised randomization | 0 | 0 | 8 | 8 | 8 |
| | Death | 0 | 0 | 4 | 7 | 7 |
| | Other attrition | 0 | 0 | 1 | 1 | 30 |
| | **Not interviewed** | **0** | **4** | **18** | **21** | **50** |
| | **Interviewed** | **122** | **118** | **104** | **101** | **72** |
| | Total | 122 | 122 | 122 | 122 | 122 |
| | Sample left | 100% | 97% | 85% | 83% | 59% |

Note: This table separates children in the original sample in four categories. Four children left the study after being randomized into the treatment group, but before having data collected on them. A girl randomized into the control group was adopted and retired from the study during preschool. We consider those 5 children withdrawn from the study. Two children were diagnosed as biologically retarded during the preschool round and considered ineligible for the study. Another two children were swapped from the control to the treatment group for being in high risk. The families of children refused their assignment to the treatment group and were considered part of the control group. The family of a control male refused his random assignment, so the child was included into the treatment group. We consider those 8 children as having compromised randomization. 4 children (2 treated, 2 control) died during the preschool phase. 6 more individuals have died since then. 99 individuals compose the rest of the sample. The age-30 data was remarkably successful in finding 98 of them. *The health data attrition (really non-response) come because the screenings were held in one office with limited office hours. The non-responders in this survey are available for further interviews.

### A.1.3 Carolina Approach to Responsive Education (CARE)

Project CARE is closely related to ABC in design and implementation. It sought to further the research on ABC by adding a family-support component to foster mother-child

[1] One treatment infant died at 3 months due to diagnosed crib death at home, and one treatment child died in 1979 at age 50 months in a pedestrian accident. One control infant died at 3 months due to cardiomyopathy and seizure disorder, and one control child died at 18 months due to cardiac arrest (Campbell and Ramey, 1994).

interaction (Wasik et al., 1990). We do not analyze it in the paper, but we add its description because it influenced the design of IHDP. Beginning in 1978, 65 low-income families in a semi-rural county in North Carolina were recruited to participate (Ramey et al., 1981; Wasik et al., 1990).

**Program Components**

There were 3 treatment conditions which differed from the ABC treatment by providing home visits: (i) center-based care and home visits, (ii) just home visits, and (iii) neither. The center-based treatment component used the same curriculum as ABC. A home-visiting component was developed specially for CARE. One month following the child's birth, families in both conditions (i) and (ii) began receiving home visits (Wasik et al., 1990). These visits were designed to be weekly for the first 3 years, but actually averaged 2.5 per month for condition (ii) and 2.7 for condition (i). During year 4 and 5, the frequency varied based on parental preference anywhere from weekly to every 6 weeks (averaged 1.4 per month for condition (ii) and 1.1 per month for condition (i)). The majority of these visits were between 30 and 60 minutes with 20% of them lasting longer than an hour (Wasik et al., 1990). This intervention was based on the belief that many families lack knowledge and skills necessary to positively influence their child's development, and that many families also experience stresses that interfere with effective parenting (Wasik et al., 1990). With the intention of fostering cognitive and social development, home visitors established a caring relationship with the parents so that they could provide support and encouragement, convey information, advocate for the family, promote effective coping, and encourage and model positive parent-child interactions (Wasik et al., 1990; Ramey et al., 1981; Burchinal et al., 1997).

The center-based treatment component used a systematic developmental curriculum very similar to the one used in ABC with the intention of helping the child's development in both cognitive and social domains (Wasik et al., 1990). All children attended the center from 7:30 a.m. to around 3:30 p.m. Some were taken home at 3:30 and others were picked up by their

parents between 3:30 p.m. and 5:30 p.m. The ratio between teacher and child was 1:3 for infants and toddlers, 1:4 for 2 year-olds, and 1:6 for 3–5 year-olds. The primary curriculum resources were *Learningames for the First Three Years* and *Learningames for Threes and Fours*, which contain activities that support the intellectual, creative, and socio-emotional domains (Wasik et al., 1990). These features of the center-based treatment resemble the ABC treatment, though the program benefited from the experience of the ABC staff.

**Eligibility Criteria**

Over an 18-month period that began in 1978, 65 families in a semi-rural county in North Carolina were identified to participate in Project CARE (Heckman et al., 2014; Wasik et al., 1990). Each family agreed to participate in a screening consisting of an interview and psychological assessment during which it was judged whether or not each infant was at an elevated risk for delayed development based on the same High-Risk Index as ABC (see Table A.2).

**Randomization and Attrition**

Once eligibility was confirmed, each of the families were assigned to 1 of 3 experimental conditions: (i) Child Development Center and Family Education (16 families), (ii) Family Education (25 families), or (ii) neither (23 families). One family of the 65 refused to participate once given their assignment (Wasik et al., 1990).

Retention remained relatively high in this program up to the age 21 follow-up, in which information on 60 of the 66 original subjects is available. In the 34 year-old health follow-up, attrition is substantial; the retention rate is 61%. Few evaluations are available for this program. Two exceptions include Campbell et al. (2008) and Campbell et al. (2013). The first paper compares the effects of CARE with the effects of ABC and finds that they are similar with respect to educational and health behavior outcomes. The second paper confirms these findings and notes that they extend to the age 34 follow-up. It also finds an important qualification: of the two branches of treatment offered in CARE, the center-based one is much more effective at boosting outcomes compared to the home-based one.

### A.1.4 The Infant Health and Development Program (IHDP)

**Summary**

The Infant Health and Development Program began in 1985 with a randomized sample of 985 infants. Treatment began after discharge from the neonatal nursery, and continued until 36 months of age. Both treatment and control group received the same medical, developmental, and social assessments. Referrals for pediatric care were provided for both groups at 40 weeks gestational age, and 4, 8, 12, 18, 24, 30, and 36 months Ballard-corrected age[2] (Martin et al., 2008). The treatment group also received home visits, child enrollment at a child development center, and parent group meetings, all for free. Primary outcomes include cognitive development, behavioral competence, and health status. The curriculum and program design of IHDP were based on and expanded from two early childhood education programs: ABC and CARE. Overall, the design of IHDP was much more similar to CARE, because it was intensive both in home visits and center-based care components, while the preschool-age treatment of ABC had no home visits (Gross et al., 1997). However, curriculum components and implementation strategies were based on both programs.

**Program Components**

Home visits were weekly for the first year, and biweekly during the second year. Family support and child health/development information was provided during visits. A curriculum of activities focusing on cognitive, linguistic and social development were provided to parents to be applied to the child. A second curriculum teaching parents how to manage self-identified problems was also provided (Gross, 1990).

Children in the intervention group attended child development centers five days a week, at least four hours a day, from ages 12 months to 36 months (Brooks-Gunn et al., 1995). Activities from home visits were utilized in the child development centers. Teacher-child ratios were low (from 1:3 to 1:4). Transportation services were provided as well (Gross,

[2]This is a measure of age corrected for prematurity using the Ballard assessment, which evaluates a premature infant's physical development.

1990). Sessions for treatment parents began when the child reached 12 months of age. Parents received information on child rearing, health and safety, social support, and other relevant parenting concerns (Gross, 1990).

**Eligibility Criteria**

Eight medical institutions were selected to participate in the study by a national competitive review. The eight sites were the University of Arkansas for Medical Sciences (Little Rock, AR); Albert Einstein College of Medicine (Bronx, NY); Harvard Medical School (Boston, MA; University of Miami School of Medicine (Miami, FL); University of Pennsylvania School of Medicine (Philadelphia, PA); University of Texas Health Science Center at Dallas (Dallas, TX); University of Washington School of Medicine (Seattle, WA); Yale University School of Medicine (New Haven, CT) (Gross, 1990). Upon initial screening at these sites, 4,551 prematurely born infants ($\leq$37 weeks of gestation) who would reach 40-weeks post-conceptional age between January 7, 1985, and October 9, 1985, and were classified as either lighter low birth weight (LLBW, $\leq$2000g) or heavier low birth weight (HLBW, 2001g–2500g), were considered for the study. Of these infants, 3,249 were excluded because they failed to meet additional eligibility criteria: (i) families must reside within 45 minutes driving distance from the center; (ii) infants must have gestational age less than 37 weeks; (iii) infants must not have any severe illness or neurological deficits[3]. Some children were excluded from the study, because they were discharged from the hospital after the screening period but before the recruitment period. This is no additional information available about infants who were excluded from the study. Of the remaining 1,302 who met the eligibility criteria, parents of 241 (21%) of those infants refused to participate, and another 43 infants were withdrawn before participating in the study. Ultimately, 985 infants comprised the primary analysis group.

**Randomization and Attrition**

One third of the sample consisted of HLBW infants, and two thirds consisted of LLBW

[3]There were 61 children excluded due to illnesses or neurological deficits (Gross, 1990).

infants. One third of the subjects within each weight group were randomly assigned to the treatment group, and the remaining two thirds were assigned to the control group. An adaptive randomization method was utilized to monitor the bias between the two study groups. Balance for birth weight, gender, maternal education, maternal race, primary language in the home, and infant participation in another study were also monitored. The targeted number of patients at each of the eight sites was 135 (Gross, 1990).

Retention rates at ages 3, 5 and 8 years of age were 92%, 82%, 89%, and 65%, respectively (McCormick et al., 2006). There is evidence to suggest that attrition was random, and cannot be predicted by pre-treatment characteristics (García, 2015). The non-participation rate of the treatment group for the IQ test at age five was 15% greater than that of the control group.

In the 36-month follow up interview, 93% of families in the primary analysis group were assessed on at least one of the primary outcome measures. However, mothers from families for which we have complete data from the time of randomization to Ballard-corrected age 36-months were more likely to be white (39% vs. 41%), less likely to be teenagers (9% vs. 15%), and have at least a high school education (65% vs. 48%) (Brooks-Gunn et al., 1998).

### A.1.5 Early Training Project (ETP)

**Summary**

The Early Training Project (ETP) was implemented in Tennessee from 1962 to 1964 and targeted 4–5 year-olds to prevent the tendency of low-income children to progressively fall behind in public school (Gray et al., 1982). The children in the study came from disadvantaged families, defined by low income, maternal educational attainment of 8th grade or less, unskilled or semi-skilled occupation or unemployment, and housing characteristics (Klaus and Gray, 1968). The children were all African-American due to racial segregation of schools, and were born in 1958 (Klaus and Gray, 1968).

Cognitive, non-cognitive, demographic, and parental data were collected from 1962 to

1966. In 1976, follow-up data were collected on the parents, as well as on the children as teenagers to gauge how they responded to the drastic social changes that occurred during the decade (Gray et al., 1982). The Consortium on Longitudinal Study continued to collect more data on the families until 1980 (Gray et al., 1982).

**Program Components**

The summer sessions each lasted 10 weeks, and involved an intensive educational environment (Klaus and Gray, 1968). In order to maximize attendance at the summer sessions, teachers fostered relationships with the parents prior to the beginning of the sessions to inform them on the structure and activities. Additionally, transportation was provided to lower the cost of attendance, and the school was made to be attractive for kids. The student-teacher ratio remained low to allow each student to receive maximum attention (4 or 5 children per adult) (Klaus and Gray, 1968).

The program focused on improving positive attitudes related to academic achievement and the cognitive skills required to achieve in this way. Skills and character traits, such as delayed gratification, were highlighted in the curriculum. The home visits encouraged parental involvement in these areas, and worked to positively affect parental attitudes towards academic achievement (Gray and Klaus, 1970). In order to maintain a positive relationship with the control group parents, play sessions were provided for their children twice a week that completely lacked educational instruction. Not only did this placate parents, but it also made it difficult for surveyed counselors later on to know which children were in which group (Gray et al., 1982).

**Randomization and Attrition**

There were initially 88 children (not including one child who died and one who became disabled), 61 of whom were from the locality, and 27 of whom were from a similar town located 60 miles away in order to measure spill-over effects from the experimental group (Gray and Klaus, 1970). In the local town, there were 2 experimental groups and one control group. One of the experimental groups received 3 summers of treatment starting in 1962, and the

other group only received 2 summers starting in 1963 (Klaus and Gray, 1968). Ultimately, no significant difference was found between these two experimental groups, so many analyses pool the samples to improve power. Because the distal group was not randomized from the same sample as the that of the local groups, comparison between the distal control and experimental group cannot be taken at face value (Gray et al., 1982). In 1976, 20 African-American, low-income families with 6-year-old children in the local town were randomly selected from a pool of 30. They were surveyed in such a way to resemble parental responses from the 1962–1964 summer treatments. While perfect comparisons are difficult due to the time gap between the two samples, differences between the data allowed the researchers to understand which aspects of parenting for approximately this population had changed over the decade (Gray et al., 1982). Of the 88 children at the start of the study, about 90% were included in the 1966 and 1968 analysis, and in 1975, data on 86 of the 88 children were collected (Gray et al., 1982). Data on survival rates of subjects is not available.

### A.1.6 Head Start Impact Study (HSIS)

**Summary**

Head Start, a government-run, free preschool education program, began in 1965 and is now the largest early childhood program in the US, enrolling about one million 3 and 4 year-olds annually at a cost of about $8 billion (Administration for Children and Families, Office of Head Start, 2014). The Head Start Impact Study (HSIS) was conducted in order to assess the effects of this national program, using a nationally representative sample of 84 grantee/delegate agencies that supported nearly 5,000 newly entering, eligible 3- and 4-year-old children. These children were placed into either a Head Start group that participated in all the services traditionally provided by the program or a control group that had no initial access to Head Start services (although some found their way in through other means) but were free to enroll in other, non-Head Start services. Data collection ran from fall 2002 to 2008, following the children through the spring of their 3rd grade year (Puma et al., 2012).

Data will continue to be collected at least through 2016.

**Program Components**

Head start provides comprehensive services that include preschool education; medical, dental, and mental health care; nutrition services; and efforts to help parents foster their child's development. While Head Start is based on a holistic model of school readiness, emphasizing both non-cognitive and cognitive development, the program allows for a wide variety of child-care settings and practices (Administration for Children and Families, Office of Head Start, 2014, 2009). Centers differ with respect to teacher characteristics, class size, instructional time, and frequency of home visits (Administration for Children and Families, Office of Head Start, 2014; Puma et al., 2012). Despite these differences, there are a set of program-wide performance standards that all grantee agencies must meet (see Administration for Children and Families, Office of Head Start, 2009).

**Eligibility Criteria**

The Head Start program was designed to service economically-disadvantaged families across the US. The program awards grants to public, private non-profit, and for-profit organizations that provide child-care services to children below 130% of the federal poverty line. That being said, up to 10% of children attending these centers come from households above this income level (Puma et al., 2012). The Head Start poverty guidelines vary state by state. Since the sample of children followed in HSIS are taken randomly from this general pool, the criteria for the study is identical to the larger program. The average household income in the sample is $1,842 per month (2014 USD) for the 3 year old cohort and $1,945 per month (2014 USD) for the four year old cohort.[4]

**Randomization and Attrition**

The study sample chosen was nationally representative, spreading over 23 different states. 84 grantees/delegate agencies were randomly selected. Among those, 383 Head Start centers were randomly selected. Finally, the total of 4,667 newly entering children (2,559 3-year-olds

[4]This is based on own calculations with HSIS data.

and 2,108 4-year-olds) were randomly selected based on a lottery. An average sample of 27 children per center was chosen, 16 assigned to the Head Start group and 11 assigned to the control group. Random assignment was done separately for the newly entering 3-year olds and newly entering 4-year olds (Puma et al., 2012). However, there was not complete compliance with this random assignment. Some children accepted into Head Start did not participate in the program (about 15% for the 3-year-old cohort and 20% for the 4-year-old cohort), and some children assigned to the non-Head Start group found a way to enter the program in the first year (about 17% for 3-year-olds and 14% for 4-year-olds) (Puma et al., 2012). By the end of the second year, about 90% of the Head Start group was in a center-based early childhood program (63% into Head Start), and a comparable percentage of the control group was also in a center-based program (about 50% into Head Start) (Puma et al., 2012).

# B Supplemental Information on Universal and Other Large-Scale Programs

## B.1 An Overview of State Preschool in the US

Most states also have their own programs with substantial funding coming from state and federal resources. National Institute for Early Education Research (NIEER) studies show that more children are enrolled in state-funded preschool than in any other publicly-funded early childhood education programs: 28% of 4-year-olds are enrolled in state-funded programs, 11% in HS, 3% in other public programs, and 3% in special education, not including special education children who are also enrolled in state-funded preschool or Head Start (Barnett et al., 2015). In 2014, 50% of all 3 and 4 year-olds attended preschool. Of the children who attended preschool, 31% attended a public program and 43% attended a private program.[5] In 2014, 50% of all 3 and 4 year-olds attended preschool. Of the children who attended preschool, 57% were enrolled in a publicly operated program, and 43% were enrolled in a privately operated program.[6,7]

There has been tremendous growth in state-funded programs over the last twenty-five years. In 1980, only 4 states subsidized any preschool programs, and by 1987, this number grew to 11. By the mid-nineties, fifteen states subsidized preschool. This number has grown steadily through 2014. In 2014, 26 states had programs only for 4-year-olds, and 10 states had no programs: (i) Hawaii, (ii) Idaho, (iii) Indiana, (iv) Mississippi, (v) Montana, (vi) New Hampshire, (vii) North Dakota, (viii) South Dakota, (ix) Utah, (x) and Wyoming (Barnett et al., 2015). Among the states that offer some preschool program, only nine (and

[5] U.S. Census Bureau, Current Population Survey, October 2014.

[6] U.S. Census Bureau, Current Population Survey, October, 2014. A public school is defined as any educational institution operated by publicly elected or appointed school officials and supported by public funds. Private schools include educational institutions established and operated by religious bodies, as well as those which are under other private control. In cases where enrollment was in a school or college which was both publicly and privately controlled or supported, enrollment was counted according to whether it was primarily public or private.

[7] Data on funding to private schools is limited and is complicated to gather, as many privately operated schools receive some funding from public streams.

Washington D.C.) had greater than 50% enrollment of 4-year-olds in 2014: (i) DC; (ii) Florida; (iii) Oklahoma; (iv) Vermont; (v) Wisconsin; (vi) West Virginia; (vii) Iowa; (viii) Georgia; and (ix) Texas.

## B.2 State Preschool in Georgia and Oklahoma

Georgia and Oklahoma both have state-wide universal programs. This means they provide access to all children aged 4 without any eligibility criteria. However, while access is universal, take-up is not (see Cascio and Schanzenbach, 2013). In Georgia, 59% of preschool-age children in the state took up the program; of these, 65% were eligible for free or reduced price lunch. In Oklahoma, 74% of preschool-age children took up the program; of these, 66% were eligible for free or reduced price lunch. Poverty is defined as household income at or below 200% of the federal poverty line. In Georgia and Oklahoma, 49% of children live in poverty, defined here as family income below the 200% poverty line (United States Census Bureau, 2014, American Community Survey). Given that a child belongs to a poor household, the probability of program take-up is 79% in Georgia and 99% in Oklahoma. Conversely, the probability of program take-up given that a child does not belong to a poor household is 40% in Georgia and 49% in Oklahoma.

Though neither preschool program is homogeneous across the all the centers in the state, they both have approved a comparable set of curriculum materials for use in the classroom. Both programs provide at least one meal per day and have vision, hearing or health support services. They have high staff quality standards, requiring teachers to have a bachelor's degree and have specialized preschool training. They also cap class sizes at 20 students and requires a staff-child ration of 1:10.

They also differ in their funding schemes. Georgia's preschool are predominately run in private centers— only 20% of providers are public schools. In Oklahoma, about 90% of the universal preschool slots are provided by public schools. In 2010-2011, Georgia spent $4,298 on average per child. For the same period, Oklahoma spent $3,461 (Bassok et al., 2014).

## B.3 Boston Public School Prekindergarten Program

The Boston Public School Prekindergarten program has programs for 3 year-olds (called K0) and 4 year-olds (called K1). The evidence presented in this chapter uses data from K1 programs. In the 2008-2009 school year, the program served about 2,045 children in 69 elementary schools. An estimated 34-43% of 4-year-olds in Boston were enrolled in the program (Georgia Department of Early Care and Learning, 2015). Any children age 4 by September first of the school year are eligible to attend preschool. Because the number of children whose families want to enroll them in preschool exceeds the number of children that can be served by the program, children enter a lottery and are randomly chosen to fill openings. Once a child is given a spot in a preschool center, he or she is guaranteed a spot for all of the years of schooling provided at the center. Of the children who were not given a preschool spot through the lottery, about 67% took up non-relative care, and almost 50% took up center-based care (Weiland and Yoshikawa, 2013).

The curricula used in preschool classrooms is similar to the curricula approved for use in Georgia's and Oklahoma's programs. However, there is more homogeneity in Boston programs classrooms, which use the Opening the World of Learning curriculum—better known as OWL (2005 version)—for literacy, and the Building Blocks curriculum in math (Georgia Department of Early Care and Learning, 2015). In addition to the academic curricula, the preschool classrooms also seek to promote social-emotional development, planning skills, community building, and executive functioning (including cognitive inhibitory control and attention shifting) (Georgia Department of Early Care and Learning, 2015).

The program maintains quality standards for staff—preschool teachers must have a Bachelor's degree and must obtain a masters degree within 5 years of being hired. They receive curriculum-specific training and 2–4 coaching sessions a month on using the curricula from experienced early childhood coaches (Georgia Department of Early Care and Learning, 2015).

## B.4 Tennessee Voluntary Prekindergarten Program (TN-VPK)

The Tennessee Voluntary Prekindergarten Program is a statewide kindergarten program, targeting disadvantaged 4 years old children one year before kindergarten. It began as a pilot program in 1998, and became statewide in 2005. The program distributes competitive grants to local school systems. In order to get funded, centers apply for one or more classrooms to implement the program. These centers include local nonprofit and for-profit childcare providers or Head Start programs, as long as they are highly rated in the licensing system administered by the Tennessee Department of Human Services (Lipsey et al., 2013). Participating centers have to meet the following requirements: (i) adult-student ratio of no less than 1:10; (ii) maximum class size of 20; (iii) approved age curriculum —which is not specified more concretely. Currently, 934 classrooms are funded by the state through this program. It serves 18,000 children across 95 counties. TN-VPK is a full-day program giving priority to children eligible for federal free or reduced-price lunch—i.e., children with family income lower than 185% of the federal poverty line. Children with disabilities or low English-speaking abilities are also eligible (Lipsey et al., 2015).

During 2009-2011, there was an attempt to perform a randomized controlled trial to evaluate the program. Across the 2009-2010 and 2010-2011 periods, 80 schools in 29 school districts in Tennessee applied to participate in the randomized controlled study, which provided a final experimental sample of 3,025 children. However, there was a major flaw in the design. The study requested parents' consent for using the information of their children in the evaluation *after* randomization, resulting in an imbalance between the treatment and the control groups. In the first cohort, 46% (32%) of the participants of treatment (control) group consented to releasing the information. In the second cohort, the rates were 71% for treatment and 74% for control. In both cohorts, there was imbalance in observed categories between the treatment and control groups (Lipsey et al., 2013). Thus, the randomization protocol did not satisfy minimal standards to assess treatment effects using straightforward, usual methodologies. Furthermore, 27% of children in the control group attended either Head

Start or other center-based childcare programs (Lipsey et al., 2013), complicating further the interpretation of standard estimates as we discussed in the main text of our paper.

## B.5 Early Childhood Education Reform in Norway

The evidence from Norway does not come from a uniform early childhood education program. Instead, it comes from a reform with nationwide effect. The reform (Kindergarten Act) took place in 1975 and provided a staged expansion across 400 municipalities. The objective of the reform was to increase childcare slots, reaching 100,000 in 1981. This quadrupled the availability in 1975. The reform subsidized operation costs and investment. It set a maximum price for childcare center services. It was universal and all families with children age 3 to 6 were eligible for the services of the centers. The slots were allocated on a first come, first served basis, i.e., according to the time a family spent on the waiting list (Havnes and Mogstad, 2011).

The subsidies were distributed by the local (municipality-level) governments to the centers based on (i) number of children served, (ii) age of children, and (iii) amount of time children spent in the center. Municipalities with the lowest childcare enrollment rates were the ones with the largest funding endowments to distribute. Centers were run by private firms, charities, and public organizations. They were subject to uniform standards with respect to: (i) educational content, (ii) group size, (iii) staff skill composition, and (iv) physical environment. Centers were evaluated with respect to these standards and assigned funding afterwards (Havnes and Mogstad, 2011).

Centers provided high-quality educational experience. Prevailing pedagogy emphasized learning through playing, thus promoting development of social, language, and motor skills. Centers provided childcare during normal working hours. Day-to-day operations were overseen by a head preschool teacher with college degree and supervised experience in early childhood education. This head teacher was responsible for satisfactory planning, observation, collaboration and evaluation of staff. This included staff guidance and collaboration

with parents and local authorities, which might be needed to provide health, child welfare, or psychological services. For small-group interactions and day-to-day parent interaction, the teacher-child ratio was capped at 16:1, and each preschool teacher had one or two adult assistants. There were no educational requirements for assistants (Havnes and Mogstad, 2011).

## B.6 A Universal Childcare Subsidy in Quebec, Canada

The evidence from Quebec is not about a uniform program. It was a reform with province-wide effect. The reform (The Quebec Family Policy) took place in 1997. It immediately extended full-time kindergarten to all 5 years old children, and subsidized the price of child-care to have a cap of 5 Canadian dollars per day (in nominal dollars for every year between 1997–2005) for 4 years old children immediately, for 3 years old in 1998, for 2 years old in 1999, and for 2 years old and below in 2000 Baker et al. (2008).

The subsidy implemented by the reform had two components. Existing nonprofit child-care centers were transformed into centers for young children (*centres de la petite infance*) and primarily served 2 year old or older children. Younger children were placed in certified centers, which emerged from home-based care providers. The policy initially generated excess demand, which led to the creation of new centers. The number of childcare slots doubled between 1997 and 2005. The reform contained components aiming to increase the quality of childcare: (i) two-thirds of staff had to have a college diploma or university degree in early childhood education; (ii) government began to provide financial support for childcare providers who were enrolled in college-level education related to early childhood; (iii) staff of certified centers which transformed from home-based care services received 24 to 45 hours of training and had to meet annual professional development requirements. The increase in demand generated by the reform increased maximum size of facilities from 60 to 80 children per year. However, the ratio of staff to children remained constant, between 1:8 and 1:10. Precise details on the curricula implementation are not available Baker et al. (2008).

# C Additional Programs

There are many early childhood education programs that are not reviewed in this chapter. This appendix briefly describes a few important examples and provides a rationale for their exclusion. In addition, there are many programs that affect the early environment through channels other than the child's education. Evaluation of these programs is outside the scope of this chapter.

## C.1 Early Head Start (EHS)

EHS is particularly important because, when combined with Head Start, it provides a program that has substantive similarities to ABC. Early Head Start (EHS) was created in 1994 during the Clinton administration as a part of Head Start and expanded Head Start program benefits to pregnant mothers and low-income families with children under age 3. EHS program services may include development services, child care, parenting education, case management, health care (including referral), and family support. Following in the Head Start tradition, EHS programs also partner with other community service providers to extend their reach and impact. These partnerships are expected to meet the same quality standards as EHS, including child to adult ratios and staff educational qualifications (see Vogel et al., 2006). In fact, EHS programs can be thought of as Head Start programs that serve families with children under age 3. It has been subject to the same policy changes as Head Start and receives its funding as a part of the overall Head Start budget (about 10%). EHS provides grants to over 700 programs serving over 60,000 children (Love et al., 2002).

The Early Head Start Research and Evaluation Project was carried out from 1996 to 2010 and followed children from recruitment until the age of 36 months. The study used EHS eligibility criteria: expectant mothers and children up to age 3 were eligible if they met federal poverty income guidelines. However, some exceptions were made. Data come from the Head Start Family Information System, which contains demographic information

on families, primary caregivers, and focus children before randomization. EHS used the same random-assignment procedure for each site making comparisons between programs more achievable (Vogel et al., 2011).

The initial sample included 3,001 families (Love et al., 2002). 1,513 were randomly assigned to the treatment group and 1,488 were assigned to the control group. Treatment families were allowed to choose from three options: center-based, home-based, or mixed approach.

1. Center-based: all services were provided through center-based child care and parent education. Each family was also required to receive a minimum of 2 home visits per year.

2. Home-Based: all services were provided through weekly home visits, which typically lasted an hour and focused on child development activities. At least two group socializations per month were required of each family.

3. Mixed Approach: services were provided through a combination of center-based and home-based strategies (Love et al., 2002).

EHS used the same random assignment procedure for each site, and used standard statistical tests to assess the similarity of the two research groups, including univariate t-tests to compare variable means for binary and continuous variables, and chi-squared tests to compare distributions of categorical variables. A more formal multivariate analysis was conducted to test the hypothesis that variable means and distributions were jointly similar (Vogel et al., 2011).

Response rates were similar for program and control group members for all data sources (Love et al., 2002). At age 10, relative to other subgroups in the sample, the high-risk group experienced high attrition over time. 71% of the high-risk group did not respond at 1 or more data-collection waves, and 13% did not respond at any wave of the follow-up (Vogel et al., 2011).

We do not review results from Early Head Start due to a scarcity of evaluations, extremely short-term follow-up, and a short duration of treatment. An evaluation of Early Head Start is Love et al. (2005). They use a standard instrumental variables framework to assess the effects of program participation on a variety of outcomes at age 3. Early Head Start had three types of implementation (i) center-based programs; (ii) home-based programs; (iii) mixed approach programs. When pooling the sample, they find important gains on mental development, cognition, and some measures of child behavior. Unfortunately, the results are not as clear when the samples are broken into type of implementation. The literature evaluating Early Head Start needs further exploration. Furthermore, it fails to provide estimates of the effects of the program in the long-term, because data are not available. Given its similarities with Head Start, evaluations also need to discuss whether control contamination is an issue.

## C.2 The Milwaukee Project

Another program that we do not study is the Milwaukee Project. The Milwaukee Project was a randomized longitudinal study of 20 treatment and 20 control children who were followed up to age 14. The population consisted of black children from a low income Milwaukee area, with mothers having IQ scores under 75. The program had two main components: a maternal rehabilitation program and an infant program, from 3 months of age until the child entered school. Children attended the program five days a week through the year. The adult-child ratio was 1:1 during infancy and was gradually decreased to 1:3 by age two. Some additional support was provided to children after entering school.

The authors claim huge impacts in IQ (30 points of initial impact) (Garber and Begab, 1988) and additional impacts on mother-child verbal interactions, and parenting attitudes. However, the extremely positive results, the charges of financial malfeasance directed against one of the researchers, the small number of publications in peer-reviewed journals, and the refusal to share their data have created much skepticism about the results of the study. The

credibility of the study remains unknown.

## C.3 Chicago Parent-Child Centers (CPC)

CPC began in 1967 and targets disadvantaged children. It is the second oldest federally funded preschool program after Head Start. Assignment is not random but follow-ups are available both for participants and non-participants up to age 26. Its evaluation compares children attending preschool and kindergarten to individuals who attended kindergarten but not preschool (Reynolds et al., 2011). It is based on only one (of many possible cohorts) of children in the program. The evaluators are unwilling to release the full data set so that replication of their claimed results is not possible.

## C.4 Nurse Family Partnership

This chapter compares and aligns evidence from early childhood education programs. Programs that affect the early environment through other channels are excluded, such as child-care subsidies and health, nutrition, or parenting interventions. One important example is the Nurse Family Partnership (NFP), which is a home-visiting program aimed at disadvantaged first-time mothers to improve maternal and infant health and well-being. While it is not a center-based childhood education program, it is means-tested and large-scale. By focusing on parenting skills, participating parents may even learn to perform activities similar to those children would learn in a center-based program. This appendix includes a description of NFP and a very brief discussion of its results as an example to demonstrate the need for future research. Except for an age limited sample from Memphis, Tennessee, evaluators of the program are unwilling to release their data to the public, so it is not possible to confirm their claims.

### C.4.1 Program Description

NFP is an ongoing large-scale program, serving over 20,000 families in many states across the US (Olds, 2006). The goal of the programs is to (i) improve pre- and perinatal mother and infant health by educating pregnant women; (ii) improve parent-child interactions by teaching parenting skills; and (iii) promote family economic self-sufficiency by encouraging planning behaviors and connecting parents with resources. The program began with a randomized control trial in 1977 in Elmira, New York, and had two later randomized control trials—one in 1988 in Memphis, Tennessee, and one in 1994 in Denver, Colorado. Each of the trials differed in key characteristics. Each program recruited at-risk, first-time mothers from clinics and provided nurse home-visiting services but differed in the populations that participated. In Elmira, the sample was made up of low-income whites. In Memphis, 92% of the women who accepted were African-American (Kitzman et al., 1997). In Denver, 15% of accepted women were African-American, and 35% were Hispanic (Olds et al., 2014). Furthermore, the Denver trial used paraprofessionals (who were trained in delivering the treatment but did not have college degrees) in addition to nurses to deliver the treatment. These trials also differed slightly in their eligibility criteria and treatment.

#### Elmira

The women were put in 4 different treatment conditions (Olds et al., 1986).

1. Control Group ($n = 90$). No services were provided through the research project. At age 1 and 2, a specialist hired by the research project screened infants for sensory and developmental problems.

2. Families in the second group ($n = 94$) were provided free transportation for regular prenatal and child care at local clinics and physicians' offices through a contract with a local taxicab company. Sensory and developmental screening were provided when

the infants were 1 and 2 years of age.

3. Families in the third group ($n = 100$) were provided a nurse home visitor during pregnancy in addition to the screening and transportation services. The nurses visited families about once every 2 weeks and made an average of 9 visits during pregnancy, each of which lasted one hour and 15 minutes.

4. Families assigned to the fourth group ($n = 116$) received the same services during pregnancy as those in treatment 3, but in addition the nurses continued to make visits until the babies were 2 years of age. For the first month after delivery, the nurses visited once a week; thereafter, they visited on a schedule of diminishing frequency until the children were 1.5 to 2 years old, when visits were made every 6 weeks.

During visits, the nurses carried out three basic activities: parent education, enhancement of the women's informal support systems, and linkage of the parents with community services. Eligibility for the program was determined by (i) maternal age (<19 years); (ii) single-parent status; (iii) low socioeconomic status (Olds et al., 1986). However, the study design allowed any women who asked to participate and who was bearing a first child to be enrolled. The randomization was stratified on the basis of mothers's marital status, race, and geographic region in which they lived.

**Memphis**

The women were put into 4 different treatment conditions (Kitzman et al., 1997).

1. Women in treatment 1 ($n = 166$) were provided free round-trip taxicab transportation for scheduled prenatal care appointments; they did not receive any postpartum services or assessments.

2. Women in treatment 2 ($n = 515$) were provided free transportation for scheduled

prenatal care in addition to developmental screening and referral services for their children at 6, 12, and 24 months of age.

3. Those in treatment 3 ($n = 230$) were provided free transportation and screening offered in treatment 2 plus intensive nurse home visitation services during pregnancy, one postpartum visit in the hospital before discharge, and one postpartum visit in the home.

4. Women in treatment 4 ($n = 228$) were provided the same services as those in treatment 3; in addition, they continued to be visited by nurses through the child's second birthday.

Eligibility was determined by the following characteristics:

1. Less than 29 weeks pregnant

2. No previous live births

3. No specific chronic illness thought to affect to fetal growth, retardation, or pre-term delivery

And, at least 2 of the following socio-demographic risk conditions:

1. Unemployed

2. Fewer than 12 years of education

3. Unmarried

The randomization was conducted within strata from a model with 5 classification factors: maternal race (African-American or non-African-American), maternal age ($< 17, 17-18$, and $\geq 19$ years), gestational age at enrollment ($< 20$ or $\geq 20$ weeks), employment status of head of household (employed or unemployed), and geographic region of residence (4 regions).

### Denver

The Denver trial sought to explore the effectiveness of staff quality in treatment delivery and included a group of paraprofessionals (who were trained but not college-educated) and a group of nurses to deliver the treatment. Women in the control group ($n = 255$) were provided developmental screening and referral services for their children at 6, 12, 15, 21, and 24 months. Women assigned to the paraprofessional group ($n = 245$) were provided the screening and referral services in addition to paraprofessional home visitation during pregnancy and infancy (the first 2 years of the child's life). Women in the nurse group ($n = 235$) were provided screening and referral and nurse home visitation during pregnancy and infancy (Olds et al., 2002).

The randomization was conducted within strata from a model with 3 classification factors: maternal race (Hispanic, white non-Hispanic, African-American, American-Indian, or Asian), maternal gestational age at enrollment (<32 or 32+ weeks), and geographic region of residence (4 regions). Women assigned to 1 of the 2 home-visitation groups subsequently were assigned at random to home visitors responsible for their geographic region.

### C.4.2 Youth Outcomes

Each trial has a different amount of follow-up; Elmira had follow-up until age 19, Memphis until age 21, and Denver until age 9. None of the programs had an effect on IQ. This is especially interesting, since evaluations of center-based programs consistently show impacts on IQ when such measurements are available. This could potentially be a benefit of center-based programs over home-visiting programs, which supports the explanation that children gain an immediate boost in IQ from initial enrollment in schooling. Exploring this relationship thoroughly and comparing center-based and home-visiting programs is outside the scope of this chapter.

### Elmira

Eckenrode et al. (2010) use a factorial structure with two covariates—child's race (white vs nonwhite) and mother's baseline education. The classification factors were treatment group (1 and 2 vs 3 vs 4), risk (unmarried and from low-socioeconomic status families at registration vs married or higher socioeconomic status), and the youth's sex. All interactions among the factors were included in the model. Quantitative variables were examined using a general linear model.

At age 19, the treatment group was less likely to have ever been arrested or convicted. As with PPP and ABC, the effect was stronger for males. Treatment did not affect rates of high school graduation.

### Memphis

Heckman et al. (2014) use a factor model for proxied skills (measured by a range of psychological instruments) to forecast how age 2 skills mediate age 6 outcomes and how age 6 skills in turn mediate age 12 outcomes. They account for measurement error and multiple hypotheses testing. They find that in the first two years of life, the program improved the quality of the home environment as measured by provision of appropriate play material, variety in daily stimulation, and maternal involvement with the child. In the first 12 years of life, treated families had significantly less reliance on food stamps and Medicaid. By age 2, females had positive effects on emotional stability and mental health measures, and males and females show positive effects on mastery.

The program had differing effects for males and females at age 6 and 12. Males improved in math and reading achievement test scores and grades for grades 1–5. Females show improved results in the Kaufman Assessment Battery at age 6 (including nonverbal skills and sequential and simultaneous processing). However, females show much greater improvements in non-cognitive skills including anxiety, hyperactivity, and aggression and conduct at age 6. There were no significant non-cognitive outcomes at age 12.

Mediation analyses show that 35% and 22%, respectively, of the cognitive effects on males and females at age 6 are due to improvements in parenting. Additionally, birth-weight explains 14% of the treatment effect on males, and maternal anxiety explains 25% of the effect on females. Maternal anxiety (along with home environment) also mediates female non-cognitive gains. Predictably, cognition at age 6 explains much of the improvement in achievement at age 12 for males. At age 12, 66% of the effect on GPA, 46% of the effect on reading comprehension, and 51% of the effect on math scores are mediated by age 6 cognitive skills. Age 6 cognition also mediated class absenteeism, internalizing behavior, and anxiety for males.

**Denver**

Olds et al. (2014) use a single classification factor for treatment (3 levels) with 6 baseline covariates (maternal psychological resource index, smoking status, whether mothers registered in the study after 28 weeks of gestation, housing density, maternal conflict with her mother or mother figure, and neighborhood disadvantage) to adjust for treatment non-equivalence among participants assessed at either the 6- or 9-year follow-ups, plus 2 additional covariates (child age at assessment and sex). Results differed substantially for the group treated by paraprofessionals and the group treated by nurses. There was also heterogeneity within those groups by various measures of disadvantage.

The children in the paraprofessional treatment group in which mothers had low psychological resources showed improvements in visual attention and task switching at age 9. There were no other statistically significant effects on the paraprofessional treatment group in cognition, behavior or grade retention.

The children in the nurse treatment group had positive effects on non-cognitive skills. At age 6, they were less likely to have emotional or behavioral problems. At age 9, they were less likely to engage in internalizing behavior and attention problems. Children born to low-resource mothers also show improvements is receptive language and sustained attention.

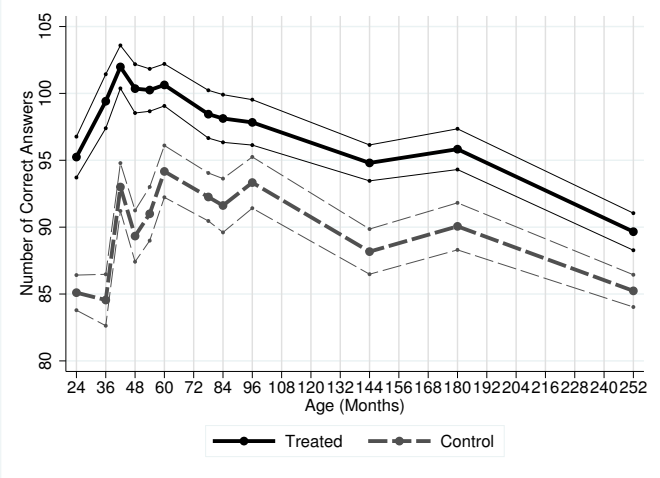
For the nurse-visited children, there were no significant effects on externalizing problems, intellectual functioning, and academic achievement (Olds et al., 2014). At age 2 and 4, these children had better receptive language scores than the paraprofessional treatment group, though the difference became insignificant at age 6. The improvements in sustained attention for the children in the nurse treatment group persisted over the age 4, 6 and 9 follow-ups. There were no differences between the nurse and paraprofessional treatment groups in visual attention/task switching, working memory, intellectual functioning, or academic achievement.

# D IQ and Achievement Dynamics in Demonstration Programs

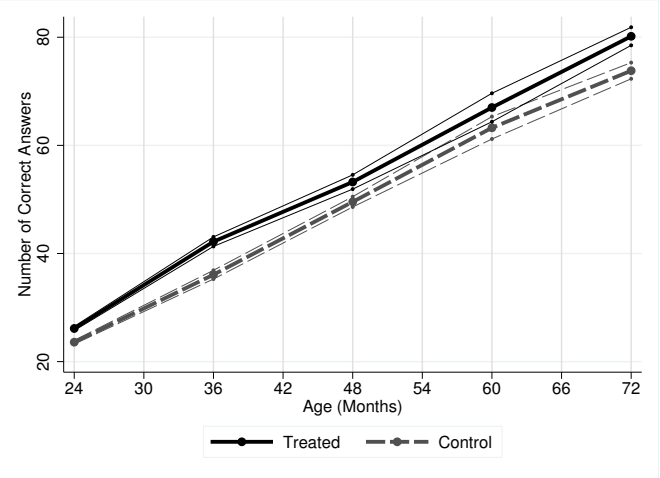
This appendix provides evidence on the fadeout phenomenon for measured IQ for programs besides PPP. As discussed in the text, IQ data is presented in two ways: using national age-by-age standardization or using raw scores. We present both types of scores when they are available. The patterns of fadeout are similar to the ones presented for PPP in the main paper: large initial impacts on IQ diminish. In the case of ABC, the impacts do not fade out completely, while in the cases of ETP and IHDP, they do. For ABC and ETP, the trends of the raw scores are similar to those presented for PPP in the main paper. Both groups have trends that are strongly increasing over time, implying that the fadeout of impacts is better interpreted as a catch-up of the control group rather than as a depreciation of the skills of the treated group (see Hojman, 2015).

Figure D.1: IQ Dynamics in ABC

(a) Standardized Scores



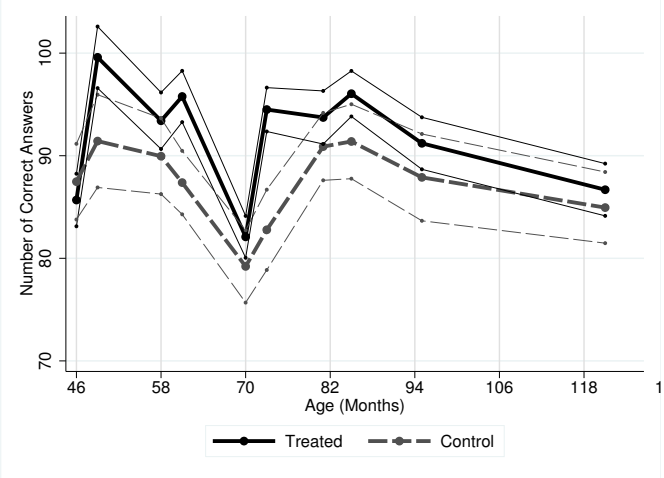

(b) Raw Scores



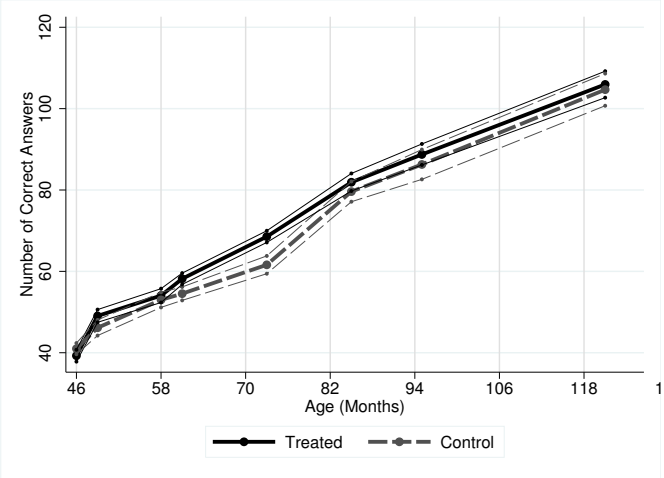

Source: Reproduced from Hojman (2015). Note: The solid line represents the trajectory of the treated group, and the dotted line represents the trajectory of the control group. Thin lines surrounding trajectories are asymptotic standard errors.

Figure D.2: IQ Dynamics in ETP

(a) Standardized Scores

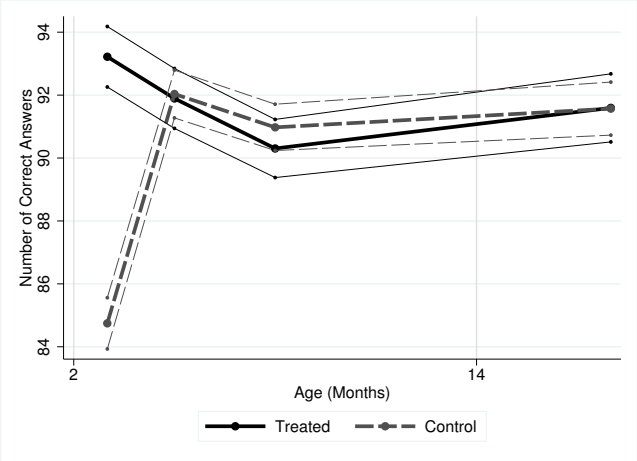



(b) Raw Scores




Source: Reproduced from Hojman (2015). Note: The solid line represents the trajectory of the treated group, and the dotted line represents the trajectory of the control group. Thin lines surrounding trajectories are asymptotic standard errors.

Figure D.3: IQ Dynamics in IHDP




Source: Reproduced from Hojman (2015). Note: The solid line represents the trajectory of the treated group, and the dotted line represents the trajectory of the control group. Thin lines surrounding trajectories are asymptotic standard errors.

# E The Formation of Skills over the Life-cycle

We draw on Heckman and Mosso (2014) and represent the vector of skills at age $t$ by $\boldsymbol{\theta}_t$ over lifetime $T$. We describe the process of skill formation as depending on three main inputs:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{f_t}\left(\underbrace{\boldsymbol{\theta}_t}_{\text{own skills}}, \underbrace{\boldsymbol{I}_t}_{\text{investment}}, \underbrace{\boldsymbol{\theta}_t^P}_{\text{parental skills}}\right). \tag{1}$$

For simplicity, we assume there is one parent. $\boldsymbol{\theta}_t$, $\boldsymbol{\theta}_t^P$, and $\boldsymbol{I}_t$ are vector-valued. Let $j$ and $j'$ denote two elements of these vectors. The production function of skills, $\boldsymbol{f_t}$, exhibits intertemporal productivity in skill $j$ if

$$\frac{\partial \theta_{t+1}^{j'}}{\partial \theta_t^j} > 0. \tag{2}$$

For the case $j = j'$, skill $j$ is self-productive when it does not fully depreciate from $t$ to $t+1$ but instead builds on itself across time. For example, a child learning to speak may use vocabulary learned at age 2 to learn more words at age 3. For the case $j \neq j'$, $\boldsymbol{\theta}_t^{j'}$ exhibits cross-productivity with $\theta_{t+1}^j$ if one skill facilitates creation of another skill. For example, a child's level of extroversion may contribute to his/her future language development. Cunha et al. (2010) find that cognitive skills in $t+1$ build on cognitive *and* non-cognitive skills in $t$, supporting cross-productivity. Interestingly, they find that non-cognitive skills in $t+1$ do not build on cognitive skills in $t$.

In addition to affecting each other dynamically, skills may also affect each other contemporaneously. Static skill-to-skill complementarity occurs if:

$$\frac{\partial^2 \theta_{t+1}^j}{\partial \theta_t^j \partial \theta_t^{j'}} > 0. \tag{3}$$

for $j \neq j'$. Similarly, investment and skill levels at age $t$ may complement each other to build skills at $t+1$ and exhibit static skill-investment complementarity if:

$$\frac{\partial^2 \theta_{t+k}^{j'}}{\partial \theta_t^j \partial I_t^{j''}},\ k \geq 1 \text{ for all } j, j', j''. \tag{4}$$

Dynamic complementarity arises when investment at age $t$ boosts the stock of skills in future periods and enhances static complementarity in those periods. Recent studies of the economic efficiency of early investment show that in early periods of life there is either static substitution between skills and investment or relatively low complementarity. That is, there exists a life-cycle period $[0, \bar{t}]$ such that

$$\frac{\partial^2 \theta_{t+k}^{j'}}{\partial \theta_t^j \partial I_t^{j''}} \leq \varepsilon,\ k \leq \bar{t} - t \tag{5}$$

for $t \in [0, \bar{t}]$ and a small enough $\varepsilon > 0$ (which may be negative).

This claim has both a theoretical and empirical basis. Complementarity increases with age. Heckman and Mosso (2014) show that it can be efficient to invest more in a disadvantaged child during an initial period under certain curvature conditions of $\boldsymbol{f_t}(\cdot)$. This increases the return to investment in a disadvantaged child and brings it closer to the return to future investment in the relatively advantaged child. While it can be economically efficient to invest in disadvantaged children due to increasing complementarity with age, it can be economically inefficient to invest in disadvantaged adolescents with a low skill base for whom returns are low. Cunha and Heckman (2008) and Cunha et al. (2010) find that static complementarity between skills and investment increases with age. This leads to two fundamental aspects of skill formation: (i) investments in relatively more skilled individuals become more productive as they age; and (ii) complementarity between skills and investments increases over the life cycle. Together, these two features imply that later life remediation investment is less efficient than early life prevention and investment because dynamic complementarity of investment increases over time. That is, complementarity increases with age.

# Appendix References


Administration for Children and Families, Office of Head Start (2009). Head Start Program performance standards and other regulations. Technical report, U.S. Department of Health and Human Services, Washington, DC.

Administration for Children and Families, Office of Head Start (2014). Head Start Program facts fiscal year 2014. Technical report, U.S. Department of Health and Human Services, Washington, DC.

Baker, M., J. Gruber, and K. Milligan (2008, August). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy 116*(4), 709–745.

Barnett, W. S., M. E. Carolan, J. H. Squires, K. Clarke Brown, and M. Horowitz (2015). The state of preschool 2014: State preschool yearbook. Technical report, National Institute for Early Education Research, New Brunswick, NJ.

Bassok, D., M. Fitzpatrick, and S. Loeb (2014). Does state preschool crowd-out private provision? The impact of universal preschool on the childcare sector in Oklahoma and Georgia. *Journal of Urban Economics 83*, 18–33.

Brooks-Gunn, J., P. K. Klebanov, and F.-r. Liaw (1995). The learning, physical, and emotional environment of the home in the context of poverty: The Infant Health and Development Program. *Children and Youth Services Review 17*(1), 251–276.

Brooks-Gunn, J., M. C. McCormick, P. K. Klebanov, and C. McCarton (1998). Health care use of 3-year-old low birth weight premature children: Effects of family and neighborhood poverty. *The Journal of Pediatrics 132*(6), 971–975.

Burchinal, M. R., F. A. Campbell, D. M. Bryant, B. H. Wasik, and C. T. Ramey (1997, October). Early intervention and mediating processes in cognitive performance of children of low-income African American families. *Child Development 68*(5), 935–954.

Campbell, F. A., G. Conti, J. J. Heckman, S. H. Moon, and R. Pinto (2013, March). The effects of early intervention on human development and social outcomes: Provisional evidence from ABC and CARE. University of Chicago, Department of Economics.

Campbell, F. A., G. Conti, J. J. Heckman, S. H. Moon, R. Pinto, and E. P. Pungello (2014). Early childhood investments substantially boost adult health. *Science 343*(6178), 1478–1485.

Campbell, F. A. and C. T. Ramey (1989, April). Preschool vs. school-age intervention for disadvantaged children: Where should we put our efforts? Report, North Carolina University, Frank Porter Graham Center, Chapel Hill. Paper presented at the Biennial Meeting of the Society for Research in Child Development (Kansas City, MO, April 27–29, 1989).

Campbell, F. A. and C. T. Ramey (1991). The Carolina Abecedarian Project. Seattle, WA. Biennial Meeting of the Society for Research in Child Development.

Campbell, F. A. and C. T. Ramey (1994, April). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development 65*(2), 684–698.

Campbell, F. A., B. Wasik, E. Pungello, M. Burchinal, O. Barbarin, K. Kainz, J. Sparling, and C. Ramey (2008). Young adult outcomes of the Abecedarian and CARE early childhood educational interventions. *Early Childhood Research Quarterly 23*(4), 452–466.

Cascio, E. U. and D. W. Schanzenbach (2013). The impacts of expanding access to high-quality preschool education. Working Paper 19735, National Bureau of Economic Research.

Clarke, S. H. and F. A. Campbell (1998). Can intervention early prevent crime later? The Abecedarian Project compared with other programs. *Early Childhood Research Quarterly 13*(2), 319–343.

Cunha, F. and J. J. Heckman (2008, Fall). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources 43*(4), 738–782.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010, May). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Eckenrode, J., M. Campa, D. W. Luckey, C. R. Henderson, R. Cole, H. Kitzman, E. Anson, K. Sidora-Arcoleo, and D. L. Olds (2010, January). Long-term effects of prenatal and infancy nurse home visitation on the life course of youths: 19-year follow-up of a randomized trial. *Journal of the American Medical Association 164*(1), 9–15.

Garber, H. L. and M. J. Begab (1988). *The Milwaukee Project : Preventing Mental Retardation in Children at Risk*. Washington, DC: American Association on Mental Retardation.

García, J. L. (2015). Childcare and parental investment: Short and long-term effects. University of Chicago, Department of Economics.

Georgia Department of Early Care and Learning (2015). Bright from the start.

Gray, S. W. and R. A. Klaus (1970). The early training project: A seventh-year report. *Child Development 41*(4), 909–924.

Gray, S. W., B. K. Ramsey, and R. A. Klaus (1982). *From 3 to 20: The Early Training Project.* Baltimore, MD: University Park Press.

Gross, R. T. (1990). Enhancing the outcomes of low-birth-weight, premature infants: A multisite, randomized trial. *Journal of the American Medical Association 263*(22), 3035–3042.

Gross, R. T., D. Spiker, and C. W. Haynes (1997). *Helping Low Birth Weight, Premature Babies: The Infant Health and Development Program.* Stanford, CA: Stanford University Press.

Haskins, R. (1985, June). Public school aggression among children with varying day-care experience. *Child Development 56*(3), 689–703.

Havnes, T. and M. Mogstad (2011). Money for nothing? Universal child care and maternal employment. *Journal of Public Economics 95*(11), 1455–1465.

Heckman, J. J., M. Holland, T. Oey, D. L. Olds, R. Pinto, and M. Rosales (2014). A reanalysis of the Nurse Family Partnership Program: The Memphis randomized control trial. University of Chicago, Department of Economics.

Heckman, J. J., S. H. Moon, and R. Pinto (2014). The effects of early intervention on abilities and social outcomes: Evidence from the Carolina Abecedarian study. University of Chicago, Department of Economics.

Heckman, J. J. and S. Mosso (2014). The economics of human development and social mobility. *Annual Review of Economics 6*(1), 689–733.

Hojman, A. (2015). Evidence on the fade-out of IQ gains from early childhood interventions: A skill formation perspective. University of Chicago, Center for the Economics of Human Development.

Kitzman, H., D. L. Olds, C. R. Henderson, C. Hanks, R. Cole, R. Tatelbaum, K. M. McConnochie, K. Sidora, D. W. Luckey, D. Shaver, K. Engelhardt, D. James, and K. Barnard (1997, August). Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing: A randomized controlled trial. *Journal of the American Medical Association 278*(8), 644–652.

Klaus, R. A. and S. W. Gray (1968). *The Early Training Project for Disadvantaged Children: A Report After Five Years*. Chicago: University of Chicago Press.

Lipsey, M. W., D. C. Farran, and K. G. Hofer (2015). A randomized control trial of the effects of a statewide voluntary prekindergarten program on children's skills and behaviors through third grade. Research report, Peabody Research Institute, Vanderbilt University, Nashville, TN.

Lipsey, M. W., K. G. Hofer, N. Dong, D. C. Farran, and C. Bilbrey (2013). Evaluation of the Tennessee Voluntary Prekindergarten Program: Kindergarten and first grade follow-up results from the randomized control design. Research report, Peabody Research Institute, Vanderbilt University, Nashville, TN.

Love, J. M., E. Eliason Kisker, C. Ross, H. Raikes, J. Constantine, K. Boller, R. Chazen-Cohen, J. Brooks-Gunn, L. B. Tarullo, C. Brady-Smith, A. Sidle Fuligni, P. Z. Schochet, D. Paulsell, and C. Vogel (2005). The effectiveness of early Head Start for 3-year-old children and their parents: Lessons for policy and programs. *Developmental Psychology 41*(6), 885–901.

Love, J. M., E. Eliason-Kisker, C. M. Ross, P. Z. Schochet, J. Brooks-Gunn, D. Paulsell, K. Boller, J. Constantine, C. Vogel, A. Sidle Fuligni, and C. Brady-Smith (2002). Making

a difference in the lives of infants and toddlers and their families: The impacts of early Head Start. Technical Report PR02-30a, Mathematica Policy Research Report.

Martin, A., P. Brooks-Gunn, Jeanne amd Klebanov, S. L. Buka, and M. C. McCormick (2008). Long-term maternal effects of early childhood intervention: Findings from the Infant Health and Development Program (IHDP). *Journal of Applied Developmental Psychology 29*(2), 101–117.

McCormick, M. C., J. Brooks-Gunn, S. L. Buka, J. Goldman, J. Yu, M. Salganik, D. T. Scott, F. C. Bennett, L. L. Kay, J. C. Bernbaum, C. R. Bauer, C. Martin, E. R. Woods, A. Martin, and P. H. Casey (2006, March). Early intervention in low birth weight premature infants: Results at 18 years of age for the Infant Health and Development Program. *Pediatrics 117*(3), 771–780.

Olds, D. L. (2006). The Nurse-Family Partnership. In N. F. Watt, C. Ayoub, R. H. Bradley, J. E. Puma, and W. A. LeBoeuf (Eds.), *The Crisis in Youth Mental Health: Critical Issues and Effective Programs*, Volume 4: Early Intervention Programs and Policies, pp. 147–180. Westport, CT: Praeger Perspectives.

Olds, D. L., C. R. Henderson, R. Chamberlin, and R. Tatelbaum (1986). Preventing child abuse and neglect: A randomized trial of nurse home visitation. *Pediatrics 78*(1), 65–78.

Olds, D. L., J. R. Holmberg, N. Donelan-McCall, D. W. Luckey, M. D. Knudtson, and J. Robinson (2014, February). Effects of home visits by paraprofessionals and by nurses on children: Age-six and nine follow-up of a randomized trial. *Journal of the American Medical Association Pediatrics 168*(2), 114–121.

Olds, D. L., J. Robinson, R. O'Brien, D. W. Luckey, L. M. Pettitt, C. R. Henderson, R. K. Ng, K. L. Sheff, J. Korfmacher, S. Hiatt, and A. Talmi (2002). Home visiting by paraprofessionals and by nurses: A randomized, controlled trial. *Pediatrics 110*(3), 486–496.

Puma, M., S. Bell, R. Cook, and C. Heid (2012). Head Start Impact Study: Final report. Technical report, Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.

Puma, M., S. Bell, R. Cook, C. Heid, P. Broene, F. Jenkins, A. Mashburn, and J. Downer (2012). Third grade follow-up to the Head Start Impact Study: Final report. OPRE Report 2012–45, Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.

Ramey, C., K. O. Yeates, and E. J. Short (1984, October). The plasticity of intellectual development: Insights from preventive intervention. *Child Development 55*(5), 1913–1925.

Ramey, C. T. and F. A. Campbell (1979, February). Compensatory education for disadvantaged children. *The School Review 87*(2), 171–189.

Ramey, C. T. and F. A. Campbell (1991). Poverty, early childhood education and academic competence: The Abecedarian experiment. In A. C. Houston (Ed.), *Children in Poverty:*

*Child Development and Public Policy*, Chapter 8, pp. 190–221. New York: Cambridge University Press.

Ramey, C. T., F. A. Campbell, M. Burchinal, M. L. Skinner, D. M. Gardner, and S. L. Ramey (2000). Persistent effects of early childhood education on high-risk children and their mothers. *Applied Developmental Science 4*(1), 2–14.

Ramey, C. T., A. M. Collier, J. J. Sparling, F. A. Loda, F. A. Campbell, D. A. Ingram, and N. W. Finkelstein (1976). The Carolina Abecedarian Project: A longitudinal and multidisciplinary approach to the prevention of developmental retardation. In T. Tjossem (Ed.), *Intervention Strategies for High-Risk Infants and Young Children*, pp. 629–655. Baltimore, MD: University Park Press.

Ramey, C. T. and R. Haskins (1981). The modification of intelligence through early experience. *Intelligence 5*(1), 5–19.

Ramey, C. T., M. C. Holmberg, J. H. Sparling, and A. M. Collier (1977). An introduction to the Carolina Abecedarian Project. In B. M. Caldwell and D. J. Stedman (Eds.), *Infant Education: A Guide for Helping Handicapped Children in the First Three Years*, Chapter 7, pp. 101–121. New York: Walker and Company.

Ramey, C. T., G. D. McGinness, L. Cross, A. M. Collier, and S. Barrie-Blackley (1982). The Abecedarian approach to social competence: Cognitive and linguistic intervention for disadvantaged preschoolers. In K. M. Borman (Ed.), *The Social Life of Children in a Changing Society*, Chapter 7, pp. 145–174. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ramey, C. T. and B. J. Smith (1977, January). Assessing the intellectual consequences of early intervention with high-risk infants. *American Journal of Mental Deficiency 81*(4), 318–324.

Ramey, C. T., J. J. Sparling, and R. Wasik (1981). Creating social environments to facilitate language development: An early education approach. In R. L. Schiefelbusch and D. D. Bricker (Eds.), *Early Language: Acquisition and Intervention*, pp. 447–476. Baltimore, MD: University Park Press.

Reynolds, A. J., J. A. Temple, B. A. B. White, S.-R. Ou, and D. L. Robertson (2011, January–February). Age 26 cost-benefit analysis of the Child-Parent Center early education program. *Child Development 82*(1), 379–404.

Sanyal, M. A., F. W. Henderson, E. C. Stempel, A. M. Collier, and F. W. Denny (1980). Effect of upper respiratory tract infectionon Eustachian tube ventilatory function in the preschool child. *Journal of Pediatrics 97*(1), 11–15.

Schweinhart, L. J. (2003). Benefits, costs, and explanation of the High/Scope Perry Preschool Program. Tampa, FL. Meeting of the Society for Research in Child Development.

Schweinhart, L. J. (2006). *The High/Scope Approach: Evidence that Participatory Learning in Early Childhood Contributes to Human Development*, Volume 4 of *The Crisis in*

*Youth Mental Health: Critical Issues and Effective Programs*. Westport, CT: Praeger Publishers/Greenwood Publishing Group.

Schweinhart, L. J., J. Montie, Z. Xiang, W. S. Barnett, C. R. Belfield, and M. Nores (2005). *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40*, Volume 14 of *Monographs of the HighScope Educational Research Foundation*. Ypsilanti, MI: High/Scope Press.

Sparling, J. and I. Lewis (1979). *LearningGames for the First Three Years : A Guide to Parent/Child Play*. New York: Berkley Books.

Sparling, J. and I. Lewis (1984). *LearningGames for Threes and Fours*. New York: Walker and Company.

United States Census Bureau (2014). American Community Survey. Technical report, United States Census Bureau.

Vogel, C. A., N. Aikens, A. Burwick, L. Hawkinson, A. Richardson, L. Mendenko, and R. Chazan-Cohen (2006, December). Findings from the survey of early Head Start programs: Communities, programs, and families. Final report. Technical Report ED498072, U.S. Department of Health and Human Services.

Vogel, C. A., Y. Xue, E. M. Moiduddin, B. L. Carlson, and E. E. Kisker (2011). Early Head Start children in grade 5: Long-term followup of the early Head Start research and evaluation project study sample. 2011-8, Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.

Wasik, B. H., C. Ramey, D. M. Bryant, and J. J. Sparling (1990, December). A longitudinal study of two early intervention strategies: Project CARE. *Child Development 61*(6), 1682–1696.

Weikart, D. P. (Ed.) (1967). *Preschool Intervention: A Preliminary Report of the Perry Preschool Project*. Ann Arbor, MI: Campus.

Weikart, D. P., J. T. Bond, and J. T. McNeil (1978). *The Ypsilanti Perry Preschool Project: Preschool Years and Longitudinal Results Through Fourth Grade*. Ypsilanti, MI: Monographs of the High/Scope Educational Research Foundation.

Weiland, C. and H. Yoshikawa (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development 84*(6), 2112–2130.