Early Childhood Development in Latin America and the Caribbean

Norbert Schady

September 9, 2005

There is considerable evidence that young children in many developing countries suffer from profound deficits in nutrition, health, fine and gross motor skills, cognitive development, and socio-emotional development. Early Childhood Development (ECD) outcomes are important markers of the welfare of children in their own right. In addition, the deleterious effects of poor ECD outcomes can be long-lasting, affecting school attainment, employment, wages, criminality, and measures of social integration of adults. We discuss ECD outcomes in Latin America and the Caribbean. The paper considers the theoretical case to be made for investments in ECD, selectively reviews the literature on the impact of ECD programs in the United States, discusses the evidence from Latin America and the Caribbean, and makes suggestions for future research. The focus is on the relation between ECD and measures of household socioeconomic status, child health, and parenting practices, as well as on the impact of specific policies and programs. We conclude that the knowledge base on ECD is still thin in Latin America and the Caribbean. There are therefore very high returns to comparative descriptive analysis of ECD outcomes in the region, as well to careful evaluations of the impact of various programs.

JEL Codes: I18, I28, J13, 015

There is considerable evidence that young children in many developing countries suffer from profound deficits in nutrition, health, fine and gross motor skills, cognitive development, and socio-emotional development. Early Childhood Development (ECD) outcomes are important markers of the "welfare" of children in their own right. In addition, the deleterious effects of poor ECD outcomes can be long-lasting. Poor ECD outcomes are associated with inadequate school readiness, and can condemn children to poor school performance. School-based interventions may therefore be less effective at promoting adequate learning than interventions that prevent deficits at earlier ages. Research from a number of developed countries also suggests that low levels of cognitive development in childhood, as measured by tests administered as early as 22 months of age, are important predictors of wages (for example Connolly, Micklewright and Nickell 1992; Currie and Thomas 1999; Feinstein 2003; Robertson and Symons 2003). Poor ECD outcomes therefore contribute to the intergenerational transmission of poverty in a variety of ways. For this reason, policies that effectively improve ECD outcomes can have long-lasting effects on welfare, as well as on a country's economic performance.

In this paper, we discuss ECD outcomes in Latin America and the Caribbean. The first part of the paper considers the theoretical case to be made for investments in ECD. We also selectively review the literature on the impact of ECD programs in the United States. In the second part of the paper, we discuss the evidence from Latin America. This section focuses on a handful of recent papers that have described ECD deficits in the region, as well as the relation between ECD and measures of household socioeconomic status, child health, and parenting practices. Some of these papers also attempt to draw causal inferences about the impact of specific programs and policies. The third part of the paper discusses directions for future research. An important message of this section is that the knowledge base on ECD is still thin in Latin America. There are therefore very high returns to comparative descriptive analysis of ECD outcomes in the region, as well to careful evaluations of the impact of various programs and policies.

1. The case for investments in ECD

The economic case for public investments in ECD programs has been made strongly by a number of authors. In a recent (2003) paper entitled "Human Capital Policy", Pedro Carneiro and James Heckman argue for ECD investments on two grounds. First, all else being equal, returns to investments in early childhood will be higher than those to investments made later in life simply because beneficiaries have a longer time to reap the rewards from these investments—a point made earlier by Becker (1964). Second, investments in human capital have dynamic complementarities, so that "learning begets learning" (p.7). This second proposition—that some forms of learning may be easiest in early childhood and make later learning easier—is consistent with evidence from the biological literature that suggests that some mental faculties are more malleable in early childhood than later in life. Carneiro and Heckman also argue that, at current levels of investment, returns to investments in early childhood in the United States are high, those to investments in the old are low. As a result, there is room for Pareto-improving efficiency gains.

Janet Currie's (2001) review of ECD programs in the United States makes a number of complementary arguments for ECD investments. She argues that it may be more effective for a government concerned with equity to equalize initial endowments through ECD programs than to compensate for differences in outcomes later in life—both because ECD investments may be more cost-effective, and because they avoid many of the moral hazard problems inherent in programs that seek to equalize outcomes in adulthood. Similar arguments are made in the World Bank's 2006 World Development Report, Equity and Development (World Bank 2005). Currie also argues that there may be a variety of market failures, including liquidity constraints, information failures, and externalities, all of which lead to under-investment in early childhood.

The theoretical case for ECD interventions is bolstered by empirical evidence that there can be high returns to high-quality investments early in a child's life. We begin with a discussion of the impact of preschool programs, drawing heavily on Currie (2001). Currie reviews the findings from a number of randomized evaluations of pilot ECD interventions, including the Perry Preschool Project, Carolina Abecederian Project, and the Early Training Program.

The Perry Preschool Program is the most-studied ECD intervention in the United States. Between 1962 and 1967, a sample of 128 low-income African-American children ages 3 or 4 who were assessed to be at high risk of school failure were randomly assigned into treatment and control groups. The treatment group received a half-day preschool every weekday plus a weekly home visit—both for eight months of the year, for two years. Project staff collected data on both study groups from ages 3 to 11, and again at ages 14, 15, 19, 27 and 40. Analysis of these data showed that the treatment group outperformed the control group on a variety of measures of educational attainment, including lower grade repetition, higher rates of high school graduation, and higher performance on various intellectual and language tests up to age 7, school achievement tests at ages 9, 10, and 14, and literacy tests at ages 19 and 27. At age 40, those who received the ECD intervention had median earnings that were more than one-third higher than those who did not, were significantly more likely to be employed, had lower fractions of lifetime arrests, and were sentenced to significantly fewer months in prison (Schweinhart 2005; see also Currie 2001; Carneiro and Heckman 2003).

The Carolina Abecederian Project provided a particularly intensive ECD intervention: At birth, children were randomized into a treatment group that received "enriched center-based child care services emphasizing language development for eight hours per day, five days a week, 50 weeks per year, from birth to age five" and a control group (Currie 2001). At school entry, the study children were again randomized into two groups, one of which received no further intervention, and another which received a "Home-School Resource Teacher". At age 15, the children who had received the preschool intervention had higher scores on achievement tests, and reductions in the incidence of grade retention and special education. (Children who are placed in a special education "track" are generally more likely to drop out of school in the future.) The effects of the Home-School Resource Teacher were either small or insignificant. At age 21, the children exposed to the ECD intervention had higher average test scores, and were twice as likely to be still in school or to have ever attended a four-year college. Finally, the Early Training Program consisted of weekly home visits to 4 and 5 year-olds, as well as part-day preschool for two, and in some cases three, summers. At age 12, children randomly assigned to

participate in the program were significantly less likely to require special education than controls.

The evaluations of the Early Training Program, the Carolina Abecederian Project, and the Perry Preschool Project provide "laboratory" evidence of the possible returns to ECD investments. However, these "model" ECD interventions are typically funded at higher levels and are administered by staff who are more motivated and better trained than staff at most publicly-funded programs. Moreover, as in many clinical trials, the sample sizes in these evaluations are generally very small—less than 100 children in the treatment and control groups each. We therefore next turn to evaluations of the impact of Head Start, the largest program for disadvantaged preschool children in the United States. Here, too, we draw heavily on the review paper by Currie (2001), as well as on work by Currie and her co-authors (Currie and Thomas 2000; Garces, Thomas, and Currie 2002; Blau and Currie 2004).

Head Start, created in 1965 as part of the Johnson administration's "War on Poverty", aims to improve the skills of disadvantaged preschool children by providing them with (predominantly part-day) preschool programs. In 1999, the program covered 800,000 children—almost 50 percent of eligible three and four year-old children—and received $4.7 billion of federal funding (Currie 2001). There has never been a large-scale, randomized evaluation of Head Start, so evaluations have relied on a variety of techniques to construct comparison groups—for example, comparing children who attended Head Start with others who did not (Lee et al. 1990), comparing siblings who attended Head Start with those who did not (Currie and Thomas 1995), or comparing children in schools which offered Head Start for two years with those which offered it for three years (Reynolds 1998). These are serious attempts to correct for the potential endogeneity of participation in Head Start. However, the results may still be biased if self-selection into Head Start on the basis of unobservables is important—for example, if parents are more likely to enroll children who have learning difficulties or, alternatively, enroll promising children they expect would benefit most from the program (see Todd and Wolpin 2003 for a general discussion of these issues).

With this caveat in mind, the studies that are methodologically most sound generally report significant effects of Head Start. Children who attended Head Start were less likely to be

enrolled in special education, more likely to make adequate grade progression, less likely to drop out of high school, and had significantly higher test scores. Longer-term studies also find higher rates of college attendance (for whites) and lower rates of delinquency and crimes (for blacks) among program participants (Garces, Thomas, and Currie 2002).

One particularly interesting line of inquiry considers the extent to which the estimated Head Start effects vary by race. Currie and Thomas (1995) find that the initial program impacts on vocabulary and reading test scores are similar for whites and blacks, but that gains for Head Start participants quickly "fade out" in primary school for black children but not for white children. There are various possible explanations for such differences in the impact of the program by race. One explanation is that there are differences in the impact of Head Start itself—the program may not serve black children as well as white children. Another explanation is that there are differences in the experience of children after they leave Head Start: Black children who attended Head Start may have family, neighborhood, or school environments that are less conducive to learning than other black children, whereas white children who attended Head Start are not disadvantaged relative to other white children. Currie and Thomas (2000) explore this question and conclude that the fade out of gains for black children is a result of the lower-quality schools they attend after Head Start. In a similar vein, Fryer and Levitt (2004) argue that the gap in test scores between black and white children becomes larger after kindergarten because of differences in the quality of schooling. Finally, Currie and Thomas (1999) find that gains in test scores for Head Start participants are at least as large for Hispanic children as for non-Hispanic whites. Moreover, the effects tended to be larger among children whose mothers had been interviewed in Spanish, a result they attribute to the importance of exposure to the "mainstream" language.

A number of studies have analyzed the effect of Head Start on child socio-emotional development. This is important because child attributes as basic as being able to sit still and pay attention are necessary for any reasonable amount of learning to take place in school (a point made by Blau and Currie 2004, among others). Children who are appropriately socialized tend to turn into better-adjusted adults, and the labor market returns to various non-cognitive skills, including motivation, enthusiasm, cooperation, and teamwork may be as large or larger as those

to "IQ" and other dimensions of cognitive development (a point made by Carneiro and Heckman 2003, among others). Behavioral problems in early childhood have also been shown to be strong predictors of high school dropout and delinquency.

There is some controversy about the effects of center-based ECD programs like Head Start on child socio-emotional development. At what age, and under what circumstances, is it beneficial for young children to spend large amounts of time away from their mothers? The analysis of model programs like the Perry Preschool Project and the Carolina Abecederian Project suggest that, whatever may be the negative effects of early childhood programs on mother-child attachment and (possibly) on child socio-emotional development, these are offset by the positive effects of program participation—at least in terms of higher academic achievement, higher wages, and lower criminality. That being said, it is not clear how easy it is to extrapolate from such model programs to large, publicly-funded or implemented programs like Head Start. Two recent papers are here informative. A randomized evaluation of the Early Head Start program, which combines center-based preschool with parent-child group socialization activities and parent education, either center-based or during home visits, suggests that children participating in Early Head Start exhibit less aggressive behavior, less negative behavior towards parents, and are better able to pay sustained attention during play (Blau and Currie 2004). This is consistent with earlier evidence that Head Start had positive effects on measures of social adjustment of children, including impulse control (Lee et al. 1990, cited in Currie 2001). On the other hand, research by Magnuson, Ruhm, and Waldfogel (2004) suggest that children in pre-kindergarten programs, as well as children participating in Head Start, were *more* likely to exhibit aggressive behavior and less able to exercise self-control. It is not clear what explains these differing findings. Part of the explanation may be related to the very large differences in quality across different child care options and Head Start sites: By some estimates, as many as 40 percent of children in the United States are attending child care judged to be of low quality (Danziger and Waldfogel 2000; see also the discussion in Currie 2001). Clearly, more research on this issue is warranted.

Head Start and similar center-based programs seek to improve school readiness by providing a nurturing learning environment in the pre-school years. These programs focus on

the child, rather than on parents. Yet there is widespread consensus that parenting also "matters" for ECD outcomes. We now briefly discuss the evidence on programs that attempt to affect parenting behavior in the United States, drawing heavily on review papers by Brooks-Gunn and Markham (2005), Brooks-Gunn, Berlin, and Fuligni (2000), and Gomby, Culross and Behrman (1999). The programs we discuss include "center plus" preschool programs like Early Head Start; home-visiting programs; parent behavior training programs; and family literacy programs.

Center plus preschool programs have been shown to have positive effects on nurturance and, in some cases, reductions in spanking, increases in the use of reasoning as a disciplining device, and improvements in parents' abilities to assist in child problem-solving activities (Brooks-Gunn and Markham 2005). As with the "regular" preschool programs, center plus programs also have positive effects on child cognitive development and school readiness. Disentangling the causality is hard: What fraction of the improvements in child outcomes are a result of the observed changes in parenting practices rather than of the participation in the preschool component of the program? These difficulties notwithstanding, a number of analyses report that as much as half of the observed changes in child outcomes can be attributed to improvements in parenting (Brooks-Gunn and Markham).

The evidence on the effectiveness of home visiting programs is decidedly more mixed. Home visiting programs send trained staff into homes of families with young children to encourage changes in parenting practices. The exact nature of the intervention varies by program, but programs typically attempt to provide parents with social support, practical assistance, and education about parenting and child development, and to prevent child neglect and abuse.

Gomby, Culross and Behrman (1999) provide evidence on the impact of six programs which jointly covered as many as 550,000 children in the United States. All six programs included a randomized evaluation. Most of the evaluations provided some evidence of improved parenting and home environment outcomes, although (worryingly) differences between treatment and control groups were more often found for parent-reported measures than for those assessed by enumerators. None of the programs provided consistent evidence of improvements on a variety of tests of child development and achievement (Gomby, Culross and Behrman; Brooks-

Gunn and Markham 2005). Results on changes in measures of children's behavior were mixed. In one evaluation, no program effects were found on parent-reported child behavior for participants in the Comprehensive Child Development Program. In another evaluation, children who had participated in the Elvira Nurse Home Visitation Program were re-visited when they were 15 years old, 13 years after the end of the intervention. The evaluation found no differences between treatment and control groups in measures such as acting out in school, suspensions, initiation of sexual intercourse, or major acts of delinquency, although children who had been exposed to the program reported fewer instances of running away, fewer arrests and convictions, fewer cigarettes smoked per day, fewer days having consumed alcohol in the last six months, and less lifetime promiscuity. Gomby, Culross and Behrman also provide some evidence of lower rates of child abuse, neglect and other forms of child maltreatment among participants in various home visiting programs.

By and large, the results from the evaluations of home visiting programs in the United States are disappointing. Part of the problem appears to be that home visiting programs suffer seriously from attrition—even among those parents who are willing to participate in home visiting programs in the first place, a large number (often more than half) drop out of the program, and those who stay generally receive many fewer visits than was originally anticipated. Gomby, Culross and Behrman (1999) make a compelling case that this is, at least in part, a result of program goals and design—often, programs are seeking to convince parents to change behaviors that they themselves may not view as negative. A further problem with home visiting programs in the United States is that they have suffered from a great deal of staff turnover. Such staff turnover is particularly debilitating in an intervention which relies on trust between home visitors and families.

A variety of programs in the United States seek to alter parenting behavior for children who have exhibited problem behaviors. One parent training program revolved around group discussions based on videotaped vignettes of typical discipline situations in the home, and how best to handle them. This program, and others discussed in Brooks-Gunn and Markham (2005), was found to increase positive parenting behaviors among participants. Programs that involved both parents and preschool teachers appeared to be most successful. Finally, literacy programs

have attempted to increase the quantity and quality of reading by parents to their children. These programs would seem to be particularly relevant for Latin America and the Caribbean, where levels of literacy of parents are much lower than those found in the United States.

2. Evidence on ECD outcomes in Latin America and the Caribbean

We now turn to a discussion of the evidence of the impact of ECD programs in Latin America and the Caribbean. Although there is a wealth of data in the medical, sociological and economic literature on the health and nutritional status of infants and young children in Latin America, relatively little is known about other dimensions of their welfare. In this section, we discuss some recent research which focuses on other markers of early development, including motor skills, cognitive development, and socio-emotional development. We draw heavily on work by Paul Gertler, Lia Fernald and various co-authors on Mexico, Christina Paxson and Norbert Schady on Ecuador, Jere Behrman, Yingmei Cheng, and Petra Todd on Bolivia, Sally Grantham-McGreggor and various co-authors on Jamaica, and Samuel Berlinski and Sebastian Galiani on Argentina.

Mexico: We begin with a discussion of evidence of ECD shortfalls in Mexico, as presented in Fernald et al. (2005), Gertler and Fernald (2004), and Behrman, Parker and Todd (2004). Fernald et al. focus on the relationship between deficits in child weight and height and the Mental Development Index (MDI) of the Bayley Scales of Infant Development for infants aged 12.5 to 23.5 months. The mental scale of the Bayley contains 178 items and is designed to assess memory, learning, problem-solving, sensory-perceptual acuities, and receptive and expressive language development. The test is normed at 100, with a standard deviation of 15, at each age. Scores below 85 suggest mildly delayed performance, scores blow 70 suggest significantly delayed performance.

Fernald et al. (2005) find significant reductions in the Bayley MDI scales with the age of the child: At age 13-14 months, the mean score was 97.4, 14.4 percent of children were one standard deviation or more below 100, and 3.0 percent were 2 standard deviations or more below. By age 21-23 months, the mean score was 87.0, almost half of children were 1 standard deviation below 100, and 11.3 were more than 2 standard deviations below. These deficits in

mental development are mirrored by an increasing fraction of children with low height for age—at age 13-14 months, 25.9 percent of the children were stunted, by age 21-23 months this fraction had risen to 42.7. (There is no comparable trend in weight for height.) However, Fernald et al. find no association between height for age and the Bayley score once family and environmental variables are included in a multivariate regression framework. More surprisingly, *none* of the parental or socioeconomic factors they control for, including income, employment, parental age, education, and whether the head of the household spoke an indigenous language, are significant predictors of the Bayley score.

Gertler and Fernald (2004) analyze the impact of *Oportunidades*, the conditional cash transfer program that built on PROGRESA, on a large set of ECD outcomes. *Oportunidades* makes large cash transfers, as much as 20 to 30 percent of household income, conditional on households meeting a set of health care requirements. Specifically, in order to receive the cash, households must make regular, monitored visits to health centers. During these visits, children are immunized, attend growth monitoring, and receive micronutrient supplementation, while parents receive education on health, nutrition, and hygiene. (Separately, households with school-aged children are required to enroll their children in school and have them attend regularly in order to receive transfers.)

To measure cognitive development, Gertler and Fernald (2004) use the Woodcock-Johnson-Munoz III tests, the MacArthur Communicative Development Inventories (CDIs, or *Inventario del Desarrollo de Habilidades Comunicativas: Palabras y Enunciados* in Spanish), and the Spanish version of the Peabody Picture Vocabulary Recognition Test (PPVT), the *Test de Vocabulario en Imágenes Peabody* (TVIP). The Woodcock-Johnson-Munoz is a co-normed set of tests that measure general intellectual ability, specific cognitive abilities, and scholastic aptitude. The tests have been used in Latin American contexts to evaluate early childhood nutritional interventions and early health insults on cognitive development of infants and older children (Lozoff et al. 1991; Rodriguez and Prewitt-Diaz 1990; Roselli et al. 2001). The TVIP is a test of receptive language that is frequently used to evaluate Spanish-speaking pre-school children (Munoz et al. 1989; Umbel et al. 1992). Children are shown a series of slides, each of which contains four pictures, and asked to point to or otherwise identify a picture that

corresponds to a word stated by the interviewer. Later slides are gradually more difficult, and the test stops when the child has made six mistakes in the last eight slides. The MacArthur CDIs are parent report forms for assessing language and communication skills in infants and young children which have been shown to provide valid assessments of early language milestones in young Spanish-speaking children, and have been linked with important biological outcomes (Marchman and Martinez-Sussmann 2002; O'Connor et al. 2001).

To measure motor development, Gertler and Fernald (2004) use the McCarthy Scales of Children's Abilities, a comprehensive battery of tests that tests children in such activities as their ability to walk backwards, stand on one foot, and skip. Finally, to assess socio-emotional and behavioral development, the authors use the Achenbach Child Behavior Checklist, which the parent or guardian uses to rate a child's problem behaviors and competencies. This test, which has been used in low-income Spanish-speaking populations (Gupta et al. 2001; Jutte et al. 2003; Basuk et al. 1997), measures behavioral problems and social competencies as reported by parents. It includes questions related to aggression, hyperactivity, bullying, conduct problems, and defiance at home and at school.

Gertler and Fernald (2004) begin with comparisons of cognitive development outcomes between the *Oportunidades* evaluation sample and the population that was used to norm a given test. They show that, on the basis of these comparisons, children in the evaluation sample appear to have very serious cognitive deficits: They place in the 17^{th} percentile for vocabulary on the TVIP, and in the 15^{th} percentile for long-term memory, the 22^{nd} percentile for short-term memory, and the 7^{th} percentile for visual integration in the Woodcock-Johnson-Munoz. The results they present are not disaggregated by the age of the child, so it is not clear whether the pattern observed by Fernald et al. (2005) using the Bayley scores—larger deficits for older children—is also apparent in this sample and with these outcomes.

To analyze the impact of *Oportunidades* on ECD outcomes, Gertler and Fernald (2004) focus on two comparisons. First, they compare children who received *Oportunidades* transfers with a comparison group of children who did not receive transfers. This comparison group was constructed with matching techniques: A group of communities that was not eligible for *Oportunidades* was selected for the evaluation, and data on a variety of outcomes was collected.

However, because these communities were, on average, somewhat better off than the communities eligible for *Oportunidades*, Gertler and Fernald further adjust for differences between households in the treatment and comparison groups with matching and regression techniques. Second, they compare households in communities who received *Oportunidades* transfers for different amounts of time. This control group was constructed by randomization—specifically, as part of the evaluation design, a lottery was used to assign communities into two groups, one of which received transfers for 12 to 18 months longer than the other.

Using the treatment and matched comparison groups, Gertler and Fernald (2004) find significant differences in motor skills: On average, outcomes are 15 percent higher among boys and 10 percent higher among girls in the treated communities than among similar children in the comparison communities. Children in the treated communities also appear to have fewer socio-emotional problems, although the effect is only statistically significant for girls. There is no clear pattern of program effects on any of the measures of cognitive development. Finally, the authors show that there is no evidence that the duration of program exposure, as measured by the difference between the two random assignment groups, has a significant impact on any outcome—motor skills, socio-emotional problems, or cognitive development.

The analysis by Behrman, Parker and Todd (2004) complements the results in Gertler and Fernald (2004). Behrman, Parker, and Todd focus on schooling outcomes. Like Gertler and Fernald, they make two comparisons—first, between households that received *Oportunidades* transfers and the matched set of communities that never received transfers, and second, between communities that were randomly assigned to different amounts of program exposure. Behrman, Parker and Todd show that children who were exposed to *Oportunidades* between ages 0 and 6 were likely to subsequently enter school at a slightly earlier age, were more likely to progress on time, and more likely to have higher years of completed schooling as they begin to enter school. As in Gertler and Fernald, program impacts are generally apparent in the comparisons between the treatment and matched comparison groups, and not in the comparisons between the two treatment groups that were randomly assigned to different amounts of program exposure.

Ecuador: Paxson and Schady (2005) use data on a sample of poor children in Ecuador to study the determinants of child cognitive development, as measured by performance on the

TVIP. They begin their analysis with a series of descriptive graphs which show that the mean age-normed TVIP score declines from close to 95 for children aged 36 months, to less than 85 for children aged between 54 and 60 months, and flattens out thereafter. This decline in the mean is accompanied by an increase in dispersion: The standard deviation of the TVIP score rises from 8 for the youngest children, to nearly 25 for the oldest children. Another way to view the increasing dispersion of the test scores is to examine percentiles. Paxson and Schady graph the 90^{th}, 50^{th}, 25^{th} and 10^{th} percentiles of the TVIP scores at each age. They show that the 90^{th} percentile scores are relatively constant with age, and there are modest declines in the median; by contrast, scores for children at the 25^{th} and 10^{th} percentiles of the distribution decline sharply with age.

What factors are responsible for the increasing dispersion in test scores with age? As a first step to answering this question, Paxson and Schady (2005) graph medians of the TVIP score at each age after splitting the sample by wealth, education, and place of residence. These graphs show that, for both wealth and parental education, there is little dispersion of the TVIP score at young ages. Children in households in the top wealth quartile, as well as children whose parents have completed secondary school or more, have approximately similar scores at older ages as at younger ages—indeed, if anything, the age-standardized scores for these groups seem to be somewhat higher among the older children. By contrast, children in the lowest wealth quartile, as well as children whose parents have only incomplete primary school or less, show very sharp drops in their age-standardized TVIP scores: On average, the score for the oldest children, age 72 months, is approximately 30 points lower than the corresponding score for children age 36 months. The decline in the median TVIP score with age affects children in both urban and rural areas, although (surprisingly) rural children appear to have higher scores at older ages.

Paxson and Schady (2005) next turn to a regression framework to analyze the determinants of cognitive development. They conclude that, as in the descriptive graphs, children from wealthier households, and children whose parents are more educated, have significantly higher test scores. These effects are large: For example, a child whose family falls at the 90^{th} percentile for wealth, maternal education and paternal education is predicted to have a score that is 35 points higher than a child at the 10^{th} percentile for each of these variables—a

difference of roughly two standard deviations of the test score. Moreover, the "protective" effect of socioeconomic status on child cognitive development becomes larger with age, so that children from households with lower socioeconomic status fall ever-further behind their better-off counterparts as they become older. As Paxson and Schady discuss, one plausible interpretation for these findings is that the effect of socioeconomic status on cognitive development is cumulative. Paxson and Schady also consider the relationship between cognitive outcomes, socio-economic status, and measures of child nutritional status. They show that nutritional status, in particular hemoglobin levels, is significantly associated with test performance. However, measures of nutritional status account for only a small fraction of the association between wealth, parental education and TVIP scores.

An innovative feature of the data collected in Ecuador is that it includes information on parenting quality, including the Home Observation for Measurement of the Environment (HOME) scale, which has been widely used in research in the United States (Bradley 1993; Caldwell and Bradley 1984; Bradley et al. 2001). This is constructed from 11 items that are assessed by enumerators at the close of the interview, and measures punitiveness (for example, whether parents yelled at or hit children during the interview), and responsivity to children (for example, whether they responded to and encouraged children in a positive way during the interview). Each item is scored as a dichotomous variable, and the final scale ranges from 0 to 11, with higher values corresponding to less responsive and harsher behavior. Paxson and Schady (2005) show that children whose parents have lower (better) HOME scores, reflecting warmer and less punitive behavior, have significantly higher TVIP test scores. These effects are large and, interestingly, the parenting measures account for a substantial fraction of the associations between socioeconomic status and cognitive development that are observed in the data.

Bolivia: Behrman, Cheng, and Todd (2004) use non-experimental data to evaluate the impact of a Bolivian pre-school program, the *Proyecto Integral de Desarrollo Infantil*, or PIDI. The outcome measures include "a battery of tests of bulk motor skills, fine motor skills, language and auditory skills, and psychosocial skills". (The authors do not provide details on the exact nature of these tests.) The PIDI program provides daycare, nutritional and educational services

to children between the ages of 6 and 72 months. Children enrolled in the program attend full-time child care centers located in the homes of women living in low-income areas targeted by the program.

Behrman, Cheng, and Todd (2004) use propensity score matching to estimate program impacts. Specifically, they provide results based on two comparisons: First, they compare program participants with a matched comparison group of non-participants—they refer to these as average treatment impacts. Second, they compare participants with different amounts of exposure to the program—they refer to these as marginal treatment impacts. The identifying assumption for the estimates of average treatment impacts is that there is no selection into the program on the basis of unobservable characteristics of households or children. This restriction is loosened somewhat in the estimates of marginal treatment impacts although, conditional on program participation, unobsevables are still not allowed to determine the duration of exposure.

Behrman, Cheng, and Todd (2004) find that, in the comparison of treated and untreated children, there is some evidence of positive program impacts on motor skills (both bulk and fine), on psychosocial skills, and on language acquisition. These impacts seem to be concentrated among children ages 37 months and older—for the younger children, ages 6-36 months, the estimated program effects are generally insignificant, and are as likely to be positive as negative. When the results are disaggregated by the length of exposure, effects are most clearly observed among children who have been exposed to the PIDI program for more than a year. Finally, Behrman, Cheng, and Todd calculate cost-benefit ratios for the PIDI program—an important concern given that the cost of the program is approximately $43 per month, in a country in which per capita annual GDP is $800 in exchange-rate converted pesos, and $2540 in PPP terms. The PIDI program is estimated to have an impact on height, cognitive development, and schooling of participating children. Behrman, Cheng, and Todd heroically combine these estimated program impacts with data on wages from Bolivia and a number of other countries to argue that there are positive cost-benefit ratios to PIDI under a variety of plausible assumptions and discount rates.

Jamaica: A number of papers by Grantham-McGregor and various co-authors use data from Jamaica to analyze ECD deficits and the short- and medium-term impact of interventions.

In one study, a lottery was used to divide a sample of stunted children ages 9-24 months into four groups: The first study group received nutritional supplements, specifically 1 kg. of milk-based formula provided weekly; a second study group received early childhood stimulation, specifically weekly home visits by social workers who demonstrated play with home-made toys and discussed parenting issues with mothers; the third study group received both the supplements and the stimulation; the fourth group served as a control group. In addition, data was collected on a sample of non-stunted children. Children in all of the study groups were then followed over time. Results after two years suggested that both the stimulation and nutritional supplement interventions had positive impacts on child development as measured by the Griffiths Mental Development Scales. The largest effects were found in the group that had received both interventions—indeed, after two years these children had caught up with the matched group of non-stunted children (Grantham-McGregor et al. 1991).

Children in the original study were re-visited at ages 7-8, and again at 11-12 years. Disappointingly, these results showed that the study group that received nutritional supplements only did *not* have better outcomes than the control group on a variety of tests of cognitive development. By contrast, the group that had received stimulation only and that which had received both interventions performed better than the control group on 9 of the 11 tests that were applied. The difference in the scores between the study groups that had received stimulation and the control group were significant at the 5 percent level or better for three tests—the Wechsler Intelligence Scales, the Ravens Progressive Matrices, and the PPVT (Grantham-McGregor et al. 1997; Walker et al. 2000). In a related study, undernourished children age 9 to 30 months and their mothers were randomly assigned to a treatment group which received stimulation, including weekly home visits by community health aides, and a control group. After one year, children who had received the stimulation intervention had significantly better outcomes in three of four subscales from the Griffiths Mental Development Scales. Mothers in the intervention group also had better knowledge of childrearing and childbearing practices (Powell et al. 2004).

Many countries in Latin America spend large amounts of public resources on feeding programs. While there is evidence that some nutrition programs have had positive long-term impacts, especially when they target pregnant mothers and continue through the first three years

of a child's life (for example, Martorell 1999 on a well-known study in Guatemala), many government feeding programs do not meet these criteria. In Peru, for example, the "Glass of Milk" program is the largest social transfer; it reaches 44 percent of households with children ages 3 to 11 years, but appears to have no impact on nutritional outcomes (Stifel and Alderman 2003). In the Jamaican population studied by Grantham-McGreggor and her co-authors, the nutritional intervention that was evaluated did not seem to have significant benefits for recipients in the long run, while there were long-lasting effects of the stimulation intervention. As with many clinical studies, the evaluations are based on very small sample sizes (less than 150 children), and it is not clear how well the stimulation programs that were studied could be replicated on a large scale. Still, these papers suggest that there could be high returns to carefully evaluated pilot interventions of stimulation interventions by governments, NGOs or others in Latin America and the Caribbean.

Argentina: In a recent paper, Berlinski and Galiani (2005) analyze the impact of a large program to construct pre-school facilities in Argentina in the 1990s on pre-primary school attendance and maternal labor supply. The intensity of the program varied by region, and Berlinski and Galiani use a differences in differences framework to identify program impact. The estimation also includes cohort and region dummies to control for nationwide trends and level differences across regions. The identifying assumption is therefore that regions that received more preschool facilities would not have different growth rates in enrollment in the absence of the construction program. (Berlinski and Galiani provide some evidence that this may be a reasonable assumption.) Berlinski and Galiani conclude that the program had a large, positive impact on preschool enrollment—indeed, they cannot reject the null hypothesis that all new preprimary school slots were taken up by children who would otherwise not have been in school. The authors take this as strong evidence for a supply constraint on preschool enrollment in Argentina.

If high-quality preschool has important returns in later life, as is suggested by much of the literature on the United States, it is important to determine how well Latin America and the Caribbean is performing in this dimension. The Word Bank's Development Data Platform (DDP) database includes data on gross preschool enrollment. To benchmark the region, we use

data for 2000, the last year for which preschool enrollment data are available for a large number of countries. Specifically, we regress preschool enrollment on a quartic in per capita GDP and a dummy variable for countries in Latin America and the Caribbean: The coefficient on the Latin American dummy is 12.7, with a robust standard error of 6.6. This suggests that, on average, countries in Latin America and the Caribbean have gross preschool enrollment rates that are approximately 13 percentage points above what would be expected for their level of development. (Results are qualitatively similar but larger in magnitude when the GDP variable is restricted to lower-order polynomials.) It also makes sense to run a comparable regression with weights given by the population of a country. In this case, the coefficient on the dummy variable for Latin America and the Caribbean provides a measure of the probability that a child age 3-5 attends school given the income of the country in which (s)he lives. In this specification, the coefficient on the Latin American dummy is insignificantly different from zero—1.6, with a robust standard error of 7.5.

Taken as a whole, Latin America and the Caribbean does not appear to have a deficit in preschool enrollment. Of course, these data say nothing about the quality of the preschool education—an important concern. Also, there are individual countries in the region with preschool enrollments levels that are well below what would be expected for their income—Argentina, for one, has a gross preschool enrollment rate that is 19.8 points below the value that would be predicted for its income level, even after the preschool construction program studied by Berlinski and Galiani (2005). Even in those countries in Latin America and the Caribbean with relatively high preschool coverage there are likely to be population groups with very low preschool enrollment rates, and the returns to offering high-quality preschool to these children, who are likely to be disadvantaged in a variety of other ways, could be very high.

3. Directions for policy and future research

In the United States, cost-benefit calculations of ECD interventions generally find large, positive cost-benefit ratios, especially for the more intensive (and more expensive) interventions (Currie 2001; Carneiro and Heckman 2003). One calculation of returns to enriched ECD programs targeted to high-risk disadvantaged minority male youth in the United States suggests

that the costs would more than pay for themselves in reduced incarceration rates alone (Donohue and Siegelman 1998, cited in Carneiro and Heckman 2003). There is good reason to believe that the returns to high-quality ECD interventions in developing countries like those in Latin America and the Caribbean could be as large, if not larger as those found in the United States. However, the knowledge base on ECD in Latin America and the Caribbean is disappointingly thin. Careful analytical work is needed to establish the basic facts about ECD outcomes and deficits in the region, and to understand the causal pathways whereby a given characteristic of households, parents, or children determines ECD outcomes. Experimentation and careful evaluation is also needed to identify specific policies and programs that can improve ECD outcomes.

Recent research from a number of Latin American countries has applied tests of motor skills, cognitive development, and socio-emotional development that have been internationally normed. In theory, norming of the test instruments could have several advantages. Many of the tests have been shown to be correlated with various biological outcomes, as well as with "economic" measures such as school performance and wage outcomes in later life. For example, the TVIP vocabulary recognition test used by Gertler and Fernald (2004) and Paxson and Schady (2005) is the Spanish version of the PPVT, and performance on the PPVT at early ages has been shown to be a strong predictor of schooling and income in Great Britain and the United States. Also, because the tests are generally normed by comparing results with those of a "reference" population, the scores are, arguably, meaningful measures of a particular dimension of ECD in some absolute sense. That is, the score on a give test can be used to provide answers to a question like "are children in the sample at the level they 'should' be for their age?" as well as questions like "do children of higher socioeconomic status in the sample perform better than those of lower socioeconomic status?" (For the latter question, there would be no need for reference populations for norming.) Finally, a number of the tests are age-normed, so that meaningful comparisons of developmental shortfalls can also be made across children of different ages. This, in theory, could help identify "critical periods" at which developmental shortfalls occur.

The appeal of the standardized, age-normed tests of ECD status is similar to the appeal of using z-scores as measures of nutritional status, cut-offs for hemoglobin levels to establish

anemia, or the fraction of households living below a dollar a day as an "international" measure of poverty. As such, many of the tests may be an improvement on ad-hoc, country-specific tests. Like the measures of nutritional status and poverty, however, they depend crucially on the extent to which the norming has been done appropriately. This is a concern because the samples of children on whom the test was normed are often small, and may not provide a meaningful comparator to the population for which the test is used. For example, the TVIP was normed on 1,219 Mexican children and 1,488 Puerto Rican children (see http://www.agsnet.com/assessments/technical/tvip.asp for details). More research is needed to assess the extent to which the reference populations and age norms are appropriate. This is an area where testing and child development specialists could help make meaningful contributions.

With a better understanding of the various tests in hand, careful description of the basic facts about ECD outcomes in Latin America and the Caribbean is an indispensable input into policy design. What is the magnitude of the deficits (if any) in cognitive development, socio-emotional development, and motor development for population-based samples of young children in the region? How do these vary with household characteristics? A well established fact from the literature on health is that there is a "gradient" between socioeconomic status and health: Households of lower socioeconomic status, as measured by income, consumption, or education have higher levels of mortality and morbidity (for reviews for developing countries, see Behrman and Deolalikar 1988; Strauss and Thomas 1998). Similar findings are often reported in the ECD literature for the United States (see, for example Smith et al. 1997; Blau 1999; Guo and Harris 2000; Waldfogel et al. 2002; Auginbaugh and Gittleman 2003; Baum 2003; Ruhm 2004; Taylor et al. 2004; Brooks-Gunn and Markham 2005). Remarkably little about this is known for countries in Latin America and the Caribbean. Indeed, two of the papers reviewed in this article find apparently contradictory results on this count—albeit, with different tests of cognitive development: Using the Bayley Mental Development Index scales, Fernald et al. (2005) find no gradients by income, parental education, or a number of other measures of socioeconomic status in Mexico, while Paxson and Schady (2005) find sharp gradients by parental education and wealth on child performance on the TVIP in Ecuador, especially for older children. It is not clear how to interpret these differences, although it is tantalizing that the sample in Fernald et al.

consisted of very young children, ages 12.5 to 23.5 months, while the associations between socioeconomic status and cognitive development that Paxson and Schady report are significantly larger for older children in their sample. Socioeconomic gradients in ECD in Latin America and the Caribbean may be stronger as children age.

Careful analysis is needed to establish whether there are "critical periods" at which faltering occurs in a particular dimension of early childhood development. For example, in most Latin American populations, the incidence of stunting, defined as height for age more than two standard deviations below that of a reference population, increases dramatically from about the age of 6 months to 24 months, and stabilizes (but does not recover) thereafter (Shrimpton et al. 2001). Are there comparable patterns in other dimensions of child wellbeing? Do specific health insults, inadequate resources, or low levels of stimulation have especially large negative effects on ECD outcomes at certain ages in the life of a child? These are hard questions to answer with a single cross-section of data, both because it is not possible to disentangle age and cohort effects, and because a variety of child outcomes tend to be correlated over time. There are therefore important benefits to the collection of panel data that span the life of children from birth onwards. Indeed, because in-utero conditions are likely to have an effect on subsequent child development, panels would ideally begin during a mother's pregnancy.

A clear understanding of what population groups are most vulnerable to ECD shortfalls in Latin America and the Caribbean is clearly indispensable for the design of effective programs. The literature from the United States is here informative. A number of papers suggest that there is considerable evidence of heterogeneity of treatment effects, and some evidence that the largest impacts are often found among populations with the biggest ECD deficits at baseline. In the Abecederian model program, all of the children were judged to be at risk of mental retardation, but the positive effects of the program were twice as large for children from the poorest and least educated families as for other children (Currie 2001). The randomized evaluations of Early Head Start and the Infant Health and Development Program both suggest that improvements in parenting behaviors among black mothers were larger than among white mothers (Love et al. 2002, cited in Brooks-Gunn and Markham 2005). Improvements in a variety of ECD outcomes were also larger for children of mothers with high school education or less than for those with

some college or more, and (less clearly) for mothers with low psychological resources, including initially higher incidences of maternal depression (Brooks-Gunn and Markham 2005). Currie (2001) and Currie and Thomas (1999) also argue that ECD improvements associated with Head Start were concentrated among participants who were most vulnerable. These results point to the importance of disaggregating analytical findings by race, gender, socio-economic status, or any one of a number of characteristics that could plausibly affect outcomes. It also suggests that serious thought needs to be given to the question of how best to target ECD interventions in Latin America and the Caribbean.

Another critical policy question is whether poor outcomes in early childhood can be made up later in life. The evidence in Paxson and Schady (2005) suggests that a substantial fraction of poor children in Ecuador arrive at the threshold of formal schooling with very large shortfalls in vocabulary recognition. Can these shortfalls, as well as other possible deficits in cognitive, motor, or socio-emotional development be reversed—for example, with access to high quality schooling? The evidence from the United States on this is inconclusive. The experimental results from the evaluation of the Abecederian Project suggest that one particular intervention during school-age, the Home-School Resource Teacher, did not enhance the effects of the ECD program. On the other hand, the results for Head Start reported by Currie and Thomas (1995) suggest that high quality schools are important to maintaining the advantages conferred by ECD investments. Little is known about this issue in Latin America. Once again, it is important to follow children from early childhood into their school years.

Descriptive work on ECD is a critical building block for more ambitious analytical work to understand the causal pathways whereby characteristics of households or children affect a given dimension of development in early childhood. In the literature on the United States, there is considerable controversy about *how* low incomes lead to poor ECD outcomes. Low income is often associated with a lack of resources that can affect child development—for example, toys or reading material that stimulate cognitive development, or high quality day care. Children in low income households also tend to have worse health and nutritional status, and generally have parents with lower levels of education. In addition, lower incomes are associated with higher levels of maternal depression and home environments which are less nurturing, both of which are

believed to have direct causal effects on ECD outcomes. Assuming that there is a gradient between socioeconomic status and ECD outcomes in Latin America and the Caribbean, careful descriptive work informed by economic theory about how (and why) households make choices about childhood investments, will be required to disentangle causal effects. Rich data sets with long histories and detailed information on household resources, maternal and child characteristics, parenting environments and access to social programs are particularly important in this regard.

In some cases, experimental evidence may also help recover "structural" parameters with a causal interpretation. The evaluation of PROGRESA is a good case in point. For the first years of the program, households were randomly assigned into a treatment group that received cash transfers and a control group. The program has been shown to have had positive effects on enrollment and attendance in school (Schultz 2004; Behrman, Sengupta and Todd 2002). If the increase in child enrollment and attendance translates into higher school attainment as adults, it should be possible to (eventually) collect data on the outcomes of children born to the (randomly-selected) PROGRESA treatment and control groups to identify the causal effect of parental education on ECD outcomes. There are likely to be other cases in which exogenous sources of variation, either from naturally-occurring "experiments" or from deliberate program design, can be put to good use.

Because the knowledge base in Latin America on what does and does not "work" to improve ECD outcomes is thin, there are very high returns to careful implementation of a variety of interventions, and rigorous evaluation of their impact. It may be that seemingly unusual combinations of programs have the potential to have the largest impact. For example, the research by Gertler and Fernald (2004) suggests that, on its own, a conditional cash transfer program like *Oportunidades* may not lead to significant improvements in child cognitive development. This might argue for an intervention that focuses on parenting skills. The descriptive evidence in Paxson and Schady (2005) for Ecuador and the small-scale experimental evidence in the work by Grantham-McGreggor and her co-authors in Jamaica both suggest that there could be very high returns to interventions that effectively improve parenting and the home environment. However, the high rate of attrition in many programs that seek to improve

parenting in the United States, in particular home visiting programs, argues for innovative combinations of parenting programs with interventions that households are keen on participating in—perhaps, conditional cash transfer or feeding programs.

Experimenting with different packages of interventions, both in small-scale "laboratory" settings and in public sector programs, perhaps implemented on a pilot basis, is indispensable to building up the knowledge base in Latin America and the Caribbean. Careful consideration needs to be given to the evaluation of these interventions. The literature on the United States shows that there can be large differences between estimates of the impact of ECD programs that are based on randomized and non-experimental methods, and that it is hard to sign the direction of the bias ex-ante (Currie 2001; Gomby, Culross and Behrman 1999). In Latin America and the Caribbean, the strongest evidence on the impact of conditional cash transfer programs is based on randomized evaluations. This has also helped ensure the political sustainability of these programs. A similar emphasis on innovative program design, careful implementation, and rigorous evaluation would build up the knowledge base on early childhood development in Latin America and the Caribbean. This would help identify programs and policies that ensure that children in the region can go on to have healthy and productive lives.

References

Aughinbaugh, A., and M. Gittleman. 2003. "Does Money Matter? A Comparison of the Effect of Income on Child Development in the United States and Great Britain." *Journal of Human Resources* 38(2): 416-40.

Bassuk, E. L., L.F. Weinreb, R. Dawson, J.N. Perloff, J.C. Buckner. 1997. "Determinants of Behavior in Homeless and Low-Income Housed Preschool Children." *Pediatrics* 100(1):92-100.

Baum, C. L. 2003. "Does Early Maternal Employment Harm Child Development? An Analysis of the Potential Benefits of Leave Taking." *Journal of Labor Economics* 21(2): 409-48.

Becker, G. 1964. *Human Capital*: *A Theoretical and Empirical Analysis with Special Reference to Education*. New York: Columbia University Press.

Behrman, J., and A. Deolalikar. 1988. "Health and Nutrition", in *Handbook of Development Economics*, Vol. 1, Eds.: H. Chenery and T.N. Srinivasan. Amsterdam: North Holland Press, pp. 631-711.

Behrman, J., Y. Cheng, and P. Todd. 2004. "Evaluating Pre-school Programs when Length of Exposure to the Program Varies: A Nonparametric Approach." *Review of Economics and Statistics* 86(1): 108-32.

Behrman, J.R., S.W. Parker and P. E. Todd. 2004. "Medium-Term Effects of the Oportunidades Program Package, Including Nutrition, on Education of Rural Children Age 0-8 in 1997." Unpublished manuscript.

Berlinski, S. and S. Galiani. 2005. "The Effect of a Large Expansion of Pre-Primary School Facilities on Preschool Attendance and Maternal Employment." Unpublished manuscript.

Blau, D and J. Currie. 2004. "Preschool, Day Care, and Afterschool Care: Who's Minding the Kids?" National Bureau of Economic Research Working Paper 10670. Cambridge, MA.

Blau, D.M. 1999. "The Effect of Income on Child Development." *Review of Economics and Statistics* 81(2): 261-276.

Bradley, R.H. 1993. "Children's Home Environments, Health, Behavior, and Intervention Efforts: A Review Using the HOME Inventory as a Marker Measure." *Genetic, Social, and General Psychology Monographs,* 119, 437-490.

Bradley, R.H., R.F. Corwyn, H.P. McAdoo, and C.G. Coll. 2001. "The Home Environments of Children in the United States Part I: Variations by Age, Ethnicity, and Poverty Status." *Child Development* 72(6): 1844-1867.

Brooks-Gunn, J. and L.B. Markman. 2005. "The Contribution of Parenting to Ethnic and Racial Gaps in School Readiness." *The Future of Children* 15(1): 139-68.

Brooks-Gunn, J., L.Berling and A. Fuligni. 2000. "Early Childhood Intervention Programs: What About the Family?" *Handbook of Early Childhood Intervention*, 2nd ed. J.P. Schonkoff and S. J. Meisels, eds. New York: Cambridge University Press.

Caldwell, B.M. and R.H. Bradley, 1984. *Administration Manual (Revised Edition): Home Observation for Measurement of the Home Environment.* Little Rock, AR: University of Arkansas.

Carneiro, P. and J. Heckman. 2003. "Human Capital Policy." National Bureau of Economic Research Working Paper 9495. Cambridge, MA.

Connolly, S., J.Micklewright, and S. Nickell. 1992. "The Occupational Success of Young Men who left School at Sixteen." *Oxford Economic Papers* 44: 460-479.

Currie,J. 2001. “Early Childhood Education Programs.” *The Journal of Economic Perspectives*, 15(2): 213-238.

Currie, J. and D. Thomas. 2000. “School Quality and the Longer-Term Effects of Head Start.” *The Journal of Human Resources* 35(4): 755-774.

Currie, J. and D. Thomas. 1995. “Does Head Start Make a Difference?” *American Economic Review* 85(3): 341-64.

Currie, J. and D. Thomas. 1999. “Early Test Scores, Socioeconomic Status and Future Outcomes.” National Bureau of Economic Research Working Paper 6943. Cambridge, MA.

Danziger, S. and J. Waldfogel. 2000. “Investing in Children: What do we Know? What Should We Do?” CASE paper 34. Centre for Analysis of Social Exclusion, London, UK.

Donohue, J. and P. Siegelman. 1998. “Allocating Resources Among Prisons and Social Programs in the Battle Against Crime.” *Journal of Legal Studies* 27(1): 1-43.

Feinstein, L. 2003. “Inequality in the Early Cognitive Development of British Children in the 1970 Cohort.” *Economica* 70: 73-97.

Fernald, L.C., L.M. Neufeld, L.R. Barton, et al. 2005 “Parallel Deficits in Linear Growth and Mental Development in Low-Income Mexican Infants in the Second Year of Life.” *Public Health Nutrition*, in press.

Fryer, R.G. and S.D. Levitt. 2004. “Understanding the Black-White Test Score Gap in the First Two Years of School.” *The Review of Economics and Statistics* 86(2): 447-464.

Garces, E., D. Thomas, J. Currie. 2002. “Longer-Term Effects of Head Start.” *The American Economic Review* 92(4): 999-1012.

Gertler, P.J. and L.C. Fernald. 2004. “The Medium Term Impact of *Oportunidades* on Child Development in Rural Areas.” Unpublished manuscript.

Gomby, D., P. Culross and R. Behrman. 1999. “Home Visiting: Recent Program Evaluations-Analyses and Recommendations.” *The Future of Children* 9(1): 4-26.

Grantham-McGregor, S.M., C.A. Powell, S.P. Walker and J.H. Himes. 1991. “Nutritional Supplementation, Psychosocial Stimulation, and Mental Development of Stunted Children: The Jamaican Study.” *Lancet* 338:1-5.

Grantham-McGregor, S.M., S.P. Walker, S.M. Chang and C.A. Powell. 1997. “Effects of Early Childhood Supplementation with and Without Stimulation on Later Development in Stunted Jamaican Children.” *American Journal of Clinical Nutrition* 66: 247-253.

Guo, G., and K.M. Harris. 2000. “The Mechanisms Mediating the Effects of Poverty on Children’s Intellectual Development.” *Demography* 37(4): 431-447.

Gupta, V.B., N.M. Nwosa, T.A. Nadel, S. Inamdar. 2001. “Externalizing Behaviors and Television Viewing in Children of Low-Income Minority Parents.” *Clinical Pediatrics* 40(6): 337-41.

Jutte, D.P., A. Burgos, F. Mendoza, C.B. Ford, L.C. Huffman. 2003. “Use of the Pediatric Symptom Checklist in a Low-Income, Mexican American Population.” *Archives of Pediatrics & Adolescent Medicine* 157(12):1169-76.

Lee, V.E., J. Brooks-Gunn, E. Schnur, F.R. Liaw. 1990. "Are Head Start Effects Sustained? A Longitudinal Follow-Up Comparison of Disadvantaged Children Attending Head Start, No Preschool, and Other Preschool Programs." *Child Development* 61(2):495-507.

Love, J., et al. 2002. "Making a Difference in the Lives of Infants and Toddlers and Their Families: The Impacts of Early head Start." U.S. Department of Health and Human Services.

Lozoff, B., E. Jimenez, A.W. Wolf. 1991. "Long-Term Developmental Outcome of Infants with Iron Deficiency." *New England Journal of Medicine* 5(325)(10): 687-94.

Magnuson, K.A., C.J. Ruhm, J. Waldfogel. 2004. "Does Prekindergarten Improve School Preparation and Performance?" National Bureau of Economic Research Working Paper 10452. Cambridge, MA.

Marchman, V. A., C. Martine-Sussmann. 2002. "Concurrent Validity of Caregiver/Parent Report Measures of Language for Children Who are Learning Both English and Spanish." *Journal of Speech, Language, and Hearing Research* 45(5):983-97.

Martorell, R. 1999. "The Nature of Child Malnutrition and its Long-Term Implications." *Food and Nutrition Bulletin* 20(3): 288-292.

Munoz, F., C. Quilodran, P. Velasquez, C. Niedmann, A Baeza, G. Silva and M Osorio. 1989. "Acquirement of the Spanish vocabulary among rural and urban students of the 9th region." *Revista Chilena de Pediatria* 60(6): 354-358.

O'Connor, D.L., R. Hall, D. Adamkin, N. Auestad, M. Castillo, W.E. Connor.2001. "Growth and Development in Preterm Infants Fed Long-Chain Polyunsaturated Fatty Acids: A Prospective, Randomized Controlled Trial." *Pediatrics* 108(2):359-71.

Paxson, C., N. Schady. 2005. "Cognitive Development among Young Children in Ecuador: The Roles of Wealth, Health and Parenting." Policy Research Working Paper No. 3605. World Bank, Washington, D.C.

Powell, C., H. Baker-Henningham, S. Walker, J. Gernay, and S. Grantham-McGregor**.** 2004. **"**Feasibility of Integrating Early Stimulation into Primary Care for Undernourished Jamaican Children: Cluster Randomised Controlled Trial." *British Medical Journal* 329(7457)89: 1-4.

Reynolds, A.J., J.A. Temple. 1998. "Extended Early Childhood Intervention and School Achievement: Age Thirteen Findings from the Chicago Longitudinal Study." *Child Development* 69(1):231-46.

Robertson, D., and J. Symons. 2003. "Do Peer Groups Matter? Peer Group versus Schooling Effects on Academic Attainment." *Economica* 70:31-53.

Rodríguez, V.L., J.O. Prewitt Diaz. 1990. "Correlations among GPA and Scores on the Spanish Version of WISC-R and the Woodcock-Johnson Achievement Subtests for 10- to 12-Year-Old Puerto Rican Children." *Psychological Reports* 66(2):563-6.

Rosselli, M, A. Ardila, J.R. Bateman, M. Guzman. 2001. "Neuropsychological Test Scores, Academic Performance, and Developmental Disorders in Spanish-Speaking Children." *Developmental Neuropsychology* 20(1):355-73.

Ruhm, C.J. 2004. “Parental Employment and Child Cognitive Development.” *Journal of Human Resources* 39(1):155-192.

Schultz , T.P. 2004. “School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program.” *Journal of Development Economics* 74(1): 199-250.

Schweinhart, L. J.2005. “The High/Scope Perry Preschool Study Through Age 40: Summary, Conclusions, and Frequently Asked Questions.” High/Scope Press.

Shrimpton, R., C. Victora, M. De Onis, R. Costa Lima, et al. 2001. “Worldwide Timing of Growth Faltering: Implications for Nutritional Interventions.” *Pediatrics* 107:75-81.

Smith, J.R., J. Brooks-Gunn and P.K. Klebanov, 1997. “Consequences of Living in Poverty for Young Children’s Cognitive and Verbal Ability and Early School Achievement.” In : Duncan, J. Brooks-Gunn J. (eds.) *Consequences of Growing up Poor*, Russell Sage Foundation, New York, 132-189.

Stifel, D., and H. Alderman. 2003. “The "Glass of Milk" Subsidy Program and Malnutrition in Peru.” Policy Research Working Paper No. 3089. World Bank, Washington, D.C.

Strauss, J., and D. Thomas. 1998. “Health, Nutrition, and Economic Development.” *Journal of Economic Literature* 36: 436-82.

Taylor, B., E. Dearing and K. McCartney. 2004. “Incomes and Outcomes in Early Childhood.” *Journal of Human Resources* 39(4):980-1007.

Todd, P.E., and K.I. Wolpin. 2003. “On the Specification and Estimation of the Production Function for Cognitive Achievement.” *The Economic Journal* 113: F3-F33.

Umbel, V.M., B.Z. Pearson, M.C. Fernandez, and D.K. Oller. 1992. “Measuring bilingual children's receptive vocabularies.” *Child Development* 63(4): 1012-20.

Waldfogel, J., W. Han, and J. Brooks-Gunn. 2002. “The Effects of Early Maternal Employment on Child Cognitive Development.” *Demography* 30(2): 369-92.

Walker SP, S.M. Grantham-Mcgregor, C.A. Powell, S.M. Chang. 2000. “Effects of Growth Restriction in Early Childhood on Growth, IQ, and Cognition at Age 11 to 12 Years and the Benefits of Nutritional Supplementation and Psychosocial Stimulation.” *The Journal of Pediatrics* 137(1): 36-41.