Can Early Childhood Interventions Decrease Inequality of Economic Opportunity?

Katherine Magnuson, University of Wisconsin–Madison

Greg J. Duncan, University of California, Irvine

DRAFT October 2014

Prepared for the Federal Reserve Bank of Boston Conference, Inequality of Economic Opportunity in the United States, October 17-18, 2014, Boston, Massachusetts.

Early childhood has emerged as a "frontier" in economic research related to the production of human capital. It comes to this role as a potentially important input into the human capital production function, and as the only period of childhood and adolescence with relatively little public investment. But as the frontier metaphor suggests, early childhood is a contested area of research. Some scholars interpret the early childhood intervention evidence as showing promising opportunities for remediating inequities in human capital and thus argue for significant expansions in public investments. Other scholars come to more cautious or even negative conclusions, worrying most about the degree of risk and uncertainty in the current evidence base regarding the long-run payoffs to early childhood investments.

In this paper we review the evidence on the potential of early childhood investments, particularly center-based early childhood education (ECE, sometimes also referred to as preschool), to reduce economic inequality. We begin by discussing the developmental processes related to human capital accumulation, and how early environments and investments may affect human capital development in the early years. We describe a taxonomy of skills formed in early childhood and how each might link to later economic outcomes. We describe what is known about socioeconomic status (SES) -based human capital gaps at the time of school entry.

Next we turn to the potential of public investments in early childhood education to improve both school readiness and later economic outcomes of children born to low-income parents. We review evidence on current patterns of enrollment in preschool patterns among U.S. children, the effectiveness of these programs in the short and long-run, and the costs and benefits of expanding programs to increase enrollments. We demonstrate that increasing enrollments for preschoolers in the year before school entry is a worthwhile investment that will have important economic payoffs in terms of increased human capital accumulation and later earnings. The

benefits of even a moderately effective early childhood education program are likely to be substantial enough to offset the costs of program expansion. We conclude with some thoughts about what other types of investments should be further studied.

Models of Child Investments and Development

Both human and animal studies establish the critical importance of the earliest years of life for establishing the brain architecture and other biological systems that will shape future cognitive, social, and emotional development, as well as physical and mental health (Knudsen, Heckman, Cameron, & Shonkoff, 2006; Blair & Raver, 2012). Infants, toddlers, and preschoolers[1] benefit from environments that provide sensitive, responsive caregiving and a variety of language-rich learning opportunities tailored to their capabilities and needs. Research on the malleability (plasticity) of cognitive and language abilities shows these skills to be highly responsive to both positive and negative influences (Fox, Levitt, & Nelson, 2010; Shonkoff, 2010). Environmental enrichment can promote cognitive development, whereas a variety of adverse experiences may shape cognitive development in ways that are ill-suited for later learning (Shonkoff, 2010).

Economic models of human development result from efforts to understand the human capital production function, with an emphasis toward bettering modeling of how investments and child endowments interact to boost a child's stock of human capital. Cunha and Heckman (2007) describe a cumulative model of the production of human capital accumulation that allows for the possibility of differing childhood investment stages as well as roles for the past effects and future development of both cognitive and socioemotional ("noncognitive") skills.

At birth, children have stocks of cognitive and socioemotional potential that reflect a combination of genetic endowment and prenatal environmental influences. The Cunha and

Heckman (2007) model highlights the interactive nature of skill building and investments from families, preschools and schools, and other agents. It posits that human capital accumulation results from "self-productivity" – skills developed in earlier stages bolster the development of skills in later stages – as well as the dynamic complementarity that results from the assumption that skills acquired prior to a given investment increase the productivity of that investment. Taken together, these two principles undergird their hypothesis that "skill begets skill."

One important strength of Cunha and Heckman-type models is that they generate clear and testable hypotheses. Although widely described and endorsed, the hypothesis of dynamic complementarity in early childhood currently rests on a thin empirical base. The most direct evidence comes from work by Aizer and Cunha (2012) who use data from a longitudinal study begun in the 1960s that spans the period surrounding the introduction of Head Start, the largest preschool intervention for low-income children, and finds the largest impacts of the program on children with higher scores on a measure of cognitive development at 8 months (Bayley Mental Development Subscale). However, evidence using more recent data from an experimental evaluation of Head Start does not find that significantly larger gains accrue to students who enter the program with higher levels of initial skills (Purtell & Gershoff, 2013).

Developmental psychologists, like economists, also describe children's development as the result of the dynamic interplay between an individual child and his or her environments. Recent developmental theory and research on how early environments affect learning and later outcomes has two foci, which draw from a variety of theoretical perspectives. The first centers on how particular contexts, and increasingly interpersonal relationships and interactions, affect children's acquisition of specific skills. These studies are focused on answering questions about which types of experiences, on average, lead to learning specific knowledge or skills. Often these

studies exploit naturally occurring variability in population-based samples, and seek to explain individual differences in developmental processes (Sameroff, 2010).

An example of this developmental approach is research on the associations between the amount and type of language input that children receive in their home and the development of children's own language skills. Several studies have documented considerable variability in the amount of speech directed at children by caregivers during the course of a typical day (Hoff, 2003; Rowe, 2012). In turn, this variability in experience of speech is strongly linked to the child's later language expression and vocabulary, even after holding constant the children's own earlier vocalizations (Weisleder & Fernald, 2013; Rowe, 2012). Similarly, studies of parenting and children's self-regulation point to associations between parents' early support of their children's autonomy with later assessments of children's executive function (Bernier, Carlson, & Whipple, 2010).

A second and newer body of developmental research has focused on understanding how environmental experiences affect the underlying neurocognitive, biological, and physiological processes of development, rather than just observed differences in skill acquisition. For example, some studies seek to link environmental experiences to individual differences in developmental outcomes through stable and permanent changes in genetic expressions (Essex et al., 2013). While genetic endowments are largely invariant during development, there is considerable change in the epigenome – the biochemical system that regulates gene expression. Moreover, the epigenome has been found to be particularly responsive to environmental conditions (Champagne & Mashoodh, 2009). For example, animal studies have shown that experimental manipulation of the amount and timing of a rat grooming her pups is related to the pups' gene expression and subsequent developmental trajectories (Meaney, 2010). Although much of this

work began with studies of adverse events and animal models, increasingly such studies are extending to humans. For example, Essex and colleagues (2013) found that early maternal stressors were related to epigenetic changes in their children during adolescence, with implications for their mental health.

Economists' and developmentalists' differing models of development generate slightly different predictions for the effects of preschool investments. If focused on the preschool period, the Cunha and Heckman model implies that school readiness is a product of the child's cognitive and socioemotional skills upon entry into the preschool period plus preschool period investments from parents and possible ECE programs. Their hypothesis of dynamic complementarity implies that impacts of parental and ECE investments on child outcomes will be largest for children who enter the preschool period with the highest levels of cognitive and socioemotional skills. Indeed, this is the very evidence provided by Aizer and Cunha (2012).

Developmental theories take a broader look at the productivity of the likely match between what a program offers and what kinds of developmental supports a child most needs (Blair & Raver, 2012). Specifically, "compensatory" models argue that preschool investments can function effectively as a substitute for sensitive or enriched home environments, rather than a complement (Ramey & Ramey, 1998). Thus, children whose skill development may be compromised by economic disadvantage or low-quality home environments are predicted to benefit more from high-quality ECE programs than more advantaged children. In particular, if preschool settings are designed to expose children to sensitive caregiving environments, developmental theory would suggest that they would increase children's socioemotional skills much more among children with less sensitive caregivers in their home environments. Recent evidence supports these patterns of association (Watamura, Phillips, Morrissey, McCartney, &

Bub, 2011). This compensatory or protective model of high-quality early childhood care and education argues for understanding specific qualities and nature of differing types of concurrent and subsequent investments, as they pertain to differing domains of development.

Which Early Skills Matter for Human Capital Accumulation?

If early childhood programs seek to build skills that will generate lasting changes in adults' human capital, which skills should they target? Economists tend to lump IQ and achievement into a "cognitive" category and everything else into a "noncognitive" or "soft-skills" category. This is unhelpful for a variety of reasons. First, the "cognitive" category mixes cognitive ability with concrete academic skills such as literacy and numeracy. Although scores on tests of cognitive ability and achievement tend to be highly correlated, there is an important conceptual difference between cognitive ability as a relatively stable trait and the concrete achievement skills that develop in response to schooling and other human capital investments, including ECE. Moreover, "noncognitive" skills such as the ability to sustain attention when performing tasks, plan ahead, and control emotions in the face of provocation involve many of the same elements of brain circuitry as learning concrete skills, and are therefore inherently "cognitive."

Third and most important, conceptualizing and measuring distinct components of "noncognitive" skills is a vital first step in understanding why ECE and other human capital inventions have the impacts they do. Personality psychologists have centered their work on the "big five" personality traits – conscientiousness, openness, agreeableness, emotional stability, and extraversion – plus general cognitive ability. Research consistently shows that conscientiousness, and the related construct of "grit," have the largest correlations with education outcomes such as overall attainment and achievement (Almlund, Duckworth, Heckman, &

Kautz, 2011). Although personality traits have traditionally been viewed as relatively stable across the life span, some evidence indicates that they can change in response to life experiences and interventions (e.g., Almlund et al., 2011; Roberts, Walton, & Viechtbauer, 2006).

The early childhood literature uses a different taxonomy for children's skills and behavior. Moreover, developmental psychologists presume that children's skills and behaviors are dynamic rather than invariant and determined by the interplay between their skill potential and the quality of their early experiences – which may include early childhood education programs (National Research Council, 2000). They classify skills and behaviors in a number of ways, and some of their categories correlate with some of the "big five" personality traits.

One recent review classified important competencies into four groups: achievement, attention, externalizing behavior problems, and mental health (Duncan & Magnuson, 2011). Attention refers to the ability to control impulses and focus on tasks (e.g., Raver, 2004). Externalizing behavior refers to a cluster of related behavioral problems that include antisocial behavior, conduct disorders, and more general aggression (Campbell, Shaw, & Gilliom, 2000). Mental health constructs include anxiety and depression as well as somatic complaints and withdrawn behavior (Bongers et al., 2003). All of these skills and behaviors might respond to ECE investments and in turn help to determine subsequent educational attainment, skill development, and labor market participation.

The evidence base on how these early skills link to later earnings in adulthood is thin. Longitudinal datasets that have collected multiple domains of early childhood data and followed subjects through adulthood are rare, and often comprise convenience samples. Second, drawing causal conclusions from these nonexperimental studies is difficult because of confounding characteristics and contexts and the likely bidirectional nature of developmental processes.

Nevertheless, analysis of the British cohort studies and recent data from U.S. studies suggest that early achievement skills directly predict later earnings (Chetty et al., 2011; Currie & Thomas, 1999). Other studies provide related evidence that early skills will matter for human capital because of their links to adolescent achievement and attainment. Data from several large studies of young children find that when a constellation of skills and behavior are taken into account and differences in family background are held constant, early achievement skills (reading and math) are most important for predicting achievement later in childhood, followed by attention skills (sometimes measured by the lack of attention and hyperactivity) (Duncan et al., 2007). Somewhat surprising is that early problem behavior such as aggression or even prosocial behaviors did not predict later achievement. Extensive replication and reanalysis of several of the datasets included in the original analysis further confirmed this pattern of generally small and inconsistent associations between behavior and later achievement, once early achievement is taken into account (Grimm et al., 2010).

A slightly different picture of the role of early behavior is found, however, if the outcomes considered are educational attainment, later criminal activity, and earnings, rather than achievement skills. In the case of high school graduation, as would be expected concrete achievement skills play an important role. However, early problem behavior, and more specifically persistent antisocial behavior during middle childhood, is also predictive of high school completion, college attendance, and years of educational attainment (Magnuson, Duncan, Lee, & Metzger, 2014).

Evaluations of high-quality early childhood interventions such as Perry Preschool and Abecedarian Program also argue for the importance of early childhood development having long-run effects on criminal activity and earnings, though in both studies it is difficult to unpack

why the programs had long-run effects as they affected multiple, and differing, domains of development. Heckman and colleagues (2013) highlight the importance of "character" in producing reductions in crime, although this doesn't account for the long-run education impacts for Abecedarian since the program did not appear to affect children's behavior.

These analyses argue for the importance of early academic skills as an important input into the human production function. Yet where to target intervention domains should also be guided by where SES differences in development are most likely to be found. On this point, the data are very clear. Differences in development between more and less advantaged children are found early in life, with recent data pointing to differences by poverty status as early as 9 months of age (Halle et al., 2009). By school entry, family SES much more sharply differentiates children's early achievement skills than their early behavior. Table 2 provides data from the Early Childhood Longitudinal Study Kindergarten Cohorts of 1998 and 2010. Twelve years apart the studies collected similar data from nationally representative U.S. children and thus provide a useful comparison.[2] These data are collected for a large cohort of children during the fall of their kindergarten year and included one-on-one achievement skill assessments, as well as teacher and parent surveys. These data provide a snapshot of skills and behavior that children have at the start of their schooling careers.

The ECLS studies measured children's emerging reading skills with items that ask about identifying letters by name, associating letters with sounds at the beginning and end of words, and recognizing common words by sight. For math skills the assessments measure identifying one- and two-digit numerals, recognizing geometric shapes, counting, and filling in number sequences. Teachers also provided reports about children's classroom behavior. The Approaches to Learning scale includes elements of attention and task persistence as well as learning

independence, eagerness to learn, flexibility, and organization. The Externalizing subscale includes reports about the frequency of the following types of child behaviors: arguing, getting angry, fighting, acting impulsively, and disturbing activities. The internalizing subscale behavior problem scale covers frequency of behaviors such as anxiety, loneliness, low self-esteem, and sadness. For ease of comparison, we present differences among groups of children scaled by the measures' respective standard deviations. Thus, the estimates represent the fraction of a standard deviation.

Disparities in children's skills are evident along a number of relevant social dimensions. Girls outperform boys in reading and are reported by teachers to be better behaved. White children outperform black and Hispanic children in terms of reading and math and are rated as having better approaches to learning. Yet, the magnitude of these differences is dwarfed by those related to SES. Table 1 and Figure 1 show differences in these early school entry skills by SES quintiles. The lowest SES quintile corresponds to an average family income of about $15,500 (in 1998) and the highest quintile corresponds to incomes over $100,000 (in 1998). The differences between children who are in the top SES quintile (measured by a composite of parental education and family income) and the bottom quintile of SES are large. The difference in math and reading skills was between 1.2 and 1.3 standard deviations for reading and math in 1998, and it was just slightly smaller by 2010 at between 1.1 and 1.2 standard deviations. It is also notable that these differences are somewhat smaller but still quite large when the comparison is made only with children who have the same teacher.

Turning to children's attention, as measured by their teachers' response to questions forming the "approaches to learning" scale, bottom-to-top quintile differences across the SES spectrum are smaller though still sizable — about .60 standard deviations in 1998 and .50

standard deviations in 2010. For externalizing and internalizing behaviors, although SES differences are apparent, they are a quarter of the size of the SES-related skill gaps for reading and math skills. All in all, SES-based differences in concrete achievement skills are by far the largest.

It might be hoped that schools would help to equalize skill and behavior differences across the SES spectrum. For the 1998 cohort, the ECLS-K study was able to follow the same children over the course of elementary school and into middle school (Figures 2 and 3). For reading and math skills the magnitude of the gaps are similar over time, with if anything a slight increase in the size of the gaps. The largest increase in gaps is for externalizing behavior, although there is no evident change in teacher reports of internalizing behaviors, and related analyses describing patterns through eighth grade look similar (Magnuson, Waldfogel, & Washbrook, 2012).

Turning to other national data to explore SES differences in adult outcomes, in earlier work we reported that, compared with children in the top SES quintile, children in the lowest SES quintile have arrest rates 15 percentage points higher, high school completion rates 31 percentage points lower, and college attendance rates 40 percentage points lower (Duncan & Magnuson, 2011). In sum, large SES-related differences in early skills and moderate differences in aspects of behavior forecast later disparities in schooling and criminal involvement that have important implications for youth's experiences in the labor market.

Current Preschool Investments

The large and enduring differences in early skills, as well as their consequences for later learning, have not gone unnoticed by educators and policymakers. Indeed these very differences were the motivating reason for the creation of Head Start, for the expansion of state and local

prekindergarten programs, and most recently for President Obama's proposed expansion of enrollment in high-quality early learning programs. Despite advocates' and critics' focus on the findings from studies of a few specific programs, it is important to understand that preschool comprises a very heterogeneous set of programs, many of which have been evaluated and studied over time.

The rubric of *preschool* includes three broad types of programs serving children two years prior to kindergarten (ages 3 to 5): private preschool programs (which may be publicly paid for with child care subsidies), Head Start, and prekindergarten programs supported by state and local education funds. Preschool is provided in center-based classrooms, typically with children of similar ages with one or more teachers. Analysis of the National Household Education Survey data found on average 6.6 children per teacher in preschool classrooms (Mulligan, Brimhall, & West, 2005). Some programs are part time, providing services for a few hours per week (sometimes no more than 2 hours per day), while others are full time, providing services for 40 or more hours a week. About 75% of U.S. children attended a center-based preschool program the year prior to kindergarten and just over half attended a center-based program the year before that (at age 3) (Federal Interagency Forum on Child and Family Statistics, 2011).

Most preschool programs, particularly those with public funding and guidelines, use some form of curriculum to organize learning activities related to early academic skills – typically general concepts, early reading, and numeracy or math. These curricula may be either developed by the program itself or purchased from a commercial provider, and they differ in terms of the specificity of their content. Some provide lesson plans designed with a "whole child approach" including aspects that focus on multiple domains of development, and others target

specific skills, such literacy or math. As is the case in all educational settings, there is often considerable variability in the extent to which teachers implement curriculum as intended.

Many preschool programs share some common features in their daily routines. The morning often starts with a window of "drop-off" in which children engage in free-play activities. The more formal class time likely starts with "circle time" in which the lead teacher gathers all the children for a discussion of the day, a book reading, and perhaps singing of a song. Many curricula are based on thematic units, and this circle time is used to introduce concepts and themes that are embedded throughout the day in various forms of activities.

Group time is often followed with instructional activities, either involving the teacher talking with all children in the class or more often working with small groups of children at activity centers. Full-day programs include lunch, rest or naptime, and typically outdoor or other types of free-play in the afternoon. One study of public prekindergarten programs found that the 4-year-old children spend 53% of the morning in academic-related activities, which tended to focus on developing literacy skills, with only 6% of the time spent on math, and 23% on science or social studies (Bryant et al., 2002). Their mornings were divided, on average, into 23% in whole group activities, 33% in free-choice activities, 6% in small groups, 14% on meals, and 21% on transitions, bathroom breaks, and other routine activities. Although similar data are not systematically available for other types of programs, it is likely that these programs are among the more instructionally oriented that children experience.

Despite the similar features of preschool programs, there are key differences among types of programs. The oldest, largest, and best known federally funded preschool program is Head Start. Conceived as part of the Johnson administration's War on Poverty, Head Start has served over 31 million children since its inception in 1965 (USDHHS, ACF, OHS, 2014). Federal

guidelines require that 90% of the families served in each Head Start program be poor (i.e., with incomes below the federal poverty threshold), with the remaining 10% of children served by Head Start having developmental disabilities.

Head Start's 2013 budget was just under $7.6 billion, and funds were distributed to 1,591 local private and public nonprofit grantees serving just over 900,000 children. This amounts to an estimated average federal cost of $8,380 per child (USDHHS, ACF, OHS, 2014), but understates total program costs since local grantees are required to contribute 20% funding matches. Slightly less than half of Head Start programs provide full-week, full-time center care (Hamm, 2006). Head Start programs are designed to enhance the development of economically disadvantaged children using a holistic approach, including the provision of educational services and dental and mental health screening and access to care for the children, and parenting education and assistance in achieving education and employment goals for the parents (Puma et al., 2005). All Head Start centers are required to adopt a "whole-child" curriculum. A high priority is placed on parents' involvement in their children's education and the local administration of Head Start programs.

Public prekindergarten (pre-K) programs are a second form of publicly provided preschool, which are funded by states or local school districts. Most pre-K programs are targeted to low-income children (31 state programs have income eligibility requirements); however, a small but growing number of states either offer, or are currently considering, funding universal access for all 4-year-olds and, in some cases, 3-year-olds (Barnett, Carolan, Fitzgerald, & Squires, 2013). Pre-K initiatives are intended to complement, rather than supplant, existing sources of funding such as Head Start. Funding and enrollment in state pre-K programs have increased dramatically over the past several years. As of 2012 to 2013, 40 states (including the

District of Columbia) had pre-K initiatives serving approximately 28% of 4-year-olds and 4% of 3-year-olds (Barnett et al., 2013).

State prekindergarten programs vary substantially in terms of funding, program design, and quality across states. Although it is difficult to measure the costs of programs, the National Institute for Early Education Research (NIEER) calculates that the average per pupil state pre-K expenditure in 2012 to 2013 was $4,026 (Barnett et al., 2013). It is important to note, however, that, as with Head Start and its federal funding, per pupil state expenditures may underestimate actual spending per student, as many state pre-K programs receive local funding, and it is difficult to calculate such support. The majority of pre-K programs are either part time or have locally determined hours. Although some programs offer an extensive set of support services, such as transportation and health screenings and referrals, others offer very few. Most states use a mixed service delivery system that provides pre-K programming in schools as well as community-based settings, by contracting with privately run preschools and federally funded Head Start programs. Approximately one-third of children receiving pre-K services in 2011 were served outside public schools (Barnett et al., 2013).

Despite expansions in Head Start and pre-K programs, a large proportion of children still attend a “private” preschool in the year before they enter kindergarten. These programs are typically licensed or regulated by states as child care providers, and include both for-profit and not-for-profit entities. Barnett and Nores’s (2012) analysis of the NHES data finds, as would be expected, that participation in private preschool is most common among higher-income families, who are less likely to qualify for public programs.

Meta-Analysis of Short-Run Preschool Program Effects

What do we know about how children's preschool attendance affects their school readiness? Despite the hundreds of evaluation studies of early childhood education programs that have been published over the past 50 years, only a handful of programs have been prominently discussed in policy circles by advocates and critics: Perry Preschool, the Abecedarian program, Head Start, and more recently some state and local prekindergarten programs. These programs provide a selective view of what the impact of early education programs might be. Empirical studies of the effects of early childhood education on children's human capital encompass a range of methodologies and a wide variety of programs. Given the range of diverse programs that children experience, attention to the broader set of impacts and averages across programs seem most relevant and important. In a collaborative research project, we have focused on evaluations of preschool programs conducted over the course of the last half-century that are based on strong experimental or quasi-experimental methods and provide impact estimates for cognitive or achievement-related outcomes.[3]

Figure 4 shows the distribution of 84 program-average treatment effect sizes for cognitive and achievement outcomes, measured at the end of each program's treatment period, by the calendar year in which the program began. Reflecting their approximate contributions to weighted results, "bubble" sizes are proportional to the inverse of the squared standard error of the estimated program impact. The figure differentiates between evaluations of Head Start and other ECE programs and also includes a weighted regression line of effect size by calendar year.

Taken as a whole, the simple average effect size for early childhood education on cognitive and achievement scores was .35 standard deviations at the end of the treatment periods, an amount equal to nearly half of race differences in the kindergarten achievement gap found in the ECLSK data, but less than a quarter of the SES-related gaps. However, as can be seen from

Figure 2, average effect sizes vary substantially and studies with the largest effect sizes tended to have the fewest subjects. When weighted by the inverse of the squared standard errors of the estimates, the average drops to .21 standard deviations (see also Leak et al., 2014).

All of the 84 programs that generated the effect size data shown in Figure 4 met minimum standards for quality of research methods. However, some of the programs lasted for only a couple of summer months, while others ran for as long as 5 years. Some of the evaluations used random assignment while others relied on less rigorous quasi-experimental methods. Almost all focused on children from low-income families, but they varied in the racial and ethnic composition of treatment groups. Programs beginning before 1980 produced significantly larger effect sizes (.33 standard deviations) than those that began later (.16 standard deviations). Declining effect sizes over time are disappointing, as we might hope that lessons from prior evaluations and advances in the science of child development would have led to an increase in program impacts over time. However, the likely reason for the decline is that counterfactual conditions for children in the control groups in these studies have improved substantially (Duncan & Magnuson, 2013).

Studies of Preschool's Long-Run Effects

Another key issue is what we know about the long-run effects of these programs. One of the motivations for investing in early childhood education programs is that they will have important long-run benefits. Indeed, any discussion of preschool's potential to mitigate economic inopportunity or inequality hinges on there being discernable long-run effects on low-income children's later education, employment, earnings, and the like. The evidence is fairly clear on two issues. First is that short-term impacts on achievement skills dissipate over time. Estimates from our meta-analysis suggest that for each year after program impact the effects decline by

.025 standard deviations (Leak et al., 2014). That suggests that if the average program impact at the end of the program was .21 standard deviations, the treatment effect would be entirely gone 8 to 9 years after the program ended. Our results are well in line with those of other researchers who have sought to answer similar questions (Aos, Lieb, Mayfield, Miller, & Pennuci, 2004).

Despite the commonly found convergence between preschool attendees and other children's IQ scores or achievement skills, the studies that have followed early education participants beyond adolescence find a range of substantial program impacts on measures of young adult and adult human capital, including increases in educational attainment, reductions in criminal activity, and greater earnings (Currie & Almond, 2011). Studying the long-run effects of large, public programs, primarily Head Start, has required different methods than haves been used for model demonstration programs, because there were not early experimentally designed studies in which the participants were able to be followed over time (Garces, Thomas, & Currie, 2002; Johnson, 2011; Ludwig & Miller, 2007). Head Start is the most examined public program because of its large size and scope. These long-run studies have used a variety of econometric methods to construct appropriate comparison groups for preschool attendees, with particular concern that the same level of disadvantage is found among preschool attendees and non-attendees. For example, Garces and colleagues (2002) sibling fixed effect study using data from of the Panel Study of Income Dynamics found that among white children, attending Head Start was linked with additional 22 percentage points higher rates of high school graduation and 19 percentage points higher rates of college attendance. Among African American children, attending Head Start was linked with lower likelihoods of being charged or convicted of a crime (12 percentage points lower). Similarly, using the National Longitudinal Survey of Youth Child Supplement, Deming (2009) created an index of adult outcomes measured after age 18

(including high school graduation, college attendance, idleness, teen parenthood and health). Also using sibling comparisons and Deming (2009) finds Head Start is associated with a .23 standard deviation increase in this adult outcome index. Though the pattern of results from the long-run studies is consistent in term of positive effects on human capital, the fact that these studies do not have many measures beyond early achievement test scores means that the processes that lead to these long-run effects are essentially a "black box".

An evaluative study of the Chicago Child Parent Centers (CPC) by Reynolds and colleagues (Reynolds & Temple, 1998) is an exception, because it provides the only longitudinal evaluation of a large, public program. This study has followed over time a cohort of children who attended the Chicago Public School prekindergarten program and a matched comparison group. The results indicate that by age 28 it had improved participants' educational attainment and income, and reduced their criminal involvement as well as substance and drug abuse. Reynolds and colleagues' (Reynolds, Temple, & Ou, 2010) efforts to understand the early foundation of the program's later effects on occupational prestige and reduced crime and depressive symptoms pointed to a complex mediation model. They evaluated the explanatory role of children's own cognitive skills, motivation, and social adjustment as well as parent involvement and school support. Their findings pointed to substantial roles of children's cognitive skills, school support, and family support. In contrast, social adjustment and motivation made far smaller contributions to these adult outcomes.

Studies of small, model demonstration programs also yield a pattern of significant long-run improvement in a variety of human capital production dimensions. In addition, these studies typically have a richer range of measures at program completion, offering the opportunity to better understand how early education programs generate these long-run effects. Yet, what

emerges from these studies is a complicated puzzle. Heckman and colleagues (2010) argue that the Perry Preschool program had its most important effects not on children's academic skills, but on their character. Indeed, reductions in criminal activity in adulthood for male participants were the most cost-saving aspects of the program's positive long-run effects, and appear to have been most closely linked to changes in children's earlier behavior. The importance of behavior may constitute a significant part of the economic case for Perry Preschool, but it doesn't appear to generalize more broadly to other preschool studies.

Abecedarian, an intensive, high-quality early education program that began in the first year of life and lasted through school entry, demonstrated positive effects on adult human capital outcomes, but both the range of outcomes affected and the possible explanatory pathways appear to differ in important ways from those found among Perry Preschool attendees. For example, the most recent study of Abecedarian participants found that by age 28 children who attended the program were more likely to be college graduates than the control group, and described substantial treatment differences in earnings, although these did not rise to the level of statistical significance (Campbell et al., 2012). However, there was no apparent treatment difference in measures of criminal conviction, and this may have been foreshadowed by early study findings that indicated no reductions in problem behavior for program participants (Clarke & Campbell, 1998).

Taken together, these studies support the argument that the implementation of these programs several decades ago may have important effects on later human capital accumulation. Studies suggest that a variety of programs, with differing designs and emphasis, have important long-run effects. Beyond that general statement, the details of which programs affected which adult outcomes through what possible mechanisms remain unclear. Yet, the fact that prior studies

show that the possible positive human capital-generating outcomes are multiple and diverse, as are the pathways by which preschool may produce these outcomes, should be reassuring. Efforts to increase children's participation in a range of programs of reasonably high quality are likely to yield long-run effects, even if the specific pathways are diverse or not fully understood, and even if programs' boost to achievement does not last in perpetuity. However, if policy goals are broader than increasing access to quality programs — for example, increasing the magnitude of long-run effects — then it would seem that greater information about the potential pathways and mechanisms by which early skills and behaviors turn into longer-run outcomes is needed. In this case, we need to know which skills and program design features to improve in order to yield larger long-run effects, and efforts that would boost early academic skills might differ from those that might more directly target socioemotional skills, behavior, or self-regulation.

The Costs of Expanding Preschool Access

Recent trends have suggested that through the late 1990s rates of preschool attendance were climbing among all SES groups, though rates of attendance continue to lag for lower-income children. In recent work, with others (Duncan & Magnuson, 2013; Magnuson & Waldfogel, 2012), we tracked enrollment trends from 1968 to 2010 for 3- and 4-year-old children, using nationally representative data from the October Current Population Survey and dividing families into five groups based on income quintiles (Figure 5). We found that while enrollment in preschool has grown for 3- and 4-year-olds from all income groups over time, enrollment rates are consistently higher for the top two income groups than for the middle and lower income groups (see Figure 3). Looking more closely at the year before children enter school, recent estimates indicate that 75% of children have attended a preschool-like program.

As would be expected, this rate is higher among the top income quintiles (nearly 90%) and lower among the three bottom income quintiles (64–69%).

Public investments have clearly played a role in boosting enrollment among low-income children. The cost of early education programs is typically expensive, with the median state average cost of full-time private preschool (center-based care) at about $8,000 per year. Without public investments to offset the price, the expense of private preschools is often prohibitive for many low-income and even middle-income families. Earlier research found that as kindergarten programs for 5-year-olds were expanded in the 1960s and 1970s, income-related gaps in enrollment for that age group disappeared (Bainbridge et al., 2005). Our analyses of trends for 3- and 4-year-olds indicated that in years when funding for subsidies, Head Start, or prekindergarten has been more generous, a greater share of low-income children have been enrolled (Magnuson, Meyers, & Waldfogel, 2007). Though expansions have no doubt been important to boosting enrollments, they have not been generous enough for all low-income children to benefit.

There are other demographic groups that have comparatively low levels of preschool enrollment — Hispanic children and children of immigrants. No doubt part of the lower rates of enrollment is attributed to the fact that on average their families have lower incomes, but African American children, in contrast, are if anything more likely than comparable white children to be enrolled in school- or center-based care (Magnuson, Lahaie, & Waldfogel, 2006; Magnuson & Waldfogel, 2005). Indeed, both language barriers and cultural factors are also likely influences that play a role in the lower levels of enrollment among Hispanic children and children of immigrants (Takanishi, 2004). A final factor that has been overlooked in discussions of increased enrollment are rural communities, in which transportation is a likely impediment. Indeed, rates of

preschool enrollment are substantially lower in rural communities and recent estimates suggest that less than half of 4-year-olds in rural communities attend preschool, compared with nearly 80% of their urban or suburban peers.

Could we reach near universal enrollment in prekindergarten or preschool programs? The answer is almost certainly, yes. In a relatively short period of time, kindergarten was introduced and became universal. Other countries, most notably France, have near universal attendance in public programs. What would it cost to do this in the U.S.? There are roughly 4 million 4-year-olds, three-quarters of whom attend some form of preschool already. The cost of providing public education per child could range from $5,000 for half-day programs to about $10,000 for full-day programs. Currently 4-year olds are in a roughly equal mix of both part-day and full-day programs (Barnett & Nores, 2012), and providing this mix of hours seems important to serving the needs and desires of parents. If we assume that those who are not currently attending would have the same distribution across full- and part-time programs, this yields an average cost of $7,500 per newly attending child. With these assumptions, the added cost for reaching 100% enrollment would be $7.5 billion ($7,500 for 1 million 4-year-olds). [4]

But of course, that exercise is foolish, because we cannot devise a policy that would pay only for those children whose parents did not otherwise enroll them. When public programs are available, some proportion of children whose parents are currently paying for care will shift to a publicly provided program (Casio & Schanzenbach, 2013). If the public program paid costs for all 4-year-olds, the price tag would be $30 billion ($7,500 for 4 million 4-year-old children). Over $5.12 billion is already being spent on state prekindergarten and $8.5 billion on Head Start, therefore the marginal new public investment would amount to $16.35 billion for a mix of full- and part-day program slots. We expect that once the cost of child care subsidies was taken into

account, this price would fall by possibly $1 billion. It is also certain that, as is the case for public education, some proportion of families would prefer to pay for a private preschool than participate in public programs. If a similar proportion chose private preschools as choose private K-12 schools (10%), that would suggest a price tag of $27 billion and a marginal public investment of $13.35 billion.

A key question, however, is whether the public investment should be attempting to offset the costs for all families, or whether resources should be targeted to those in lower-income families. With limited public resources, there is a compelling reason to focus on providing access for low-income families, rather than offsetting costs for higher-income families. As a result, for non-universal public preschool programs it is useful to think about categorical eligibility based on residence in low-income communities or family income (or a combination of similar characteristics). In particular, the bottom three income quintiles (60%) all share similarly low rates of enrollment, compared with the upper two income quintiles (see Figure 5). Combined, roughly 52% of these income groups are either not enrolled or enrolled in private programs (Barnett & Nores, 2012). Thus, the cost for publicly providing for all these children who are currently not enrolled in publicly funded programs would require new public investments on the order of $9.36 billion ($7,500 per child for 1.248 million children). For purposes of comparison, this is a little more than currently being spent on Head Start, and a little less than twice what is already being spent on state pre-K programs.

What Are the Benefits of Increasing Preschool Enrollment Likely to Be?

Research efforts to quantify the benefits of early childhood education are accumulating. Three studies that have followed children through adulthood have based their estimates on actual

collected data, and at least one other study of the Tulsa pre-K program by Bartik and colleagues (2012) has projected benefits. We generally adopt the methods used in projections with some minor modifications. First, it is relevant to note that Bartik and colleagues' paper uses estimates from Chetty et al.'s (2011) analysis of the Tennessee Star experiment that correlates early test scores to later adult earnings. This work finds that at the end of kindergarten a percentile increase in test scores is associated with approximately .5% increase in adult earnings. Predicting adult earnings based on a program's impacts on test score percentiles, Bartik and colleagues (2014) compare predicted earnings impacts based on Chetty et al.'s (2011) estimates to measured earnings impacts for available long-run follow-up studies, and find that prediction model seems to do reasonably well. For example, based on the prediction model earnings would increase by about 16%, 10%, and 8% for Perry Preschool, Abecedarian, and the CPC, respectively. The actual earnings effects in these programs were 19%, 14%, and 7%, respectively. This provides Bartik et al. (2014) with the necessary supporting evidence to conclude that Chetty et al.'s (2011) parameter provides a good approximation for the effects of test scores at the end of a preschool program rather than just kindergarten.[5]

We take the approach of estimating the lowest possible effect size that would justify an expenditure of $7,500 per child to fund the program expansion. That is, how much of an effect size might the program have to generate to "break even" in terms of costs and benefits? In prior work, we have used the 2013 March Current Population Survey (CPS) data to estimate the present value of lifetime earnings for adult high school graduates and adults with some college (ages 20–65, including 0 earnings for non-workers, we think a reasonable group for this exercise) (Magnuson, Brooks-Gunn, & Waldfogel, 2014). In predicting lifetime earnings, there are numerous assumptions to make. Here, we offer a low and high estimate of lifetime earnings. For

the low estimate, we assume no wage growth over time, and 3% discount to age 5, the resulting present value of lifetime earnings for workers with a high school degree or some college is about $382,392. At $7,500 per child, the preschool investment would represent 2% of lifetime earnings, suggesting that the program would need to result in an on average 4% increase in test scores (given the .5% association between test score percentile and earnings). Translating from test score percentiles to effect sizes suggests that this would be an effect size of .095 standard deviations for someone who was at the 50% percentile, but closer to .15 effect size for those scoring at just the 25^{th} (or 75^{th}) test score percentile.

For the high estimate of lifetime earnings, we assume that earnings should be inflated by 20% to reflect the value of benefits (such as health insurance and retirement contributions) and that there is a 1% wage growth per year, but we still use the 3% discount rate to age 5. This results in a much larger estimate of lifetime earnings of about $681,544. Under these assumptions, $7,500 amounts to just 1.1% of earnings. Based on Chetty et al.'s (2011) estimates, this would suggest that impacts of just 2.5 percentile points would be minimally necessary to equalize costs and benefits in terms of later earnings. This would translate into an effect size of .035 at the 50^{th} percentile and to about .07 effect size for closer to the 25^{th} (or 75^{th}) percentile test scores. These projections should be interpreted as rough approximations, and with the appropriate understanding that any projection into the future involves significant uncertainty.

Taken together, these rough benefit estimates seem to suggest that even if expansions to preschool programs yield relatively small effect sizes in improvements in academic skills, the spending and expected returns are likely in the very least to break even, and to bring increased income and economic opportunity. As has been found in long-run evaluations and in recent projections, a program that yields substantial impacts on academic skills will have earnings

benefits that well exceed the program costs (Heckman et al., 2010). Even if program impacts on later earnings were the only benefit to be considered, the exercise indicates that increasing enrollments in preschool programs is likely to be a wise investment. That said, as the benefit-cost evaluations of model preschool programs show, earnings is a major component of the societal benefits, but by far not the only one. They have amounted to from one-third to one-half of estimated program benefits in prior studies. Reductions in spending for special education, grade repetition, child protection services, public welfare benefits, and crime are important potential benefits, with potentially large payoffs (Heckman et al., 2010; Barnett & Masse, 2007; Temple & Reynolds, 2007). Yet, in each study the particular of the program and the children and families being served has resulted in a different set of non-earnings benefits. In Perry Preschool, a large category was reductions in crime. For CPC, both reductions in crime and reductions in participation in the child welfare system were important. For Abecedarian, benefits were counted from increased maternal employment early in life and later improvements in health. All of this suggests that estimating the specifics of likely additional benefits is hard to do, and requires a careful understanding of the populations that will be served by expanded funding. Nonetheless, if prior studies provide a useful guide, then other benefits are likely to amount to an important return on the investment.

If the focus of policy attention is on improving inequality in economic fortunes, then it is important to note that our calculations of increased earnings do not differentiate between the magnitudes of earnings gains for children from differing economic backgrounds. The extent to which both short- and long-run program impacts differ by family background along a number of relevant demographic characteristics is not fully understood. There is some evidence that effect sizes might be slightly larger for children from more disadvantaged backgrounds, but the

differences are often relatively small and not always significantly different (Burchinal, Magnuson, & Powell, 2014). Thus, although preschool may increase later economic productivity, it is likely to do so for all participating children. As such, it is unlikely to function as an equalizer of economic opportunity between disadvantaged and affluent children.

What Other Types of Investments Should Be Considered?

Focusing only on preschool enrollment in the year prior to kindergarten as the key target of investments is a very narrow approach to improving young children's development and building human capital. Several other types of investments need greater policy attention, and in particular development and evaluation. First, improving the effectiveness of preschool instruction is an important pathway to improving children's skills. Increasingly, evidence suggests that one of the most important opportunities to increase children's learning is by selecting good curricula and supporting teachers as they implement it. Curricula and related professional development that are intensive, focused, developmentally appropriate, and sequential can have especially positive impacts on early childhood instruction and on children's learning (Burchinal, Magnuson, & Powell, 2014).

A seminal comparative evaluation of preschool curricula was conducted by the Preschool Curriculum Evaluation Research (PCER) initiative. The impact of 14 different curricula implemented in early childhood classrooms serving primarily low-income children was assessed (Preschool Curriculum Evaluation Research Consortium, 2008). In each of 12 different projects, early childhood classrooms or centers were randomly assigned to a target curriculum or to a control condition, typically the local business-as-usual curriculum. Unfortunately, the studies were challenged by their cluster design, and the low statistical power it generated for analyzing

impacts. During the pre-K year, 8 of the 14 curricula had a positive impact on teacher instruction, but only two had statistically significantly positive effects on child outcomes (effect sizes of .32 to .96). A recent reanalysis of these data by Duncan and colleagues (2014), which pools across curricula based on their content in order to better detect significant small to moderate effects, concluded that content-specific curricula focused on literacy and math are better able to promote academic skills, compared with more general "whole-child" curricula.

The Building Blocks math program is illustrative of a recently developed curriculum focused on a specific developmental domain. Developed by Sarama and Clements (2004), the curriculum includes large- and small-group instruction focused on teaching math skills in a focused and sequential manner, and hands-on and computer activities that promote children's active involvement in solving problems and explaining their solutions. An experimental evaluation found that the curriculum resulted in large improvements in children's math knowledge when compared with a different math curriculum (effect size of .47) and a business-as-usual control group (effect size of 1.07) (Clements & Sarama, 2008).

An example of a public preschool program that has taken seriously the need to identify exemplary curricula and implement them well is the Boston Pre-Kindergarten Program. The program developed their curriculum by integrating proven literacy, math, and social skills interventions. The academic component combined two curricula, Building Blocks for math instruction and Opening the World of Learning for language and literacy. Extensive teacher training and coaching was provided. The rigorous evaluation indicated large impacts on vocabulary, math, and reading (effect sizes of .45 to .62) and somewhat smaller impacts on executive functions (effect sizes of .21 to .28) (Duncan & Murnane, 2013; Weiland & Yoshikawa, 2013).

While evidence is accumulating, much more research related to preschool curriculum development and evaluation is needed. This work is critically important, but not easy for several reasons. First, the costs associated with successful implementation are not negligible, often requiring substantial investments in materials and teacher training time. Second, there are often non-pecuniary obstacles to overcome. In general, the early childhood education workforce often works long hours for low salaries, which often results in workers with low levels of education and high rates of job turnover. Sometimes, these circumstances can make implementation challenging, especially in community-based settings. Finally, the associated research costs are often quite high because it is expensive to conduct multisite experimental evaluations that include individual child assessments.

Finally, all the discussion of preschool leaves out infants and toddlers. These earliest years of life are also an important period of development, and warrant greater policy and programmatic attention. The models of early learning programs that are developmentally appropriate for preschoolers cannot be simply extended downward for younger children at the same cost for the same effect. Some model home visiting programs and parenting programs for mothers of infants have also demonstrated the potential to have important impacts on children's trajectories, with potential implications for human capital accumulation (Olds, Sadler, & Kitzman, 2007). Yet, at this time what is most needed is continued efforts to innovate and evaluate the feasibility and effectiveness of theoretically informed interventions for very young children.

Conclusions

Development during early childhood provides an important foundation for human capital development, with important long-run links to economic earnings and opportunity later in life. The accumulated evidence suggests that there are multiple aspects of early skills – achievement, behavior, and mental health—that if improved early in life can improve children's life chances. Moreover, there is accumulating evidence that attending good quality preschools for a year or two results in long-lasting improvements in educational attainment and earnings, even when short-term improvements in concrete achievement skills fade during the elementary school years. The processes by which these changes occur, however, seem to vary depending on the populations being served and the emphasis of the programs. Taken together, this argues for the importance of early childhood investments as a way to increase economic opportunity.

Currently, about 25% of children do not attend preschool before they enter kindergarten. Because low-income children are least likely to be enrolled compared with higher-income children, and because income gaps in early development forecast lower levels of human capital accumulation, improving attendance should be a first priority for policy. Efforts to expand enrollment will also need to consider other potential barriers such as language and program location. We estimate the costs of providing publicly funded preschool, a mix of part- and full-day programs, for all children in the bottom three income quintiles. We estimate that this would cost an additional $9.6 billion. Our consideration of the potential benefits finds that programs that have relatively small effects on children's achievement are projected to "break even" (our high estimate of income, which is more conservative, would require a .09–.15 standard deviation impact on achievement to break even). Prior studies have demonstrated that programs that produce more substantial effects will yield much larger returns on investments. Although all

efforts to forecast years in the future involve uncertainty, we read the evidence to point toward the importance of increased investments in public preschool programs. Other targets for investment include improving learning through research-based curricula and programs for infants and toddlers.

References

Aizer, A., & Cunha, F. (2012). *The production of human capital: Endowments, investments and fertility* (No. w18429). National Bureau of Economic Research.

Almlund, Mathilde, Angela L. Duckworth, James J. Heckman, and Tim Kautz. 2011. Personality Psychology and Economics. IZA Discussion Paper 5500.

Aos, Steve, Roxanne Lieb, Jim Mayfield, Marna Miller, and Annie Pennucci. 2004. Benefits and Costs of Prevention and Early Intervention Programs for Youth. Olympia: Washington State Institute for Public Policy.

Bainbridge, J., Meyers, M. K., Tanaka, S., & Waldfogel, J. (2005). Who gets an early education? Family income and the enrollment of three- to five-year-olds from 1968 to 2000. *Social Science Quarterly, 86*(3), 724-745.

Barnett, W. Steven and Leonard N. Masse. 2007. Comparative Cost-Benefit Analysis of the Abecedarian program and Its Policy Implications. *Economics of Education Review 26*(1): 113-125.

Barnett, W. S., & Nores, M. (2012). Estimated participation and hours in early care and educaiton by type of arrnagement and income at ages 2 to 4 in 2010. New Brunswick, NJ: Rutgers, NIEER. Retrieved from: http://www.nieer.org/publications/nieer-working-papers/estimated-participation-and-hours-early-care-and-education-type.

Barnett, W. Steven, Megan Carolan, Jen Fitzgerald, and J. H. Squires. 2013. The State of Preschool 2011: State Preschool Yearbook. National Institute for Early Education Research, Rutgers. Retrieved from: http://nieer.org/sites/nieer/files/yearbook2013.pdf

Bartik, T. J., Gormley, W., & Adelstein, S. (2012). Earnings benefits of Tulsa's pre-K program for different income groups. *Economics of Education Review, 31*(6), 1143-1161.

Bernier, A., Carlson, S. M., & Whipple, N. (2010). From External Regulation to Self-Regulation: Early Parenting Precursors of Young Children's Executive Functioning. *Child Development, 81*(1), 326-339. doi: 10.1111/j.1467-8624.2009.01397

Blair, C., & Raver, C. C. (2012). Child Development in the Context of Adversity Experiential Canalization of Brain and Behavior. *American Psychologist, 67*(4), 309-318.

Bongers, Ilja L., Hans M. Koot, Jan van der Ende, and Frank C. Verhulst. 2003. The Normative Development of Child and Adolescent Problem Behavior. *Journal of Abnormal Psychology 112* (5): 179–92.

Bryant, D., Clifford, R., Early, D., Pianta, R., Howes, C., Barbarin, O., et al. (2002, November). Findings from the NCEDL Multi-State Pre-Kindergarten Study. Annual meeting of the National Association for the Education of Young Children, New York.

Burchinal, M. Magnuson, K., & Powell, D. (in press). Children in Early Care and Education Chapter to appear in M. Bornstein and T. Levelthal (Eds). *Handbook of Child Psychology and Developmental Science, Volume 4: Ecological Settings and Processes in Developmental Systems.*

Campbell, S. B., Shaw, D. S., & Gilliom, M. (2000). Early externalizing behavior problems: Toddlers and preschoolers at risk for later maladjustment. *Development and Psychopathology, 12*(3), 467-488.

Campbell, Frances A., Craig T. Ramey, Elizabeth Pungello, Joseph Sparkling, and Shari Miller-Johnson. 2002. Early Childhood Education: Young Adult Outcomes from the Abecedarian Project. *Applied Developmental Science 6* (1): 42-57.

Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., . . . Ramey, C. T. (2012). Adulty Outcomes as a Function of an Early Childhood Educational Program: An Abecedarian Project Follow-Up. *Developmental Psychology, 48*(4), 1033-1043.

Cascio, E. U., & Schanzenbach, D. W. (2013). The Impacts of Expanding Access to High-Quality Preschool Education. *Brookings Papers on Economic Activity*, 127-192.

Champagne, F. A., & Mashoodh, R. (2009). Genes in Context Gene–Environment Interplay and the Origins of Individual Differences in Behavior. *Current Directions in Psychological Science, 18*(3), 127-131.

Chetty, Raj, John Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Schanzenbach, and Danny Yagan. 2011. How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR, *Quarterly Journal of Economics* 126(4): 1593-1660.

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

Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal, 45*(2), 443-494.

Cunha, Flavio, and James J. Heckman. 2007. The Technology of Skill Formation. *American Economic Review 97* (2): 31–47.

Currie, J., & Almond, D. (2011). Human capital development before age five. *Handbook of labor economics, 4*, 1315-1486.

Currie, J., & Thomas, D. (1999). *Early test scores, socioeconomic status and future outcomes* (No. w6943). National bureau of economic research.

Deming, D. (2009). Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start. *American Economic Journal-Applied Economics, 1*(3), 111-134.

Duncan, Greg, and Katherine Magnuson. 2011. The Nature and Impact of Early Achievement Skills, Attention and Behavior Problems. In Greg J. Duncan and Richard J. Murnane (eds.), *Whither Opportunity: Rising Inequality, Schools, and Children's Life Chances*, New York: Russell Sage, pp. 47-69.

Duncan, G. J., & Magnuson, K. (2013). Investing in Preschool Programs. *Journal of Economic Perspectives, 27*(2), 109-131.

Duncan, G. J., & Murnane, R. J. (2103). *Restoring opportunity : the crisis of inequality and the challenge for American education*. Cambridge, MA: Harvard Education Press.

Duncan, G. J., Ludwig, J., & Magnuson, K. A. (2007). Reducing poverty through preschool interventions. *Future of Children, 17*(2), 143-160.

Duncan, G. J., Claessens, A., Huston, A. C., Pagani, L. S., Engel, M., Sexton, H., et al. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428-1446.

Duncan, G., Auger, A., Burchinal, M., Domina, T. & Bitler, M. (2014). Boosting School Readiness with Preschool Curricula and Quality. UC Irvine, Manuscript.

Essex, M. J., Thomas Boyce, W., Hertzman, C., Lam, L. L., Armstrong, J. M., Neumann, S. M. A., et al. (2013). Epigenetic Vestiges of Early Developmental Adversity: Childhood Stress Exposure and DNA Methylation in Adolescence. *Child Development, 84*(1), 58-75.

Fox, S. E., Levitt, P., & Nelson, C. A. (2010). How the Timing and Quality of Early Experiences Influence the Development of Brain Architecture. *Child Development, 81*(1), 28-40.

Federal Interagency Forum on Child and Family Statistics. (2011). *America's children: Key national indicators of well-being, 2011*. Washington, DC: U.S. Government Printing Office. Retrieved from http://www.childstats.gov/pdf/ac2011/ac_11.pdf

Garces, E., Thomas, D., & Currie, J. (2002). Longer-term effects of head start. *American Economic Review*, *92*(4), 999-1012.

Gibbs, Chloe, Jens Ludwig, and Doug Miller, D. 2012. Does Head Start Do Any Lasting Good? NBER Working Paper 17452.

Grimm, K. J., Steele, J. S., Mashburn, A. J., Burchinal, M., & Pianta, R. C. (2010). Early behavioral associations of achievement trajectories. *Developmental Psychology, 46*(5), 976-983.

Halle, T., Forry, N., Hair, E., Perper, K., Wandner, L., Wessel, J., & Vick, J. (2009). Disparities in early learning and development: lessons from the Early Childhood Longitudinal Study–Birth Cohort (ECLS-B). *Washington, DC: Child Trends*.

Hamm, K. (2006). *More than meets the eye: Head Start programs, participants, families, and staff in 2005*. Policy Brief No. 8. Washington, DC: Center for Law and Social Policy (CLASP). Retrieved from http://files.eric.ed.gov/fulltext/ED494132.pdf

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review, 103*(6), 2052-2086.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics, 94*(1-2), 114-128.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child development*, *74*(5), 1368-1378.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences of the United States of America, 103*(27), 10155-10162.

Leak, James, Greg Duncan, Weilin Li, Katherine Magnuson, Holly Schindler, and Hirokazu Yoshikawa. 2014. Is Timing Everything? How Early Childhood Education Program Cognitive and Achievement Impacts Vary by Starting Age, Program Duration and Time since the End of the Program. Under review.

Ludwig, J., & Miller, D. L. (2007). Does head start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics, 122*(1), 159-208.

Magnuson, Katherine, Claudia Lahaie, and Jane Waldfogel. 2006. Preschool and School Readiness of Children of Immigrants. *Social Science Quarterly 87*: 1241-1262.

Magnuson, Katherine, Marcia Meyers, and Jane Waldfogel. 2007. Public Funding and Enrollment in Formal Child Care in the 1990s. *Social Service Review, 81* (1): 47-83.

Magnuson, K. A., & Waldfogel, J. (2005). Early childhood care and education: Effects on ethnic and racial gaps in school readiness. *Future of Children, 15*(1), 169-196.

Magnuson, K., Waldfogel, J. & Washbrook. (2012). The Development of SES Gradients in Skills during the School Years: Evidence from the US and UK. In J. Ermisch, M. Jantti, & T. Smeeding (eds.), *From Parents to Children: The Intergenerational Transmission of Advantage.* NY: Russell Sage Foundation.

Magnuson, K., Duncan, G., Lee, Y & Metzger, M. (2014b). Elementary School Adjustment and Educational Attainment? Manuscript.
Meaney, M. J. (2010). Epigenetics and the Biological Definition of Gene × Environment Interactions. *Child Development, 81*(1), 41-79.
Mulligan, G. M., Brimhall D., & West, J. (2005). *Child care and early education arrangements of infants, toddlers, and preschoolers: 2001*. Statistical Analysis Report, NCES 2006-039. Jessup, MD: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubs2006/2006039.pdf
National Association of Child Care Resource & Referral Agencies. 2011. *Parents and the High Cost of Child Care.* Arlington, VA: NACCRRA.
National Research Council. 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Shonkoff, Jack, and Deborah Phillips, eds. Washington, D.C.: National Academy Press.
Olds, David L., Lois Sadler, and Harriet Kitzman. 2007. Programs for Infants and Toddlers: Recent Evidence from Randomized Trials. *Journal of Child Psychology and Psychiatry, 48* (3-4): 355-391.
Preschool Curriculum Evaluation Research Consortium (PCER). (2008). *Effects of Preschool Curriculum Programs on School Readiness: Report from the Preschool Curriculum Evaluation Research Initiative.* Washington DC: National Center for Education Research. Retrieved from http://www.researchconnections.org/childcare/resources/14449/pdf
Puma, M., Bell, S., Cook, R., Heid, C., & Lopez, M. (2005). *Head Start Impact Study: First year findings*. Washington, DC: US Department of Health and Human Services, Administration for Children and Families. Retrieved from http://www.acf.hhs.gov/sites/default/files/opre/first_yr_finds.pdf
Purtell, K. M.,& Gershoff, E. T. (2013). The "skill begets skill" hypothesis and the experimental effects of Head Start on children's academic skills and social behaviors. Paper presented at the Association for Public Policy Analysis and Management Annual Meeting, Washington, D.C.
Ramey, Craig T., and Sharon Landesman Ramey. 1998. Early Intervention and Early Experience. *American Psychologist 53*: 109-120.
Ramey, C. T., & Ramey, S. L. (2004). Early learning and school readiness: Can early intervention make a difference? *Merrill-Palmer Quarterly, 50,* 471–491. doi:10.1353/mpq.2004.0034
Raver, C. Cybele. 2004. Placing Emotional Self-Regulation in Sociocultural and Socioeconomic Contexts." *Child Development 75* (2): 346–53.
Reynolds, A. J., & Temple, J. A. (1998). Extended early childhood intervention and school achievement: Age thirteen findings from the Chicago Longitudinal Study. *Child Development, 69*(1), 231-246
Reynolds, A. J., Temple, J. A., & Ou, S. R. (2010). Preschool education, educational attainment, and crime prevention: Contributions of cognitive and non-cognitive skills. *Children and Youth Services Review, 32*(8), 1054-1063.
Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological bulletin, 132*(1), 1.
Rowe, M. L. (2012). A Longitudinal Investigation of the Role of Quantity and Quality of Child-Directed Speech in Vocabulary Development. *Child Development, 83*(5), 1762-1774.

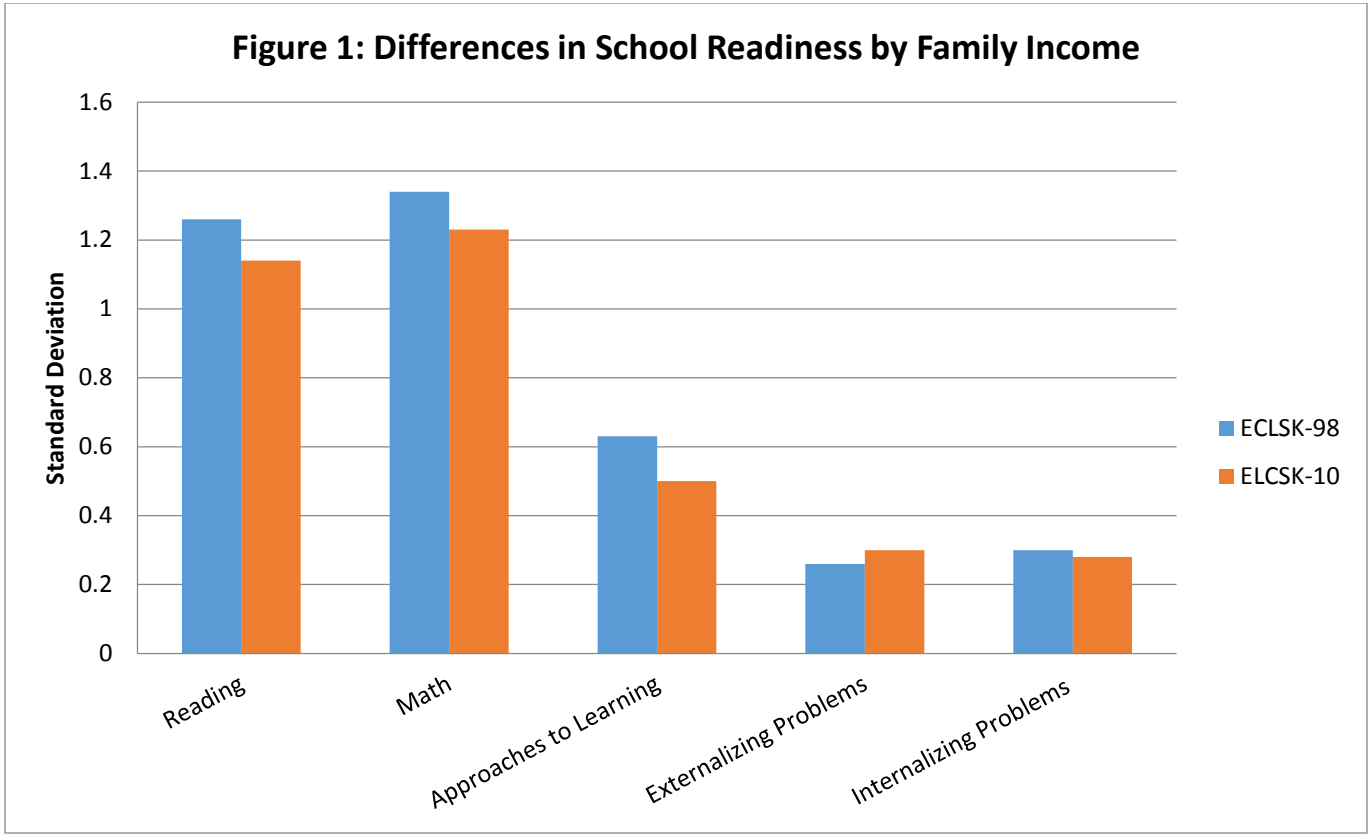
Sameroff, A. (2010). A Unified Theory of Development: A Dialectic Integration of Nature and Nurture. *Child Development, 81*(1), 6-22.
Sapolsky, Robert. 2004. Mothering Style and Methylation. *Nature Neuroscience 7* (8): 791-792.
Sarama, J., & Clements, D. H. (2004). Building Blocks for early childhood mathematics. *Early Childhood Research Quarterly, 19*(1), 181-189.
SECCYD Experiencing Home and Child-Care Environments That Confer Risk. *Child Development, 82*(1), 48-65
Shonkoff, J. P. (2010). Building a New Biodevelopmental Framework to Guide the Future of Early Childhood Policy. *Child Development, 81*(1), 357-367.
Takanishi, R. (2004). Leveling the playing field: Supporting immigrant children from birth to eight. *Future of Children, 14*(2), 61-79.
Temple, J. A., & Reynolds, A. J. (2007). Benefits and costs of investments in preschool education: Evidence from the Child-Parent Centers and related programs. *Economics of Education Review, 26*(1), 126-144.
U.S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start. (2014). *Head Start program fact sheet fiscal year 2013*. Retrieved from http://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/docs/hs-program-fact-sheet-2013.pdf
Watamura, S. E., Phillips, D. A., Morrissey, T. W., McCartney, K., & Bub, K. (2011). Double Jeopardy: Poorer Social-Emotional Outcomes for Children in the NICHD
Weiland, C., & Yoshikawa, H. (2013). Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills. *Child Development, 84*(6), 2112-2130.
Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science, 24*(11), 2143-2152.

Table 1: Gaps in Children's Academic and Behavior Skills in the Fall of Kindergarten

	Reading		Math		Approaches to Learning		Lack of Externalizing Behavior[a]		Lack of Internalizing Behavior	
ECLSK-10	Unadj.	Teacher FE	Unadj.	Teacher FE	Unadj.	Teacher FE	Unadj.	Teacher FE	Unadj.	Teacher FE
Boys/Girls	0.11	0.11	-0.03	-0.03	0.43	0.44	0.42	0.42	0.07	0.06
Black/White	0.30	0.23	0.57	0.36	0.25	0.32	0.29	0.29	0.08	0.06
Hispanic/White	0.52	0.31	0.64	0.35	0.12	0.12	0.02	0.10	0.01	-0.02
SES: 1st quintile/5th quintile	1.14	0.95	1.23	0.87	0.50	0.72	0.30	0.39	0.28	0.39
SES: 1st quintile/3rd quintile	0.52	0.39	0.61	0.39	0.25	0.22	0.12	0.09	0.15	0.16
SES: 3rd quintile/5th quintile	0.62	0.49	0.62	0.43	0.26	0.34	0.17	0.16	0.13	0.12

	Reading		Math		Approaches to Learning		Lack of Externalizing Behavior[a]		Lack of Internalizing Behavior	
ECLSK-98	Unadj.	Teacher FE	Unadj.	Teacher FE	Unadj.	Teacher FE	Unadj.	Teacher FE	Unadj.	Teacher FE
Boys/Girls	0.17	0.15	0.03	0.01	0.40	0.39	0.41	0.39	0.06	0.05
Black/White	0.43	0.30	0.62	0.40	0.36	0.30	0.31	0.28	0.06	0.04
Hispanic/White	0.53	0.29	0.77	0.36	0.22	0.14	-0.01	-0.07	0.05	0.03
SES: 1st quintile/5th quintile	1.26	0.85	1.34	0.85	0.63	0.63	0.26	0.17	0.30	0.31
SES: 1st quintile/3rd quintile	0.59	0.45	0.72	0.46	0.36	0.35	0.14	0.08	0.21	0.23
SES: 3rd quintile/5th quintile	0.67	0.47	0.62	0.40	0.27	0.27	0.12	0.13	0.09	0.08

Note: In this table, all positive numbers represent gaps in reference to the advantaged group indicated on the right hand side of the first column . Negative numbers indicate that the lefthand group has better scores, on average. [a]For both externalizing and internalizing behaviors, a positive gap indicates better behavior (i.e., less externalizing and internalizing) for the advantaged group.



Notes: The graph shows standard deviation differences in skills and behavior for children in the lowest income quintile and the highest income quintile based on estimates in Table 1.

Figure 2: Trajectories of Reading and Math Achievement by Parental Income, ECLSK-98 Data

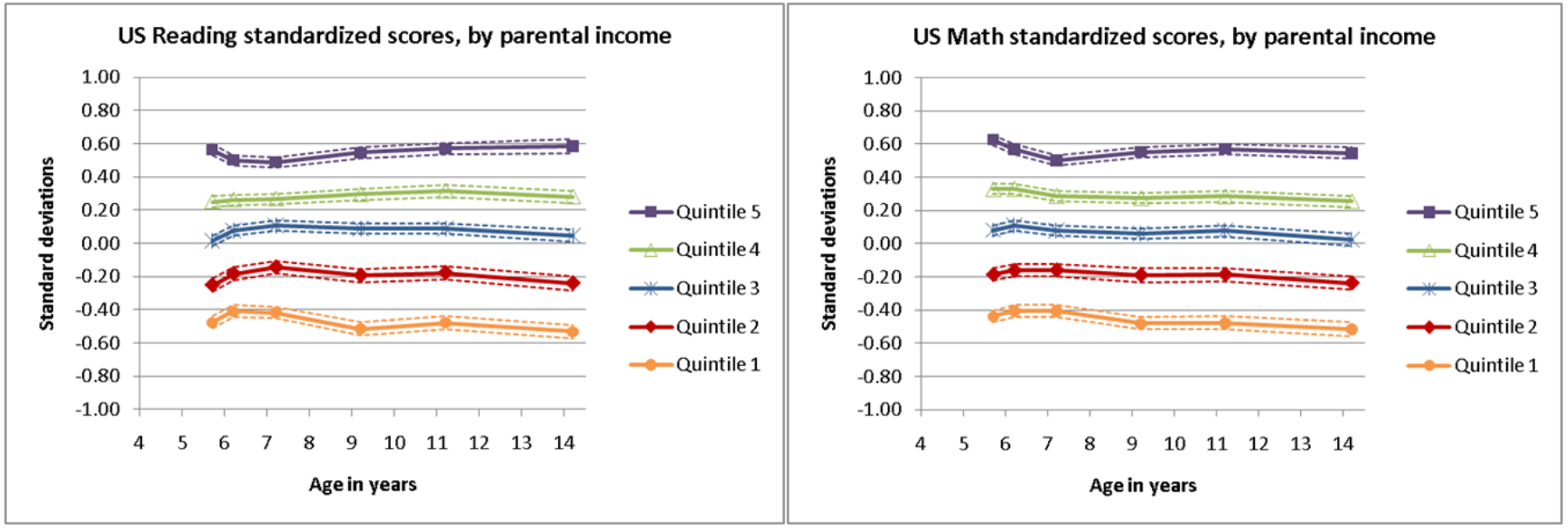


Figure 3: Trajectories of Children's Externalizing and Internalizing Behavior by Parental Income, ECLSK-98 Data

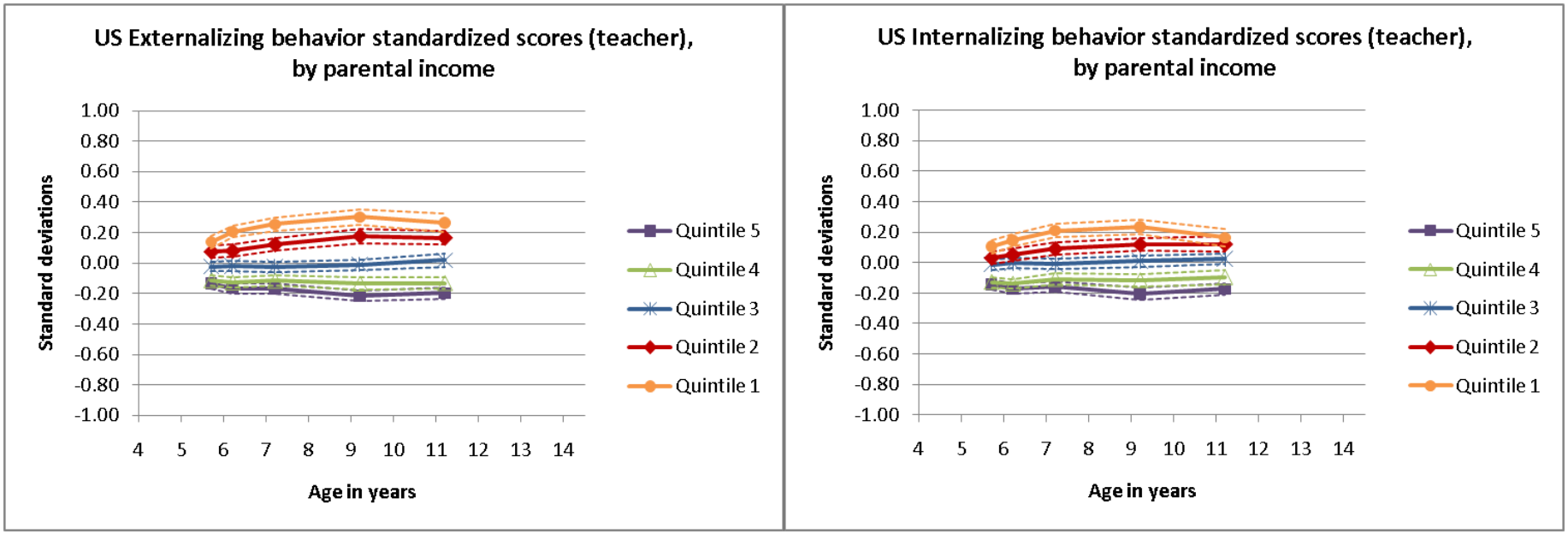


Figure 4: Average cognitive impact at end of treatment

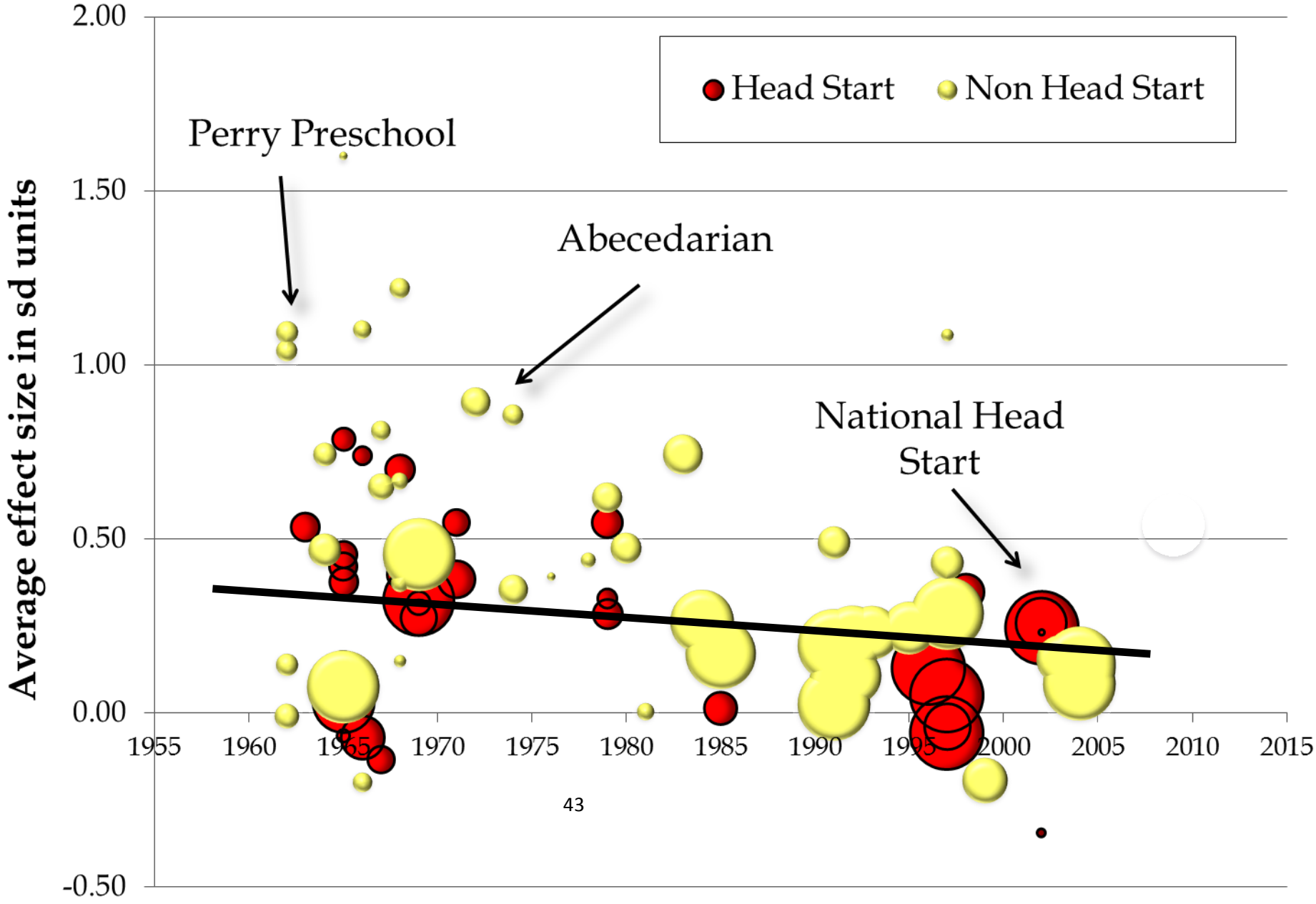
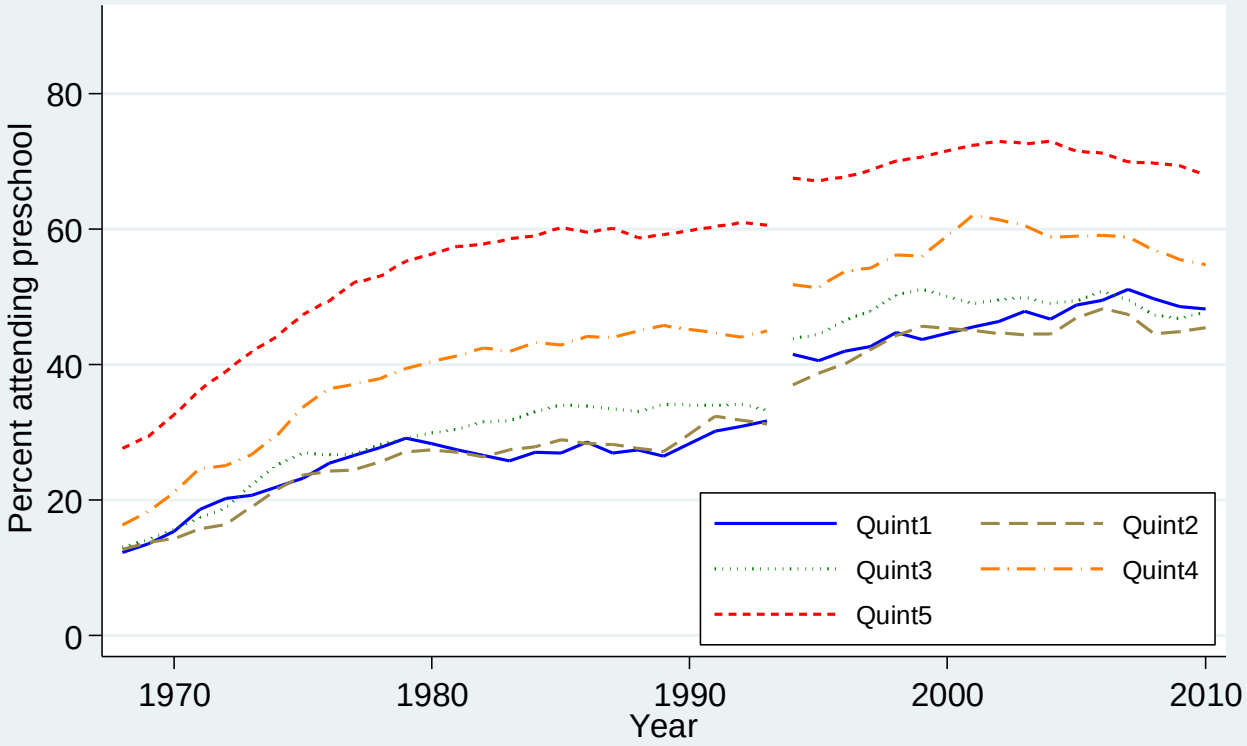


Figure 5: Percent of children enrolled in preschool by family income quintile:

3- and 4-year olds



Note: Data from October CPS, data shown are from 3 year moving averages

[1] We define preschoolers as children age 3 to 5 who are not yet attending elementary school, and distinguish them from infants and toddlers.

[2] The Early Childhood Longitudinal Birth Cohort data are also similar, and yield substantively similar results to those found in the ECLS Kindergarten studies; see Isaacs & Magnuson, 2011).

[3] Programs selected for our analysis had both treatment and control/comparison groups, included at least 10 participants in each condition, incurred less than 50% attrition, and measured children's cognitive development close to end of their "treatment" programs. Studies had to have used random assignment or one of the following quasi-experimental designs: change models, fixed effects modes, regression discontinuity, difference in differences, propensity score matching, interrupted time series, instrumental variables, and some other types of matching. Studies that used quasi-experimental designs must have had pre- and post-test information on the outcome or established baseline equivalence of groups on demographic characteristics determined by a joint-test. See also http://developingchild.harvard.edu/activities/forum/meta_analytic_database/.

[4] These price points are meant to reflect the real costs of providing higher quality programs, but note that even $10,000 is less than the average per pupil cost for K-12 schooling, in which class sizes can be significantly larger than for the early years. This is also less than the current cost of programs such as Perry Preschool. It is a midpoint between the costs of the recently studied Tulsa Pre-K program ($4,403 for part-day and $8,803 for full-day) and the Boston preschool program ($12,000 for full-day).

[5] It is hard to know whether the association would be higher or lower if the program affected other domains, such as behavior, as we would expect these to be correlated with achievement in both contexts, but whether the correlations would differ across experimental vs. non-experimental settings is unclear. A bigger issue is the question of whether these associations would hold when moving from a small-scale intervention to a large-scale intervention, in which rather than a few students shifting positions on test score percentile rank, the entire distribution shifted upward.