The Effect of Early Childhood Developmental Program Attendance On Future School Enrollment and Grade Progression in Rural North India

Gautam Hazarika*
Department of Finance and Economics
School of Business
The University of Texas at Brownsville
80 Fort Brown
Brownsville, TX 78520
Email: gautam.hazarika@utb.edu
Ph: 956-882-8953
Fax: 956-882-5808

Vejoya Viren
Department of Educational Psychology and Leadership Studies
College of Education
The University of Texas at Brownsville
80 Fort Brown
Brownsville, TX 78520
Email: vejoya.viren@utb.edu

* Corresponding author

The Effect of Early Childhood Developmental Program Attendance
On Future School Enrollment and Grade Progression in Rural North India

Abstract

This paper examines the effect of prior participation in early childhood developmental programs, considered endogenous, upon 7 – 14 years olds' school enrollment and grade progression in rural North India. It hopes both to contribute to extending to less developed countries recent influential research on the long-term benefits of early childhood interventions in the United States, and to make a case for the inclusion of such interventions amongst developing nations' policy initiatives toward expanding schooling. Analysis of data from the World Bank's 1997-98 Survey of Living Conditions in Uttar Pradesh and Bihar yields the findings that early childhood developmental program attendance at ages 0 – 6 raises the probability of school enrollment among average 7 – 14 year olds by about 27.5 percentage points, and that this beneficial early experience also significantly hastens students' grade progression.

JEL Classification: 012, 015, I21

Keywords: early childhood, schooling, India

1. Introduction

This study has two objectives. It aims to help extend to less developed countries the exciting new burgeoning body of economic research on early childhood development in the United States[1] (e.g., Heckman, 2008). It intends also to investigate the efficacy of a somewhat overlooked, by academics and policy makers, means of boosting school attendance and attainment in less developed countries, namely, formal early childhood care and education.

Whilst early childhood development has perennially interested educationists and psychologists, it has only recently piqued the attention of economists. Economists in the United States have begun to view early childhood developmental interventions as cost effective means of redressing rising income inequality. Heckman (2008) observes that "American society is polarizing". Cunha and Heckman (2008) show that factors determined by as early as age 18, such as family background, account substantially for interpersonal variation in lifetime earnings. It would appear that American adolescents' destinies, though not set in stone, are etched in rapidly hardening cement. That children cannot purchase their family backgrounds, and that parents without means cannot borrow to invest in their young children, lead to market failure[2] (Cunha, Heckman, Lochner, and Masterov, 2006), remediable by public early childhood developmental interventions. Cunha, Heckman, Lochner, and Masterov (2006) summarize some experimental evidence, gleaned from randomized controlled trials, of the long-term benefits to children from early intervention programs in the United States. Exposure to superior preschool environments may lead to higher achievement test scores, higher grades at school, higher rates of high school graduation, higher rates of employment, better jobs, lower propensities to crime, more healthful lifestyles, lower rates of teen pregnancy, and higher rates of marriage. A recent widely publicized analysis[3] by Chetty et al. (2010) of data also yielded by an experiment utilizing a randomized protocol indicates that children who learn more in kindergarten, mainly from better teaching, enjoy higher earnings by age 27, are likelier to attend college, likelier to become home owners by age 28, likelier to save for retirement, and less likely to become single parents. Such studies are

1. While it is clear development economists' scrutiny of early childhood hasn't lagged the blossoming of economists' interest in early childhood in the United States, it's principal object has been early childhood health and nutrition (e.g., Alderman, Behrman, Lavy, and Menon, 2001), in that there are few economic studies of the benefits of preschool attendance in developing countries. Of the 56 studies of early childhood interventions worldwide surveyed by Nores and Barnett (2010), but 17 were recognizably economic, of which merely three (Behrman, Cheng, and Todd, 2004; Armecin, Behrman, Duazo, Ghuman, Gultiano, and King, et al., 2006; Berlinski, Galiani, and Manacorda, 2008) examined the benefits of preschool attendance.

2. in that the existing 'market' for early childhood human capital investment delivers, from society's viewpoint, a suboptimal equilibrium quantity.

3. described in the New York Times (Leonhardt, 2010)

yet to be replicated in the developing world[4], where, it might be argued, the accident of birth has farther reaching consequences.

Economists have long considered schooling a key factor in economic growth and development. Macroeconomists have found a strong positive relation between a nation's rate of economic growth and the educational attainments of its citizenry in cross-country growth regressions (e.g., Barro, 1991), though the direction of causation isn't wholly established[5] (e.g., Bils and Klenow, 2000). Microeconomists, focused on the private gains from education, have estimated high rates of return to schooling in developing countries (e.g., Psacharopoulos and Patrinos, 2004[6]). More educated farmers have been found to be likelier to adopt new 'Green Revolution' technologies (e.g., Foster and Rosenzweig, 1996). Further, parents' education, especially mothers' education, has been seen to improve child health and nutrition, and promote children's schooling (e.g., Schultz, 2002). So it has been common for developing nations to pursue policies to boost school enrollment, attendance, and attainment.

Glewwe and Kremer (2006) summarize the principal obstructions to schooling in developing countries. These are: lack of access to schools, their high cost, their low quality, and, lastly, poor child health. Numerous empirical studies employing experimental, quasi-experimental, and non-experimental-survey data largely confirm that removal of these obstructions stands to substantially expand schooling. Since randomized experiments are now often considered the 'gold standard' in social policy assessment[7] , we dwell below on a few such studies which employ experimental or quasi-experimental data.

Burde and Linden (2010) analyze data generated by an experiment in the random assignment of new primary schools to villages in north-west Afghanistan, to find that the intervention raised enrollment in formal schools by 42 percentage points. Hence, improving access to schools may be a very effective means of raising school enrollment in the developing world.

---

4. MIT's Abdul Latif Jameel Poverty Action Lab (J-PAL) has participated in a small number of randomized controlled trials in preschool settings in the developing world in recent years, but the studied children haven't yet been followed into subsequent schooling and beyond.

5. Instead growth may raise schooling. e.g., growth, even if it is skill-neutral, may increase demand for schooling by raising its effective rate of return, as students giving up current earnings to attend school stand to earn larger sums, owing to economic growth, upon leaving school (Bils and Klenow, 2000). Growth, yielding resources for the expansion of national schooling infrastructure, may even raise the supply of schools.

6. The mean private rate of return to schooling in low income countries is about 11%.

7. though this new eminence accorded them has been pointedly questioned (e.g., Deaton, 2009).

Schultz (2004), in a study of Mexico's *Progresa* poverty program, which confers cash grants on parents provided their children attend school, conditional cash transfers which may be viewed as lowering the cost of schooling, finds that the Program raised school enrollment in grades 1 through 8 by 3.4%. These *Progresa* data are experimental in that they pertain to a subset of eligible communities divided in random fashion into a treatment group and a control group. Evens, Kremer, and Ngatia (2009) analyze experimental data yielded by a randomized controlled trial in Kenya which awarded children school uniforms by lottery. The cost of schooling of lottery winners may be considered reduced. It is found that school absenteeism among all winners of uniforms fell by 44% on average, with a steeper reduction on average, 62% , in absenteeism among winners who didn't previously own a uniform. Hence, reducing the monetary cost of schooling in less developed countries may significantly raise school enrollment and attendance.

Case and Deaton (1999) examine the effect of the quality of schools in South Africa upon children's educational outcomes. It is significant that the data studied by these authors date to 1993, a year before the end of apartheid, when blacks, denied political representation, did not control funding for their children's schools. Further, internal black migration was then strictly controlled, ruling out the migration of families to localities with better schools. So, the assignment of black children to schools of varying quality may be considered random, that is, determined within a natural experiment. The authors discover that a lowering of the pupil-teacher ratio, their primary measure of school quality, in black schools from 40 to 20 would have added 0.75 years to grade attainment by age 10 and 1.5 years by age 15. There is a modicum of evidence, therefore, that raising school quality improves educational outcomes in less developed countries[8].

Finally, Miguel and Kremer (2004), in an analysis of data generated by a randomized controlled trial in Kenya, conclude that initiatives to improve children's health are a potent means of raising school attendance. The experiment introduced a mass de-worming treatment in schools by random assignment in that some schools were randomly chosen to receive the treatment earlier than others. This simple and inexpensive health initiative is found to reduce school absenteeism by as much as 25%, with untreated children in treatment schools benefiting as well from the positive externality of a lower rate of inter-pupil infection.

It is notable that this voluminous body of economic research devoted to uncovering the effective means of promoting schooling in the developing world makes no mention of one potentially very effectual means, namely, formal

8. On the other hand, literature surveys by Hanushek (1986, 1995) find but a tenuous link between school quality, as measured by expended resources, and student outcomes.

early childhood care and education. This is an unfortunate omission given the highly publicized long-term benefits of early childhood interventions in the United States, an established theoretical literature in education and psychology fully predictive of these benefits, and the intriguing sizeable positive correlation between low and lower-middle income nations' preschool enrollment rates and their primary school enrollment rates in data from the UNESCO Institute for Statistics[9]. To be fair, development economists have been hampered by a paucity of data. It isn't common for survey data from less developed countries to inform of children's participation in preschool programs, simply because this is rare, especially amongst the poor: data from the UNESCO Institute for Statistics reveal that the mean gross preschool enrollment rate[10] in low income countries was merely about 14% in 2007. Experimental data are scarce as well since there have been but few randomized controlled trials in preschool settings in developing countries, all of them recent enough as to rule out long-term assessments.

This study aims to contribute to stopping this gap in the literature by uncovering evidence of a strong beneficial effect of children's participation in early childhood development programs upon their future schooling in a developing country. It hopes, thereby, to help build a case for the inclusion of early childhood interventions amongst developing nations' policy initiatives toward expanding schooling. The data examined, from India, are well suited to this purpose since India is rare among low income countries in having a 35 year old nation-wide network of public early childhood developmental facilities called *Anganwadi*[11] Centers (AWCs), established as part of the Government of India's Integrated Child Development Services (ICDS) Program. It is creditable that the Government of India has long considered early childhood care and education vital to its efforts at universalizing primary education[12]. This analysis reveals that the average 7 – 14 year old rural North Indian is 27.5 percentage points likelier to be currently enrolled in school as a result of having previously participated in an early childhood developmental program, whether in an *Anganwadi* Center, an analogous facility run by a non-governmental organization[13], or a preschool classroom[14]. Further, of those presently

9. Linear regression of 2004-08 average primary enrollment rates in low and lower-middle income countries against these nations' 2004-08 average preschool enrollment rates yields the estimated regression equation *primary enrollment rate = 72.9 + 0.29×preprimary enrollment rate*, with R^2 and slope-coefficient t-ratio of, respectively, 0.22 and 4.13.

10. preschool enrollment as a percentage of all preschool age children

11. The Hindi word *anganwadi* means 'courtyard garden'.

12. According to the UNESCO Institute for Statistics, 34.2% of Indian children enrolled in the first grade of primary school in 2005 were not expected to complete their primary education.

13. These include urban *Balwadi* (kindergarten) Centers run by the NGO *Pratham*. Urban India is relatively underserved by the

enrolled, children who participated in early childhood developmental programs are found to have enjoyed faster age-normed grade progression, due presumably to more timely admission to primary school as well as lower rates of grade repetition.

The remainder of this paper is organized as follows. Section 2 draws on a vast literature in neurobiology, nutrition, and psychology to explain the pivotal nature of early childhood. Section 3 describes the empirical methodology. Section 4 describes the utilized data and the empirical findings, with section 5 offering a brief conclusion.

2. The Pivotal Nature of Early Childhood

It is well understood that early childhood is a critical period in human development. Early childhood experiences are capable of leaving indelible marks upon abilities and personality. Abilities are usually taken to mean cognitive abilities, such as intelligence as measured by an IQ test. There is no doubt cognitive abilities are important determinants of life success. However, it is now realized that personality traits like self-regulation, motivation, and sociability have significant bearing on life outcomes as well. Indeed, Heckman, Stixrud, and Urzua (2006) find that such 'noncognitive abilities' are often as strongly correlated with schooling, labor market outcomes, and social (mis)behavior as cognitive abilities[15].

Early childhood is a critical period in the development of cognitive and noncognitive abilities because the requisite wiring of neural circuitry largely occurs at this time. Early life experiences strongly shape this wiring both because neural circuitry is highly plastic at this stage, and because patterns of neural connections are more easily formed when there aren't any established patterns to begin with, that is, when the slate is blank (Knudsen, Heckman, Cameron, and Shonkoff, 2006). For example, when a young child has worse vision in one eye than in the other, that is, is more reliant on one of her eyes, neural connections from it to her brain proliferate whereas connections from the eye less relied on wither away, and this process is *irreversible* once a sensitive period in early childhood is past.

Similarly, children unexposed to language during a sensitive period of early childhood may become largely

---

Government's ICDS Program.

14. According to India's National Institute of Educational Planning and Administration (NIEPA), 14.27% of public primary schools contained preschool sections in 2003.

15. For example, noncognitive skills have approximately the same positive effect on wages as cognitive skills, with noncognitive skills exerting a stronger salubrious effect on employment propensities than cognitive skills.

incapable of language because of attrition of their brains' neural circuitry responsible for language acquisition. So called 'feral children', of whom there are many historical examples[16] (Benzaquen, 2006), may have suffered precisely this form of neural atrophy, for no rescued feral child ever learnt to speak fluently. It is curious that these linguistically challenged feral children were also severely cognitively challenged. It is possible that they were abandoned by their families *because* they were cognitively challenged, but this could not have been true of feral children who, when rescued, were far too young[17] to have been sensibly assessed as mentally impaired. In light of the highly influential human developmental theories of the Soviet psychologist Lev Vygotsky[18], it is likelier that these children's cognitive impairment was tied to their lack of language. Vygotsky held that language, paramount of mental tools, played a central role in the acquisition of other mental tools. By Vygotskian theory, language makes thought possible, that is, thoughts come into existence through words[19]. Indeed, Vygotsky held that words play a central role in the growth of consciousness. Thus, acquisition of the cognitive skill of language, largely impossible outside a fairly narrow window in childhood, may open doors to higher cognitive functioning. Language may also influence the development of noncognitive skills. For example, young children use 'private speech' or self-talk as a tool of self-regulation or self-control (Leong and Bodrova, 2006), an important noncognitive skill linked with future success[20].

There is evidence that other personality traits, like reactivity to stress, are also closely formed in early childhood. For example, biological experiments have found that rat pups reared in the first week of their lives by a relatively indifferent dam, not necessarily their mother[21], grew up to be more reactive to stress, more anxious, more fearful, and less adventurous than pups reared by a more solicitous dam. These permanent changes in temperament seem rooted in

16. These were children who'd ostensibly been brought up by wild animals after being abandoned in infancy. Victor of Aveyron was one such child, finally rescued in 1800 in southern France.

17. For example, the feral child Amala was only 18 months old when rescued near Midnapore, eastern India, in 1920.

18. 1896 – 1934

19. Indeed, it is held (e.g., Watson, 1929) that thinking involves unconscious movement of the vocal chords.

20. There is a considerable literature, to which belongs Stanford psychologist Walter Mischel's famous 'Marshmallow Experiment', supportive of a strong positive relation between patience in young children, a type of self-regulation, and their future cognitive and social competence (e.g., Mischel, Shoda, and Rodriguez, 1989). In the Marshmallow Experiment, 4-year olds in a laboratory school at Stanford University in the late 1960s were each presented with a marshmallow and told that if they waited 20 minutes before eating it, they would receive another. It was observed that some children were able to wait longer than others. The researchers then followed these children into adolescence to discover striking patterns.

21. since the experiments attempted cross-fostering in order to sever genetic links between the pups and their caregivers

epigenetic[22] changes wrought by the excessive release of stress hormones in early life[23]. Their cause being epigenetic, such modification to temperament may be passed on to future generations. It is believed that this model is applicable to humans as well (e.g., Lords, 2009)[24].

There can be no starker or more poignant demonstration of the critical nature of early childhood than the grim lifelong consequences of the *in utero* and early childhood malnutrition common in less developed countries[25]. Malnutrition *in utero* and during the first two or three years of life may irreversibly stunt physiques, reduce cognitive development so as to permanently lower IQ, decrease attention and focus, hinder learning, impede educational attainment, and lead to behavioral difficulties and poor social skills (Martorell, 1999).

In sum, it is clear that early childhood is a momentous phase in human development. A young child's experiences may well set the tone for the rest of her life. Moreover, her experiences may, via epigenetic transmission, pass on to her descendents.

3. Empirical Models and Estimation

This study aims to gauge the effect of rural Northern Indian children's participation in early childhood developmental programs upon their subsequent school attendance and, given attendance, their grade progression. School attendance in measured in binary fashion, and grade progression continuously as 'schooling for age', or *SAGE* [26], defined as

$$[\text{ grade presently attending} \div (\text{age} - 6)] \times 100 .$$

The effect of participation in early childhood developmental programs, the 'treatment' in this instance, upon future school attendance may be estimated via the equation

$$attend_i^* = X_i'a_1 + a_2 . ECD_i + e_{1i} , \qquad (1)$$

22. Epigenetic changes are changes in gene expression (the 'turning on' or 'turning off' of genes) caused by mechanisms other than changes in the underlying DNA sequence. Such changes are inheritable, at least within proximal generations.

23. The brains of the more fearful pups had far fewer glucocorticoid receptors, responsible for dampening down the release of stress hormones, owing to these pups' early childhood stress 'turning off' certain genes responsible for the growth of these receptors.

24. It has been found, for example, that post-traumatic stress disorder in Holocaust survivors is a risk factor in the development of this disease in their adult offspring (Yehuda and Bierer, 2007).

25. For example, 41% of 0 – 5 year children in South Asia are malnourished in the sense of being underweight.

26. used, for example, by Ray and Lancaster (2005)

where $attend_i$* is a latent variable underlying its observed binary counterpart, $attend_i$, an indicator of whether child i, 7 to 14 years old, is presently attending school, the X_i are personal, household, and community characteristics with bearing on school attendance, ECD_i is an indicator of whether child i participated in an early childhood developmental program between the ages of 0 and 6, and e_{1i}, the error term, taken to be standard normally distributed, denotes unobserved influences upon school attendance. (1) may not simply be estimated by probit ML because it is likely that participants in early childhood developmental programs are different from non-participants in unobserved ways. For example, parents with strong taste for educated children may both enroll them in an early childhood developmental program and insist that they continue formal schooling long afterwards. Alternatively, certain early childhood developmental programs target 'at risk' children from disadvantaged families unlikely to foster educational attainment. If the unobserved aspects of such family attributes were subsumed within the error term e_1, estimates of the coefficient a_2 would be inconsistent. It would be necessary, then, to estimate (1) together with a second equation

$$ECD_i^* = X_i'b_1 + u_i , \qquad (2)$$

where ECD_i* is the latent variable underlying the binary ECD_i, and u_i, the standard normally distributed regression error, denotes unobserved influences upon early childhood developmental program participation that are correlated with unobserved influences upon subsequent school attendance. (1) and (2) constitute a recursive simultaneous equations system with binary dependent variables, corresponding to Maddala's and Lee's (1976) Model 1. Notice that (1) has exactly the same exogenous regressors as (2), that is, there is no identifying exclusion restriction. How, then, is (1) identified? Wilde (2000) makes the important observation that because the model comprising (1) and (2) features both latent variables as well one of these variables' realized qualitative counterpart (the indicator *ECD* is a regressor in (1) whereas the dependent variable in (2) is rather the latent *ECD** that underlies *ECD*), the usual identifying restriction that the exogenous regressors in (2) include at least one variable excluded from the exogenous regressors in (1), is, in empirical practice, unnecessary. (1) would be unidentified if the linear combination of (1) and (2), namely,

$$\lambda_1 . attend_i^* + \lambda_2 . ECD_i^* = \lambda_1 . (X_i'a_1 + a_2 . ECD_i + e_{1i}) + \lambda_2 . (X_i'b_1 + u_i),$$

which yields

$$attend_i^* = X_i' [a_1 + (\lambda_2 / \lambda_1).b_1] + a_2 .ECD_i - (\lambda_2 / \lambda_1).ECD_i^* + e_{1i} + (\lambda_2 / \lambda_1). u_i \qquad (3)$$

contained the very same variables as (1). But it is evident that (3), unlike (1), contains the term $(\lambda_2 / \lambda_1).ECD_i$*, which makes it structurally different. "Thus the classical identification problem does not exist", argues Wilde (2000), as did

Maddala (1983)[27]. However, Wilde (2000) notes, as did Maddala (1983)[28], that the parameters of the model would be unidentified yet if $X_i = [1]$, that is, if the regressors in (1) included but a constant term and ECD_i , and the only regressor in (2) were a constant term. In this special case, the likelihood function for the joint estimation of (1) and (2) would involve the four probabilities

$$\text{pr.}(attend_i^* > 0, ECD_i^* > 0) = \text{pr.}[\, e_{1i} > -a_1 - a_2 \,,\, u_i > -b_1 \,] = \text{F}(a_1 + a_2 \,,\, b_1; \rho) \,,$$

$$\text{pr.}(attend_i^* > 0, ECD_i^* \leq 0) = \text{pr.}\,[\, e_{1i} > -a_1 \,,\, u_i \leq -b_1 \,] = \text{F}(a_1 \,,\, -b_1; -\rho) \,,$$

$$\text{pr.}(attend_i^* \leq 0, ECD_i^* > 0) = \text{pr.}[\, e_{1i} \leq -a_1 - a_2 \,,\, u_i > -b_1 \,] = \text{F}(-a_1 - a_2, b_1; -\rho) \text{ , and}$$

$$\text{pr.}(attend_i^* \leq 0, ECD_i^* \leq 0) = \text{pr.}\,[\, e_{1i} \leq -a_1 \,,\, u_i \leq -b_1 \,] = \text{F}(-a_1 \,,\, -b_1; \rho) \,,$$

where $\rho = \text{corr.}(\, e_{1i} \,,\, u_i)$ and F( ) denotes the cumulative distribution function of the bivariate standard normal distribution. Of these probabilities, only three would be independent since the four must sum to one. *Three* independent probabilities are insufficient for the purposes of estimating the *four* parameters of the model, namely, a_1, a_2, b_1, and ρ. However, Wilde (2000) observes that the model would be identified if $X_i' = [1 \;\; x_i]$, where x_i is an exogenous value-varying regressor. By introducing two slope coefficients, this modification would raise the number of parameters in the model to six. But there would now be three independent probabilities for *each* value of x_i , so that if x_i took even two different values, six independent probabilities are yielded, which are sufficient in number for the estimation of the model's six parameters. Hence, Wilde (2000) contends that "the existence of one varying exogenous regressor in each equation is sufficient to avoid small variation identification problems in multiple equation probit models with endogenous dummy regressors". In sum, as long (1) and (2) include value-varying regressors, the parameters of the model would be identified even without exclusion restrictions[29]. This is a most useful result since it obviates inevitably controversial 'instrumenting' of the regressor ECD_i in (1).

What of estimation? Greene (1998) argues that the equations (1) and (2) might be jointly estimated by bivariate probit ML "as if there were no simultaneity problem".

Similarly, the effect of participation in early childhood developmental programs upon grade progression in school,

---

27. p. 122

28. ibid.

29. This isn't akin to identification based on non-linear functional form, as when the selection equation contains exactly the same variables as the primary equation in the case of Heckman's Two-Step Estimation. Rather, identification in this instance is based on the variation in data so common in empirical practice.

conditional on enrollment, may be estimated via the equation

$$SAGE_i = X_i'a_3 + a_4 \,.\, ECD_i + e_{2i}, \qquad (4)$$

where, as described, $SAGE_i$ is a continuous variable measuring the pace of enrolled child *i*'s grade progression, the X_i, personal, household, and community characteristics with bearing on school attendance, are plausible correlates of grade progression as well, ECD_i is, as before, an indicator of whether child *i* participated in an early childhood developmental program between the ages of 0 and 6, and e_{2i}, the error term, denotes unobserved influences upon grade progression. (4) may not simply be estimated by OLS since it is probable that participation in an early childhood developmental program is endogenous. For example, parents with strong taste for educated children may both enroll them in an early childhood developmental program and, later, ensure that they don't lag in school. If unobserved aspects of such parental taste were subsumed within the error term e_2, OLS estimates of the coefficient a_4 would be biased. So it would be necessary to estimate (4) together with (2). Equations (4) and (2) constitute a recursive simultaneous equations system with one of its dependent variables continuous and the other binary, corresponding to Maddala's and Lee's (1976) Model 2, and may be estimated by maximum-likelihood[30]. Notice that (4) has exactly the same exogenous regressors as (2), that is, there is, again, no identifying exclusion restriction. Wilde's (2000) argument that such a restriction is unnecessary in multiple equation probit models with dummy endogenous regressors so long as there is sufficient variation in the data, may be extended to this model. Indeed, identification in this case is aided not only by variation in the regressors X_i but also by variation in the continuous dependent variable $SAGE_i$.

4. Data and Empirical Findings

The empirical models described in section 3 above are fitted to data from the 1997-98 Survey of Living Conditions in Uttar Pradesh and Bihar, which is unusual in having inquired of children's participation in early childhood development programs. The survey is part of the Living Standards Measurement Survey (LSMS) series of the World Bank. It covered 2,250 households drawn from 120 villages in 25 districts of rural southern and eastern Uttar Pradesh and rural northern and central Bihar. With nearly 176 million residents, Uttar Pradesh is India's most populous state. Bihar, with a population of nearly 83 million, is India's third-most populous state31. Both states are ranked amongst the lowest

30. using, STATA's *treatreg* command, for example

31. In 2000, the two states were subdivided. Uttaranchal, renamed Uttarakhand, was carved from Uttar Pradesh and Jharkhand from

in the Indian Union by almost all indicators of poverty and socioeconomic development.

These data yield a sample of 2,802 7 – 14 year olds. Table 1 presents the sample mean values of all the variables featuring in the statistical analysis. 66.8% of these 7 – 14 year olds were enrolled in school at the time of the survey. 54.6% of the sample was male. These 7 – 14 year olds' average age was about ten. 10% of them were Muslim[32]. The average child's household agricultural assets (land, livestock, and agricultural capital equipment) were worth about 210,000 rupees[33]. 19.7% of these 7 – 14 year olds dwelt in households owning *pucca* homes, that is, homes built typically of brick and mortar rather than less durable material. The average child's household annual unearned income, from pensions and other benefits as well as gifts and remittances, was about 1073 rupees. Household earned income was excluded from the analysis since child work at the expense of schooling might contribute to it, making it endogenous. The average child lived 0.6 kilometers away from the closest primary school, 2.7 kilometers away from the closest middle school, and at a distance of about 4.9 kilometers from the nearest secondary school. Only 53.7% of these children lived in villages serviced by a paved road. 31.2% lived in villages in which all agricultural land was irrigated. 44.9% of these 7 – 14 year olds hailed from the state of Bihar.

Table 2 presents estimates of (1), arrived at by joint-estimation of (1) and (2) as described in section 3[34]. They indicate that boys are significantly likelier than girls to have been enrolled in school at the time of the Survey. The probability of school enrollment increases in the number of older household members, presumably because the availability of more such persons to shoulder chores reduces the burden of work upon 7 – 14 year olds. On the other hand, since children in this setting commonly assist in the rearing of their younger siblings, often at the cost of schooling, a 7 – 14 year old's propensity to enroll in school is expected to be inversely related to the number of younger members of her household. This expectation isn't belied, though the variable is not significant at customary levels. A child from a household whose head[35] is illiterate[36] appears significantly less likely to be enrolled in school. School enrollment seems

---

Bihar.

32. 13.4% of India's population is Muslim.

33. The USD-INR exchange rate at this time was about $1 = Rs. 36.

34. Corresponding estimates of (2) may be found in column 1 of table 4.

35. a parent in 74.1% of cases, a grandparent in 21.2% of cases

36. Educational attainment in these data is measured as a categorical variable, whose category 'illiterate' applied most widely to the surveyed adults.

subject to a wealth effect in that the probability of enrollment increases in household ownership of a *pucca* home. School enrollment appears deterred by distances to the closest middle and secondary schools, and the monetary cost of schooling. Children in Bihar are less likely to have been enrolled in school at the time of the Survey than their peers in Uttar Pradesh.

By these estimates, 7 – 14 year olds who participated in early childhood developmental programs when they were 0 – 6 years old are significantly likelier to have been enrolled in school at the time of the Survey. Indeed, the average 7 – 14 year old is 27.5 percentage points likelier to be enrolled in school as a result of previously participating in an early childhood developmental program. Further, unobserved influences upon, respectively, current school enrollment and past participation in an early childhood developmental program appear significantly negatively correlated, indicating that the treatment of the regressor ECD_i in (1) as endogenous is warranted.

Table 3 presents estimates of (4), arrived at by joint-estimation of (4) and (2) as described in section 3[37]. It is found that boys enjoy more rapid grade progression than girls, and that age slows grade progression. Grade progression is more rapid in households with greater numbers of older members. A student from a household whose head is illiterate is likelier to be attending a lower grade than is age-appropriate. The pace of grade progression increases in household wealth as measured by agricultural assets and *pucca* home ownership. Grade progression is slowed by household non-farm enterprises. Perhaps students must tend to their households' non-farm enterprises, such as village shops, at the expense of their studies[38]. The monetary costs of primary schooling seem to raise students' pace of grade progression. Perhaps a high cost of schooling serves to winnow out less able or motivated students. Finally, the estimates indicate that 7 – 14 year old students who previously participated in an early childhood developmental program progress through grade levels at a distinctly speedier rate than students without such beneficial early experience. It is notable that unobserved influences upon, respectively, grade progression and previous participation in an early childhood developmental program seem strongly negatively correlated, which suggests that OLS applied to (4) will yield a downwardly biased estimate of the benefit of early childhood developmental interventions.

6. Conclusion

In sum, it is found that rural North Indian 7 – 14 year olds with prior experience of preschool care and education

37. Corresponding estimates of (2) may be found in column 2 of table 4.

38. It has been argued (e.g., Wydick, 1999) that household retail enterprises in these settings are particularly at risk of petty pilferage by hired hands, making them more reliant on household labor.

are significantly likelier to be enrolled in school. Indeed, the average such child is about 27.5 percentage points likelier to be enrolled in school as a result of early childhood developmental interventions. Further, students with such beneficent early experience enjoy faster age-normed grade progression. These findings make a strong case for the inclusion of early childhood interventions amongst developing nations' policy tools for the expansion of schooling.

This study also contributes to extending to less developed nations recent influential research upon the many benefits of early childhood interventions in the United States. Its findings will, it is hoped, remind educationists and policy makers in less developed countries, as well as development economists, of the pivotal nature of early childhood.

TABLE 1
Sample means

	7 – 14 year olds **Sample Mean (S.D.)**	Enrolled 7 – 14 year olds **Sample Mean (S.D.)**
Dependent Variables		
Currently enrolled in school (= 0,1)	0.668 (0.471)	
SAGE (schooling for age)		101.692 (55.362)
Child Attributes		
Male (= 0,1)	0.546 (0.498)	0.623 (0.485)
Age (years)	10.042 (2.212)	10.090 (2.185)
Household Attributes		
No. of household members aged 15 or older	4.061 (2.364)	4.326 (2.476)
No. of household members aged 6 or younger	1.525 (1.408)	1.522 (1.471)
Household head is illiterate (= 0,1)	0.483 (0.500)	0.372 (0.484)
Household head is female (= 0, 1)	0.032 (0.176)	0.027 (0.161)
Household's religion is Islam (= 0,1)	0.101 (0.301)	0.089 (0.284)
Value in ′00,000 Rupees of household agricultural assets	2.100 (6.145)	2.635 (7.200)
Household owns a *pucca* home (built of brick & mortar) (= 0,1)	0.197 (0.398)	0.242 (0.429)
Value in ′000 Rupees of annual household unearned income	1.073 (7.020)	1.335 (8.338)
No. of household non-farm enterprises	0.520 (0.748)	0.535 (0.777)
Village Attributes		
Distance to nearest primary school (km)	0.560 (0.847)	0.574 (0.806)
Distance to nearest middle school (km)	2.687 (2.353)	2.537 (2.223)
Distance to nearest secondary school (km)	4.868 (3.983)	4.599 (3.859)
Village-average annual monetary cost in ′00 Rupees of primary schooling	3.687 (2.384)	3.639 (2.278)
Village is accessible by a paved road (= 0,1)	0.537 (0.499)	0.530 (0.499)
All village agricultural land is irrigated (= 0,1)	0.312 (0.464)	0.350 (0.477)
Village is located in the state of Bihar (= 0,1)	0.449 (0.497)	0.404 (0.491)
Key Variable		
Child attended an early childhood developmental program at ages 0 – 6 (= 0,1)	0.095 (0.294)	0.100 (0.301)
n =	2802	1873

TABLE 2

Determinants of School Enrollment Among 7 – 14 Year Olds (ML Estimates)

	Coeff. (S.E.)
Constant	0.666*** (0.213)
Child Attributes	
Male (= 0,1)	0.626*** (0.067)
Age (years)	–0.017 (0.013)
Household Attributes	
No. of household members aged 15 or older	0.075*** (0.018)
No. of household members aged 6 or younger	–0.045 (0.028)
Household head is illiterate (= 0,1)	–0.763*** (0.085)
Household head is female (= 0, 1)	–0.013 (0.178)
Household's religion is Islam (= 0,1)	–0.160 (0.113)
Value in ′00,000 Rupees of household agricultural assets	0.030 (0.019)
Household owns a *pucca* home (built of brick & mortar) (= 0,1)	0.317*** (0.095)
Value in ′000 Rupees of annual household unearned income	0.013 (0.012)
No. of household non-farm enterprises	–0.007 (0.045)
Village Attributes	
Distance to nearest primary school (km)	0.002 (0.041)
Distance to nearest middle school (km)	–0.030** (0.015)
Distance to nearest secondary school (km)	–0.021** (0.010)
Village-average annual monetary cost in ′00 Rupees of primary schooling	–0.036*** (0.013)
Village is accessible by a paved road (= 0,1)	–0.002 (0.068)
All village agricultural land is irrigated (= 0,1)	0.088 (0.074)
Village is located in the state of Bihar (= 0,1)	–0.271*** (0.079)
Key Variable	
Child attended an early childhood developmental program at ages 0 – 6 (= 0,1)	1.113*** (0.404)
Rho = corr.(e_{1i} , u_i)	–0.539 (*p*-value = 0.054)
Log-Likelihood	–2321.872
n =	2802

Note: *, **, & *** indicate, respectively, significance at the 10%, 5%, & 1% levels

Standard errors corrected for correlated errors within a household

Dependent Variable = Currently enrolled in school (= 0,1)

TABLE 3

Determinants of Grade Progression Among Enrolled 7 – 14 Year Olds (ML Estimates)

	Coeff. (S.E.)
Constant	175.399*** (9.175)
Child Attributes	
Male (= 0,1)	10.546*** (2.624)
Age (years)	–9.685*** (0.696)
Household Attributes	
No. of household members aged 15 or older	2.055*** (0.750)
No. of household members aged 6 or younger	–2.159 (1.389)
Household head is illiterate (= 0,1)	–6.870** (3.239)
Household head is female (= 0, 1)	0.966 (7.654)
Household's religion is Islam (= 0,1)	–5.700 (6.214)
Value in ′00,000 Rupees of household agricultural assets	0.777** (0.337)
Household owns a *pucca* home (built of brick & mortar) (= 0,1)	9.372*** (3.435)
Value in ′000 Rupees of annual household unearned income	0.043 (0.087)
No. of household non-farm enterprises	–5.292** (2.078)
Village Attributes	
Distance to nearest primary school (km)	–0.521 (1.989)
Distance to nearest middle school (km)	0.447 (0.780)
Distance to nearest secondary school (km)	–0.191 (0.401)
Village-average annual monetary cost in ′00 Rupees of primary schooling	1.402** (0.632)
Village is accessible by a paved road (= 0,1)	1.691 (3.045)
All village agricultural land is irrigated (= 0,1)	1.736 (3.411)
Village is located in the state of Bihar (= 0,1)	4.813 (3.545)
Key Variable	
Child attended an early childhood developmental program at ages 0 – 6 (= 0,1)	73.623*** (6.231)
Rho = corr.(e_{2i} , u_i)	–0.803 (*p*-value = 0.000)
Log-Likelihood	–10392.986
n =	1873

Note: *, **, & *** indicate, respectively, significance at the 10%, 5%, & 1% levels

Standard errors corrected for correlated errors within a household

Dependent Variable = SAGE (schooling for age)

TABLE 4
Determinants of Prior Participation in Early Childhood Developmental Programs (ML Estimates)

	Col. 1	Col. 2
	7 – 14 year olds (n = 2802)	Enrolled 7 – 14 year olds (n = 1873)
	Coeff. (S.E.)	**Coeff. (S.E.)**
Constant	–1.020*** (0.247)	–1.756*** (0.327)
Child Attributes		
Male (= 0,1)	0.034 (0.069)	0.004 (0.092)
Age (years)	–0.007 (0.014)	0.049* (0.026)
Household Attributes		
No. of household members aged 15 or older	–0.028 (0.022)	–0.065*** (0.020)
No. of household members aged 6 or younger	0.017 (0.054)	0.106* (0.056)
Household head is illiterate (= 0,1)	–0.137 (0.097)	–0.252** (0.105)
Household head is female (= 0, 1)	0.285 (0.208)	0.143 (0.270)
Household's religion is Islam (= 0,1)	0.114 (0.171)	–0.179 (0.205)
Value in ′00,000 Rupees of household agricultural assets	0.005 (0.009)	–0.00008 (0.009)
Household owns a *pucca* home (built of brick & mortar) (= 0,1)	– 0.176 (0.128)	–0.159 (0.131)
Value in ′000 Rupees of annual household unearned income	–0.003 (0.007)	–0.004 (0.012)
No. of household non-farm enterprises	0.084 (0.067)	0.117* (0.065)
Village Attributes		
Distance to nearest primary school (km)	0.028 (0.052)	–0.019 (0.059)
Distance to nearest middle school (km)	0.016 (0.022)	0.015 (0.028)
Distance to nearest secondary school (km)	–0.012 (0.014)	0.0003 (0.013)
Village-average annual monetary cost in ′00 Rupees of primary schooling	0.011 (0.017)	0.027 (0.020)
Village is accessible by a paved road (= 0,1)	–0.177* (0.099)	–0.012 (0.114)
All village agricultural land is irrigated (= 0,1)	–0.029 (0.113)	0.084 (0.117)
Village is located in the state of Bihar (= 0,1)	–0.186* (0.103)	–0.331*** (0.118)
Log-Likelihood	–2321.872	–10392.986
n =	2802	1873

Note: *, **, & *** indicate, respectively, significance at the 10%, 5%, & 1% levels
Standard errors corrected for correlated errors within a household
Dependent Variable = Child attended an early childhood developmental program at ages 0 – 6 (= 0,1)

References

Alderman, H., Behrman, J. R., Lavy, V., & Menon, R. (2001). Child health and school enrollment: A longitudinal analysis. *Journal of Human Resources, 36(1),* 185-205.

Armecin, G., Behrman, J. R., Duazo, P., Ghuman, S., Gultiano, S., & King, E., et al. (2006). Early childhood development through an integrated program: Evidence from the Philippines. World Bank Policy Research Working Paper No. 3922, Washington, DC.

Barro, R. (1991). Economic growth in a cross-section of countries. *Quarterly Journal of Economics, 106(2),* 407-443.

Behrman, J. R., Cheng, Y., & Todd, P. E. (2004). Evaluating preschool programs when length of exposure to the program varies: A nonparametric approach. *Review of Economics and Statistics, 86(1),* 108–132.

Benzaquen, A. S. (2006). *Encounters with wild children: temptations and disappointment in the study of human nature*. Montreal & Kingston: McGill-Queen's University Press.

Berlinski, S., Galiani, S., & Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics, 92(5-6),* 1416-1440.

Bils, M. & Klenow, P. (2000). Does schooling cause growth? *American Economic Review, 90(5),* 1160-1183.

Bobonis, J., Miguel, E., & Sharma, C. P. (2004). Iron deficiency anemia and school participation. Abdul Latif Jameel Poverty Action Lab Paper No. 7, Massachusetts Institute of Technology, Cambridge, MA.

Burde, D. & Linden, L. L. (2010). The effect of village schools: evidence from a RCT in Afghanistan. Abdul Latif Jameel Poverty Action Lab Academic Publication, Massachusetts Institute of Technology, Cambridge, MA.

Case, A. & Deaton, A. (1999). School inputs and educational outcomes in South Africa. *Quarterly Journal of Economics, 114(3),* 1047-1085.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D., & Yagan, D. (2010). How does your kindergarten classroom affect your earnings? Evidence from project star. National Bureau of Economic Research Working Paper No. 16381, Cambridge, MA.

Cunha, F. & Heckman, J. J. (2008). A new framework for the analysis of inequality. *Macroeconomic Dynamics, 12(S2),* 315-354.

Cunha, F. & Heckman, J. J. (2010). Investing in our young people. Institute for the Study of Labor Discussion Paper No.

5050, Bonn.

Cunha, F., Heckman, J. J., Lochner, L. J., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the economics of education.* Amsterdam: North-Holland.

Deaton, A. (2009). Instruments of development: randomization in the tropics, and the search for the elusive keys to economic development. National Bureau of Economic Research Working Paper No. 14690, Cambridge, MA.

Evans, D., Kremer, M., & Ngatia, M. (2009). The impact of distributing school uniforms on children's education in Kenya. Abdul Latif Jameel Poverty Action Lab Academic Publication, Massachusetts Institute of Technology, Cambridge, MA.

Glewwe, P. & Kremer, M. (2006). Schools, teachers, and educational outcomes in developing countries. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the economics of education.* Amsterdam: North-Holland.

Greene, W. (1999). Gender economics at liberal arts colleges: Further results. *Journal of Economic Education, 29(4),* 291-300.

Hanushek, E. (1986). The economics of schooling: production and efficiency in public schools. *Journal of Economic Literature, 24(3),* 1141-1177.

Hanushek, E. (1995). Interpreting recent research on schooling in developing countries. *World Bank Research Observer, 10(2),* 227-246.

Heckman, J. J. (2008). Schools, skills, and synapses. *Economic Inquiry, 46(3),* 289-324.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24(3),* 411-482.

Knudsen, E. I., Heckman, J. J., Cameron, J., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences, 103(7),* 10155-10162.

Miguel, E. & Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of externalities. *Econometrica, 72(1),* 159-217.

Leong, D. J. & Bodrova, E. (2006). Developing self-regulation: The Vygotskian view. *Academic Exchange Quarterly, 10(4),* 33-38.

Leonhardt, D. (2010). The case for $320,000 kindergarten teachers. The New York Times, July 27.

Lords, S. (2009). Bridging nature and nurture: a look at epigenetic effects of stressful social environments on childhood development. *The University of Toronto Medical Journal, 87(1),* 50-52.

Maddala, G. S. (1983). *Limited dependent and qualitative variables in econometrics.* Cambridge: Cambridge University Press.

Maddala, G. S. & Lee, L.F. (1976). Recursive models with qualitative endogenous variables. In S. Berg (Ed.), *Annals of Economic and Social Measurement, 5(4).* Cambridge, MA: National Bureau of Economic Research, Inc.

Martorell, R. (1999). The nature of child malnutrition and its long-term implications. *Food and Nutrition Bulletin, 20(3),* 288-92.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244(4907),* 933-938.

Psacharopoulos, G. & Patrinos, H. A. (2004). Returns to investment in education: a further update. *Education Economics, 12(2),* 111-134.

Ray, R. & Lancaster, G. (2005). The impact of children's work on schooling: Multicountry evidence. *International Labour Review,144(2),* 189-210.

Schultz, T. P. (2002). Why governments should spend more to educate girls. *World Development, 30(2),* 207-225.

Schultz, T. P. (2004). School subsidies for the poor: evaluating the Mexican Progressa poverty program. *Journal of Development Economics, 74(1),* 199-250.

Wilde, J. (2000). Identification of multiple equation probit models with endogenous dummy regressors. *Economic Letters, 69(3),* 309-312.

Wydick, B. (1999). The effect of microenterprise lending on child schooling in Guatemala. *Economic Development and Cultural Change, 47(4),* 853-869.

Yehuda, Y., & Bierer, L. M. (2007). Transgenerational transmission of cortisol and PTSD risk. *Progress in Brain Research, 167,* 121-135.