Does Preschool Entry Age Affect Children's Cognitive and Social Development? Evidence from India

Saikat Ghosh*, Subhasish Dey**

* Leibniz Institute for Educational Trajectories, Germany. (**Corresponding author**)

Address: Wilhelmplatz 3, 96047 Bamberg, Germany.

Email: saikat.ghosh@lifbi.de

**University of Warwick, UK.

Email: subhasish.dey@warwick.ac.uk

***Abstract*:**

There is considerable debate regarding the appropriate age for children to enter preschool. The literature on such debate is sparsely available in developing country contexts. This study uses primary survey data for 1369 children (mean age as 40.61 months, with 50% girls, 78% Hindu and 22% minority religions, 71% General Caste and 29% other Castes). By deploying Heckman-Instrumental Variable (Heckman-IV) regression approach, we estimate the association between preschool entry age and children's cognitive and social skill accumulation in grade 1.

Our findings suggest that delaying preschool entry significantly negatively impacts children's social and cognitive skills in elementary school. Hence we advocate for early entry into preschool, particularly for children from the disadvantaged section of society.

Keywords: Kindergarten, Kindergarten entry age, cognitive skill, social skill, Heckman-IV

JEL Classification Code: I21, I24, I26, I28, I29.

Introduction:

What should be the appropriate age for children to enter preschool or kindergarten is a longstanding question which still has no proper answer. Researchers have studied relative entry age effects for many years. However, there is no consensus regarding the impact of relative maturity due to different entry ages in preschool on student outcomes in the later stage of their life. There is considerable debate regarding the age at which children are ready to enter preschool. Besides, whether there is any heterogeneous outcome of entry into preschool at different ages is still an open question. This study examines the effect of different entry ages in preschool on children' cognitive and social skill formation in grade 1. Alternatively, we examine the time durations spend in preschool by children on their development (both cognitive and social development) in the context of India.

Findings from developmental neuroscience tell us that the brain structures undergo their most dramatic development during the first years of life (Benton, 2010; Johnson, 2001; Shore, 1997). Therefore, life's early years profoundly impact a child's future, and the early years of childhood form the basis of intelligence, personality, social behaviour, and capacity to learn (UNICEF, 2017). Evidence already suggests that interventions such as early preschool attendance are enormously effective for skill development (Currie and Almond, 2011; Heckman, 2008; Cunha and Heckman, 2007; Cunha et al., 2006). However, scholars researching the relative age effect on children's development are not equivocal in this context, and there are two principal views on this issue that shapes the age of entry debate both at the policy and practice level: *First*, early preschool entry, and *Second*, late entry with maturity.

Scholars, who advocate for early preschool entry, rely on chronological age as the sole determining factor for entry into preschool. The argument is that age is exogenous and less susceptible to cultural or social characteristics (Brent et al., 1996; Kagan, 1990; Stipek,

2002). They emphasized more on spending more time in preschool rather than additional biological maturation or general out-of-school experience. Moreover, researchers also argue that defining the threshold of development needed before entering primary school cannot be easily identified as development is multidimensional and often uneven (Stipek, 2002, 4). Besides, evidence showed that the skill difference among children in the early years of their life stems from their family environment and prekindergarten learning. The skill differences before kindergarten tend to fade over time as children spend more time in preschool due to a specific entry age (Stipek, 2002).

Several pieces of evidence show that the early entrants are doing relatively better at later schooling. Among numerous studies available, a recent study by Cornelissen and Dustmann (2019) estimated the effects of receiving additional schooling before age five on cognitive and non-cognitive outcomes in England. It found significant effects on cognitive and non-cognitive outcomes at ages five and seven, particularly for boys with a disadvantaged parental background. Lincove and Painter (2006) have shown that, concerning long-term outcomes, young students have slightly better outcomes on average than older students, as long as there is no retention of the young students in any grade.

Moreover, considering the children from developing countries, where various forms of inequalities are already present, several differences may exist between children of lower socioeconomic strata families and those of higher SES families even before they enter preschool. Lee and Burkhan (2002) have shown that, even before entering preschool, children in the highest SES group may have cognitive scores above those of students in the lowest SES group. Therefore, delaying these children's entrance into preschool indicates that children are deemed to fail before they begin (Siegel & Hanson, 1991). It is argued that keeping children at home instead of sending them to preschool may potentially increase the gap between children of low SES and middle-high SES families (Vecchiotti, 2001; Stipek, 2002). The

evidence also suggests that the children from the lower socioeconomic strata and the disadvantaged families benefit more from preschool (Cornelissen et al., 2018; Felfe & Lalive, 2018; Havnes & Mogstad, 2015, 2011; Blanden et al., 2014). Havnes & Mgstad (2015, 2011) found that the effects of child care expansion were positive in the lower and middle parts of the earnings distribution. Therefore, providing an early start to the disadvantaged children could help reduce child development and human capital formation inequalities.

On the contrary, the alternative view is a maturational point of view that expects the child to mature and be ready for school. The proponents of the maturational view believed in what is popularly known as ‘academic redshirting or ‘holding back’. They suggest delaying children entering preschool who are not ready for school, and such delay gives the child an extra year to become developmentally ready. They argue that school readiness requires a threshold level of development the child should achieve before starting school. Reaching a specific age does not ensure that a child is ready for school nor guarantees a specific level of development. The conventional wisdom is that older children are more likely to have the necessary skills and maturity to succeed in school and learn more in each grade (Lubortsky & Kaestner, 2016; Krauerz, 2005; Graue & DiPema, 2000; Cmic & Lamberty, 1994). With this belief, parents send their children to preschool deliberately late with the assumption that as the child ages, he or she will mature sufficiently to handle schoolwork. That age alone should not be an adequate basis for determining the time of school entrance.

The questions discussed in the literature in this context are: whether younger children in preschool benefit lesser than their older peers? Moreover, would delay the entry into preschool of the youngest children in their cohort results in better or improved academic results? (Stipek, 2002). Advocates of the maturational view claimed that older children are better until they achieve the prerequisite level of development needed to succeed in school. This "gift of time" gives children an extra year to become developmentally ready for the

formal classroom structure and instruction (Grau & DiPema, 2000). An important implication of this interpretation is that age-related differences in educational outcomes tend to persist or grow as children progress through school. There exists substantial evidence that shows that older children tend to perform better compared to their younger peers within a specific grade or class. Such as, children who start preschool late tend to get relatively higher test scores. (Bedard & Dhuey, 2006; Datar, 2006; Dobkin & Ferreira, 2007).

Research on this entry age debate in preschool in developing countries, especially in the Indian context, is nonexistent. India has the world's largest preschool-aged (under six years) population (around 114 million as per census 2011) and the world's most extensive preschool programme (Integrated Child Development Services). Nevertheless, no systematic study contributes to the preschool entry debate. Based on a primary survey of 1369 households in the Indian state of West Bengal, the study examines if the child's age of entry into preschool impacts children's cognitive and socioemotional development at a later stage. The outcome variables are the indices of cognitive and socioemotional development. These indices were generated using principal component analysis of several cognitive and socioemotional development responses. Using an Instrumental Variable regression after Heckman correction or Heckman-IV regression (Wooldridge, 2010; Kumar, 2016), we found that delaying the preschool entry decreases the skill accumulation of children in grade 1 in primary school. Hence our study significantly contributes to the sparse body of literature on the preschool entry age debate in the context of the global South. It advocates for early entry in preschool as it improves children's social and cognitive skills, particularly in developing countries like India.

With this introduction, the following sections are as follows: Section 2 will provide a background and context of preschool in the Indian context. Section 3 will introduce data and descriptive statistics. Section 4 will explain our identification strategy, and section 5

will present the study's findings. Finally, section 6 will discuss the findings and will conclude the study.

Preschool in the Indian context:

The Indian case is unique compared with the rest of the world. Preschool education in India is available through mainly two channels: Public and private preschools. The public provisions are known as Anganwadi centres and belong to one of the most extensive child-development programmes, the Integrated Child Development Scheme (ICDS), which has been operating since 1985. Children aged 0–6 years are eligible to enroll in these public preschools and receive supplementary nutrition and non-formal education, free of any financial cost to parents. Besides, unregulated and privately operated preschools exist in different parts of the country (for an overview, see Ghosh, 2019; Ghosh & Dey, 2020; Rao et al., 2021; Prochner, 2002; Rana and Sen, 2008).

In most developed and developing countries, the entry age in any preschool is around 36 months. Children continue in preschool until they get admission to primary school based on a school entry cut-off. Admission to primary schools in developed countries is regulated. Children born between 01 April and 31 August can usually start at the beginning of September if they turn five years (Like in the UK and the Netherlands) or six years in most countries (like Germany, France, Denmark, Sweden, and the USA).

Unlike other countries (mainly developed countries), the incidence of preschool enrollment of children in India is very much parents' decision. It primarily depends on their socio-economic-cultural background and motivation, typically unobserved. In India, if enrolled, children usually continue their preprimary experience until they reach six years of age. Once a child turns six years, and if the parents want to send the child to primary school, then the primary school is bound to admit that child whatever that month. According to the Right to Education Act in India, any child can get admission to primary school at any time of

the year, irrespective of the actual start of an academic session, as soon as the child becomes six years old. Article 21-A (since 2002) in the Constitution of India to provide free and compulsory education for all children in the age group of six to fourteen years as a Fundamental Right in such a manner as the state may, by law, determine. The Right of Children to Free and Compulsory Education (RTE) Act, 2009 (which represents the consequential legislation envisaged under Article 21-A) means that every child has a right to full-time elementary education of satisfactory and equitable quality in a formal school which satisfies certain essential norms and standards. Among many other clauses, this act says that every child in India between 6 and 14 years must be provided with access to quality education in a formal school whenever they come. This aspect of India's Right to Education Act is called the Universal Access to Education.

This unique Indian case implies our study and, more specifically, finding the exact identification strategy to get the impact of entry age in preschool on later development. While we are interested in finding the impact of entry age in preschool on a child's later development, we are referring to the impact of the duration spent in preschool. This duration in preschool depends precisely on the age a child enters the preschool and at what age the child leaves the preschool. In developed countries, the preschool entry age is around 36 months, depending on the child's birth date. Although the beginning of elementary school sessions differs from region to region in each country, the school entry age regulation is followed in most cases. For example, in Germany, in some regions, schools begin in September each year whereas in some regions it is August. Therefore, in any specific year, children six-year or older on the entry month are eligible to enter a school that year. The same is true for the US, where school beginning varies between September and December each year. Hence, in these countries, the duration spent in preschool becomes random and exogenous, and that would only depend on the date of birth. Therefore, a simple OLS

specification would be appropriate to see the impact of variation of `duration spent in preschool` on a child's later development, and arguably this variation would be completely exogenous. However, in the Indian context, every child can enter primary school at any time of the year as soon as they age six (under the Right to Education Act). Therefore the duration spent in preschool exclusively depends on at what age a child enters preschool. As in India, preschool education is not compulsory; the decision to send a child into a preschool centre depends entirely on parents' decisions which in turn depends on parents' socioeconomic status, such as their income, level of education, and ethnic identities. Therefore, especially in the Indian context, duration in preschool depends on at what age a child enters preschool. Therefore, to see the impact of preschool on a child's later development, we need to look at the entry age in the preschool.

Moreover, this entry age depends on parents' and household-specific observed and unobserved characteristics (like parents' motivation), which could also affect the child's later development, which would be our dependent variable. Thus our primary explanatory variable, namely the `entry age` in preschool endogenous. In the identification section, we have discussed more on this endogenous explanatory variable and how to get an unbiased causal estimate of the impact of preschool on a child's later development. Moreover, we will also discuss the potential sample selection issue and its way out in dealing with the selected sample of children who have attended preschool from a random sample of children.

Methodology:

This cross-sectional study employs rigorous empirical estimation to answer the key research question: Does the age of entry to preschool influence children's accumulation of cognitive and social skills in primary school? Studies so far failed to reach any convergence in the early vs late preschool entry debate. Besides, no such evidence from India shows relative advantages or disadvantages of late preschool entry. However, considering the developing

nature of the country and prevailing socioeconomic inequalities, there is no certainty that most children will get a developmental stimulus environment at home (Sigurdsen et al., 2011; Belsky et al., 2007; Engle et al., 2007). Therefore, there is a risk of further perpetuating inequalities among disadvantaged children if they are kept home instead of sent to preschool (Vecchiotti, 2001). Hence, in the Indian context, the study hypothesises that children benefit from early preschool entry rather than staying home until they enter primary school. As the study is nonexperimental, the study's findings showed association rather than claiming causality.

Data:

This study is based on a primary survey of 1369 households collected from the state of West Bengal, located in the eastern part of India. The choice of the state was based on pragmatic reasons. The entire sampling process is divided into four stages: (1) Selection of districts, (2) selection of sub-districts, (3) selection of schools, and (4) selection of children. In step two, two of the districts in West Bengal named: Howrah and Murshidabad were randomly chosen for the study through multi-stage and purposive sampling. Available secondary information provided by the state administration showed that these two districts consisted of over 3000 primary schools, including mainly publicly sponsored schools, publicly sponsored religious schools known as Madrasa, and private schools. At the third step of sampling, 70 primary schools were randomly selected from the list provided, and 20 children in grade 1 in each of these selected schools were sampled using a random number generator. However, during the field visit, it was found that in some schools, there were less than 20 children in grade 1. Therefore, some other schools were selected to compensate for this shortage of children. Finally, 84 primary schools were selected to sample about 1400 children for the study. Eventually, 1369 households out of 1400 were located during the household interview and these households and children were then included in the final sample. The unit of our analysis was children in Grade-I in primary schools. The data shows a

great deal of variation concerning preschool attendance of children and the socioeconomic status of their households. Out of the 1369 children selected primarily, 904 children had attended preschool, 645 children had attended public preschool (Anganwadi), and the rest 259 had attended private preschools.

Variables of the study:

***Cognitive and Social Skills (Dependent variables)*:** The study's outcome variables are children's cognitive and social skills during grade 1 in primary school. Children's cognitive and social skills accumulation was measured using the twelve indicators shown in Table 1. These indicators were chosen based on the context and the type of curriculum practised in Indian preschools. These indicators were adopted from the national-level longitudinal study in Germany called National Educational Panel Study (NEPS, see Blossfeld & Roßbach, 2019), starting cohort 2: Kindergarten. However, considering the preschool context in the Indian context, these indicators were modified to fit the context. Unlike studies that primarily used the three R's (Read, write, and arithmetic) or standardised tests for evaluating children's cognitive and social skills, this study focuses more on cognitive and socio-emotional development attributes. Using standardised test modules to assess children's skills may lead to difficulties as these modules are often designed for high-income countries (Rao et al., 2019). Hence, these may not be suitable for the Indian early childhood education context. Therefore, we focus on relatively general concepts of cognitive skills such as attention, and working memory, which are also considered essential for children's academic performance worldwide. Studies showed that children who can regulate attention may more easily engage in classroom instruction, focus on assessments, and complete assignments (Rudasill et al., 2010; Rothbart & Jones, 1998). Besides, working memory (Miller-Cotto & Byrnes, 2020; Peng et al., 2020) and children's ability to comprehend ideas (Baumann, 2010) significantly predict children's academic performance. Furthermore, children's ability to integrate with others, emotional self-control, and socially acceptable behaviour is also considered critical social skills that predict school performance (Baumeister and Vohs, 2004; Boekaerts & Corno, 2005; Zimmerman, 2000; Brownell and Kopp, 2007).

Table 1: Indicators for child development used in the study

Cognitive skills	• Attention	Attention toward class activities
	• Spontaneous	Ability to answer spontaneously if asked questions
	• Assignments	Ability to deliver if given an assignment in class
	• Memory	Ability to recall previous lessons
	• Own ideas	Ability to apply their own ideas
	• Assessment	How the child performed in the last class assessment
Social skills	• Friendliness	Ability to make friends
	• Share	Share food and other items with peers
	• Group activities	Participate in group activities with other peers
	• Help peers	Volunteer to help peers if needed
	• Control temper	Control temper in conflicting situations
	• Compromise	Agree to compromise in conflicting situations

During the fieldwork, a respective class teacher in grade 1 in each school was requested to evaluate all sample children in the classroom based on these twelve indicators Likert scale of 1 to 5 (very bad to very good) was used for this evaluation. Therefore, each of the sample children in grade 1 received a value between 1 and 5 for all twelve cognitive and social skills indicators, respectively. This skill measure took place once for each sampled child in grade 1. The evaluation took place towards the end of the academic session. Therefore, the teacher spent ample time with the children, observing them and knowing their behaviour which in turn helped correctly evaluate them for the study. In total, there were 84 teachers involved in this evaluation in all 84 elementary schools selected for this study. These schools comprised publicly sponsored schools, publicly sponsored religious schools known as Madrasa, and private schools. Summary statistics of all these 12 indicators are reported in Appendix 1. Principal Component Analysis (PCA) was used to conduct exploratory factor analysis separately for the indicators of cognitive skills and social skills to generate latent variables *'cognitive skill' and 'social skill*'. Table 2 presents the summary statistics of all the variables used in the regression analysis.

***Preschool entry age (Independent variable)*:** The variable of interest in this study is the preschool entry age measured in months. During the household visit, parents were asked questions about their preschool experience, including a child's age of entry into preschool. The entry age was converted to months for ease of analysis. The descriptive statistics in Table 2 show that the average age of entry to the preschool was slightly above forty months. In contrast, a good number of children entered at a much earlier age, bringing a great degree of heterogeneity at the entry age. Though there is a great degree of variation within and across public and private preschools, the average entry age does not vary across the type of preschool.

***Control Variables**:* Drawing from the literature and considering the Indian context, several control variables were introduced at each level of the regression analysis. Monthly household income was included as an indicator of the household's economic status. Furthermore, the parents' education levels and their occupational status were included as an indicator of their social status. In the Indian context, social stratification based on caste and religion is utterly vital as significant variation exists in the socioeconomic status of people belonging to different social strata (Census of India, 2011). The caste system is a social hierarchy that originated in ancient India and still exists today. This system defines people's social and economic rights and participation based on their social origin. People from specific castes are often considered subject to various discrimination (Chancel et al., 2021; Oxfam, 2022). People belonging to specific castes (often referred to as 'lower castes') faced oppression for years from their social, economic, and political rights. The government of India has taken affirmative action to provide representation to the historically disadvantaged groups in education, employment, and politics. They adopted inclusive policies by introducing positive discrimination in favour of people belonging to specific castes (SC, ST and OBC). Article 366 (24) and (25) of the Indian Constitution classified caste into four categories, viz. 'General

Caste' Scheduled Caste' (SC), 'Scheduled Tribes' (ST), and 'Other Backward Castes' (OBC). Reservation in India is based on this classification in access to government jobs, educational institutions, and even legislatures to specific sections of the population. Therefore, this study includes caste origin as it may influence children's preschool enrollment and development. The categorical variable for caste is coded into Lower caste [Scheduled Castes (SC), Scheduled Tribes (ST), and Other Backward Caste (OBC)] and Upper castes (General caste). Since the two main religions of the survey region are Hindu and Muslim, the religion dummy has two categories: ' Hindu' and 'Muslim' (the largest group among religious minorities).

Apart from these, the list of control variables includes child characteristics such as age and sex and health status. Furthermore, it also includes district-wise dummies in the models to control for the district to which a child belongs. Other than these, we have also controlled for the 'type of primary school' children are currently attending and the 'type of preschool' they have attended. Finally, the regression specification includes the school-specific class teacher fixed effects. Despite having a fixed guideline and setting the template for internal assessment for school children, there can be school-level heterogeneity via class teacher's assessment, which could influence the test results. Our school-specific class teacher fixed effects in the regression specification will control for that heterogeneity.

Table 2: Summary statistics

Variable Name		**Nature of the variable**	**Mean Value**	**SD.**
PCA Score for Cognitive Skills*1000000		Dependent variable (type -2)	-0.002	2.24
PCA Score for Social Skills*1000000		Dependent variable (type -3)	-0.008	2.18
Entry age (In months)		Independent Variable	40.61	9.52
Religion	1=Hindu	Control Variable	77.65	0.41
	2= Islam and others		22.35	
Household Income (In INR)		Control Variable	6118.40	4220.11
Number of Children in Family		Control Variable	1.752	0.76

Caste	1=Backward casts (SC, ST & OBC together)	Control Variable	29.15	0.45
	2=General Caste		70.85	
Father Edu	1= up to primary	Control Variable	47.34	0.69
	2= up to secondary		40.10	
	3=higher secondary or above.		12.56	
Mother Edu	1= up to primary	Control Variable	42.68	0.66
	2= up to secondary		45.75	
	3=higher secondary or above.		11.57	
Father Job	1=Regular	Control Variable	62.11	0.48
	2= casual or no job.		37.89	
Mother Job	1=Regular	Control Variable	5.05	0.21
	2= casual or no job.		94.95	
Child Sex	1= Male	Control Variable	49.96	0.50
	2= Female		50.04	
Child health	1=Poor	Control Variable	10.52	0.30
	2=Average or good		89.48	
district	1=Howrah	Control Variable	34.55	0.47
	2= Murshidabad		65.45	
Preschool type	1=Public	Control Variable	71.35	0.45
	2=Private		28.65	
Preschool distance	1= within 500 m.	Control Variable	40.54	0.49
	2= more than 500 m.		59.46	
Primary School Type	1=Public	Control Variable	87.66	0.32
	2=Private		12.34	
Source: Authors' calculation based on primary data				

Empirical Strategy:

This paper tries to investigate whether variation in age of entry and hence duration spent in preschool can explain the variation in the child's cognitive and social development after completing one year of primary school education. In our approach, we use entry age at preschool as the primary explanatory variable, which by default has an inverse relation with the duration of stay in preschool as there is a standard age for entry in primary school.

Therefore we worked with a selected sample, namely only those samples of children who had attended preschool. However, considering only the children who had attended preschool would imply that we are dealing with a selected sample (904 out of 1369 children) from our random sample of households. This selection may lead to the classical case of sample selection bias (Heckman, 1979) in our estimates. Therefore, we first estimate the

selection equation (for results from the selection equation, see Appendix 2) and the outcome equation with sample selection correction by Heckman methods (Van de Ven and Van Pragg, 1981). In the Heckman model, the *Selection Equation* is a probit regression to see the factors affecting the probability of attending preschool. The *Outcome equation* estimates the impact of the age of entry into preschool on children's cognitive and social skills for those children who had attended preschool.

Furthermore, there could be many observed and unobserved household-specific and parent-specific factors that could simultaneously influence the entry age at preschool and a child's developmental outcome at the primary level. For example, parents' educational and economic status is often associated with school enrollment and child development (Jonsson & Erikson, 2000; Schober and Spiess, 2013). It has been found that educated parents are more likely to send their children to preschool (Ghosh, 2019) and also provide better 'home education' (Ceka & Murati, 2016). Furthermore, economically well-off parents can afford more toys, book, and other learning materials at home which eventually helps children develop faster. There is evidence that household income has a positive effect on children's cognitive and social development as well as preschool enrollment (Cooper & Stewart, 2021; Magnuson & Waldfogel, 2016). Household income has causal effects on 'intermediate outcomes' that are important for children's development, including maternal mental health, parenting and the home environment (Cooper & Stewart, 2021).

Therefore, the primary explanatory variable in the outcome equation, i.e. *entry age* in preschool, could be endogenous, and simple OLS estimation in the outcome equation would result in biased estimates. Therefore, even if the study begins with a simple OLS specification to estimate the outcome equation, eventually, we propose an instrumental variable regression after Heckman correction or Heckman-IV regression (Wooldridge, 2010; Kumar, 2016). For this purpose, we used a heteroskedastic-based instrumental variable

regression as done in Lewbel's method for our preferred estimation method for the outcome equation. This technique allows the identification of structural parameters in the regression model with endogenous regressors without traditional identifying information. In this form of Lewbel's method, instruments would be constructed as simple functions of the model's data (Lewbel, 2012; Baum and Schaffer, 2012). Econometrically, the identification strategy is as follows:

$$y_i^{attended\_preschool} = \gamma q_i + u_{1i} \dots \dots \dots \dots \dots (1) \text{ (Selection Equation)}$$

$$y_i^{skill} = \beta_1 x_{1i} + \beta x + u_{2i} \dots \dots \dots \dots (2) \text{ (Structural form of the outcome equation)}$$

Furthermore, considering 'x_{1i}' as endogenous, the outcome equation (2) can be disentangled as follows:

$$x_{1i} = \alpha_1 z_i + \alpha x + u_{3i} \dots \dots \dots \dots \dots \dots \dots \dots (2.1) \text{ (First Stage of Outcome Equation)}$$

$$y_i^{skill} = \beta_1 \hat{x}_{1i} + \beta x + u_{4i} \dots \dots \dots (2.2) \text{ (Second Stage of Outcome Equation)}$$

Where $u_{1i} \sim N(0,1)$; $u_{3i} \sim N(0,1)$; $u_{4i} \sim N(0,1)$; $corr(u_1 u_2) = \rho$; $Cov(x_1, u_{2i}^2) \neq 0, i = 1,2$ and $Cov(z, (u_{21} u_{22})) = 0, i = 1,2$. When $\rho \neq 0$, standard probit estimations using only the outcome equation, taking only the children who attended preschool, would yield biased and inconsistent estimates. Hence, probit regression with sample selection is applied, following the two-step Heckman method.

In the selection equation (1), q_i is the vector of independent variables affecting the probability of sending children to preschool of the i^{th} household, γ is the vector of coefficients of independent variables, and u_{1i} are the error terms. N (0, 1) represents the standard normal distribution of the error terms. In the first stage, we estimate a probit model of $y_i^{attended\_preschool}$ on q_i and obtain the estimate $\hat{\gamma}$. Then compute the Inverse Mills Ratio

(imr), $\widehat{\alpha}i = \alpha(q_i\,\hat{\gamma}) = \varphi(q_i\,\hat{\gamma})/\emptyset(q_i\,\hat{\gamma})$ [it is the ratio between the standard regular pdf and the standard normal cdf] for those with $y_i^{attended\_preschool} = 1$.

In the outcome equation (2), x_{1i} is the entry age of the i[th] child, affecting the child's cognitive and social skills $separately$ β_1 is the coefficient of x_{1i}, x and β are the vectors of independent variables and their coefficients, respectively, and u_{2i} is the error term. In the second step, using the selected sample, i.e. we observe only the outcome y_i^{skill} if $y_i^{attended\_preschool} = 1$ or $\gamma q_i + u_{1i} > 0$ (Wooldrige 2006, 618-620). This procedure will give an estimator $\hat{\beta}_1$, which is consistent and normally distributed. The usual t-test was followed to test the selection bias on the coefficient on 'imr,' i.e. coefficient on $\widehat{\alpha}$ as a test of $H_0 = \rho = 0$. However, this $\hat{\beta}_1$ would be unbiased only if x_1 (the entry age in preschool) is exogenous.

However, we have already discussed that x_1 could arguably be endogenous as there could be unobserved factors that could influence x_1, and can also directly influence the outcome variable y_i^{skills}. Therefore, to estimate the unbiased and consistent estimator of x_1 to see the effect of x_1(i.e. entry age in preschool) on outcome variable y_i^{skills}, we instrumented x_1 by a suitable instrumental variable (IV) z_i in equation (2.1). Here z is the heteroscedasticity-based instrumental variable (Lewbel, 2012; Baum and Schaffer, 2012) which is constructed under the following conditions. First, $Cov(x_1, u_{2i}^2) \neq 0, i = 1,2$ and second, $Cov(z, (u_{21}u_{22})) = 0, i = 1,2$. Here the instrumental variable z in equation (2.1) is a subset of regressors or can be assumed as a specific function of regressors in equation (2), like $z=(x_i - \bar{x}) * u_{3i}$ (Lewbel, 2012). Consequently, $\hat{x}_{1i}$ is the estimated value of endogenous x_{1i} from equation (2.1). Eventually, in equation (2.2), we used this $\hat{x}_{1i}$ instead of x_{1i} to see the impact of different preschool entry-age on the developmental outcome of the child at the primary level. Therefore our estimation strategy eventually addressed two econometric issues,

namely sample selection biased and endogeneity, simultaneously by adopting IV regression with Heckman Sample Selection corrections or, in short, Heckman-IV regression (Wooldridge, 2010; Kumar, 2016)

In the sample selection correction model, x is a strict subset of q in equation (1) (Wooldridge, 2006, 618-620). Otherwise, the functional form identifies the model, and the coefficients have no structural interpretations (Cameron & Trivedi 2009). In this case, the exclusion variable chosen is '*distance of nearest preschool*', which is assumed to be exogenous to households. However, in general, critics may argue that due to endogenous sorting across geographic regions of both people and schools, this '*distance'* variable might fail to satisfy the homogeneity restriction of the exclusion variable. However, we argue that other than major urban conglomerates in India, the household's domicile or residence status does not change for school or preschool choices. Hence the decision to send children to preschool directly depends on the availability and the distance of preschool.

Results:

The detailed findings from the regression-based confirmatory analysis (both OLS and IV analysis for the outcome equation) are presented in Table 3. It also shows the goodness of fit of the model as well as the validity of the instrument variables by the F-statistics of each stage of the estimation.

Table 3: Effect of preschool *entry age* on child's later development: Results from Outcome equation

	Outcome equation with OLS estimation		**Outcome equation with IV estimation**	
	Cognitive	Social	Cognitive	Social
Entry Age	-0.00151	0.00376	-0.0451*	-0.0714***
	(0.00997)	(0.00875)	(0.0212)	(0.0212)
Household Income			-6.45e-06	-1.23e-05
	0.0000158	-0.00000801	(1.87e-05)	(1.85e-05)
Religious Origin (Ref. Hindu)	(0.0000202)	(0.0000177)		
Muslim and others			-0.754*	-0.237

	-0.749*	-0.663*	(0.330)	(0.326)
Caste (Ref. General Caste)	(0.372)	(0.325)		
Other Castes (SC, ST OBC)			-0.504**	-0.418**
	-0.328	-0.177	(0.178)	(0.176)
Father's Education (Ref. Up to Primary)	(0.174)	(0.152)		
Secondary			0.376*	0.630***
	0.164	0.320	(0.186)	(0.184)
Higher secondary or above	(0.190)	(0.166)	0.962***	1.220***
			(0.252)	(0.249)
Mother's Education (Ref. Up to Primary)	0.557*	0.471*		
Secondary	(0.264)	(0.231)	0.728**	0.221
			(0.243)	(0.240)
Higher secondary or above	0.811***	-0.141	1.046**	0.870**
	(0.233)	(0.204)	(0.339)	(0.336)
Father Employment (Ref. Regular job)				
Casual or no job	1.000**	0.135	0.140	-0.0359
	(0.323)	(0.283)	(0.223)	(0.220)
Mother Employment (Ref. Regular job)				
Casual or no job	0.0846	0.0419	0.0840	0.0173
	(0.237)	(0.208)	(0.317)	(0.313)
Child Age	-0.000409	0.0227*	0.0223	0.0337**
	(-0.03)	(2.05)	(1.74)	(2.66)
Sex of the child (Ref. Male)				
Female child	-0.104	-0.0670	0.0188	0.143
	(0.322)	(0.282)	(0.134)	(0.133)
Child's health status (Ref. Average)				
Good health	0.151	0.177	-0.247	-0.00443
	(0.132)	(0.115)	(0.358)	(0.354)
Residing District (Ref. Howrah)				
Murshidabad	0.368	0.0717	-0.447	-0.629
	(0.384)	(0.336)	(0.513)	(0.507)
Preschool type (Ref. Public)				
Private Preschool	0.845	2.066*	0.327	0.177
	(0.998)	(0.874)	(0.206)	(0.204)
Primary School Type (Ref. Public)				
Private school	0.441	0.411	-0.0682	-0.469*
	(0.283)	(0.248)	(0.220)	(0.218)
IMR	-0.134	-0.210	-0.342	0.399
	(0.857)	(0.750)	(0.838)	(0.828)

School-specific class teacher-fixed effects (for 84 primary schools)	Yes	Yes	Yes	Yes
Constant	-1.455 (-1.042)	-1.405 (0.913)	1.960* (-1.078)	2.647** (-1.075)
Observations	894	893	894	893
Standard Errors in parenthesis.*** p<0.001, ** p<0.01, * p<0.05 Standard errors are also clustered at the school level.				

Here columns 1 and 2 represent the results from the outcome equation with OLS estimation, and column 3 and 4 shows the results from the same outcome equation with IV estimation. As expected, the outcome equation estimated with OLS resulted in insignificant coefficients for entry age on the outcome variable. We suspect this is due to the endogeneity of the primary explanatory variable, entry age. Once this endogeneity issue is addressed via an IV regression for the outcome equation, we find significant coefficients for the entry age on the outcome variable. Findings from the IV analysis show that preschool entry age negatively affects the child's cognitive and social skill accumulation. On average, with one month increase in the entry age in preschool, there is a 0.04 point decrease in the cognitive skills index and a 0.07 point reduction in social skills. Thus children who enter preschool at a relatively higher age and hence spend less duration in preschool possess lower social and cognitive development at the primary school level compared with children who entered preschool at an earlier age and spend relatively more time. Results also imply that entry age has the most potent effect on the accumulation of social skills among all skills considered.

Moreover, the analysis shows that control variables representing households' socioeconomic status influence children's development outcomes significantly. Children from the marginalised section of the society, that is, children belonging to the minority religion and lower castes, were found to have a significantly lower score in both outcomes. For example, a child with lower caste background (that is 'scheduled caste', 'scheduled tribe', or 'other backward caste') exhibits a 0.50 point less score in the cognitive index and a 0.40 point less

score on the social skill index compared to a child from the ‘general caste’ background. Parents’ education seems to have an important implication for children's later development. Children with parents with a higher education level scored better concerning all the outcome variables compared with low-educated parents. There was also variation in children's development based on the type of primary school that children were currently attending. Attending private school was associated with a 0.46 point lesser score in social skills.

Discussion:

The study's empirical findings indicate that the greater the duration spent in preschool, the better the child performs at a later stage in terms of cognitive and social development. The effect was more substantial for social skills because more time spent in preschool with peers allows them to integrate socially. Attending preschool allows children to interact, communicate, and integrate with other children, eventually stimulating bonding among children in school and enhancing learning (Howes et al., 1998; Ladd, Price, & Hart, l988; Coolahan et al., 2000; Konold & Pianta, 2005).

Furthermore, different school types may have a differential effect on child development. Private preschools seem to be particularly lagging with children's development of social skills. In a recent study, private preschools were found not to contribute to socio-emotional development because of their more rigorous pedagogical approach, which does not significantly change the extent of social interaction (Dean & Jayachandran, 2020). Another study shows that children who attended private preschools have significantly higher cognitive achievements than those in public preschools in India (Singh and Mukherjee, 2018).

Moreover, a considerable variation in child development can be attributed to the families' socioeconomic status, with parents' education having the most considerable impact. The results confirm that pre-existing preschool inequality (if any) might perpetuate over time

if children's entry to preschool is delayed. On the contrary, early entry and thereby spending more time in the preschool may counter the pre-existing inequality and, thus, result in a better outcome in the latter part of their life.

Conclusion:

Our paper tries to shed light on the debate on the impact of *early* versus *late* entry in preschool on the child's later cognitive and social development in primary school. We presented theoretical and empirical literature in respect of this debate across different countries with different socioeconomic backgrounds and including the context of India. We paid particular attention to constructing the identification strategy in our empirical analysis considering the influence of the Right to Education Act of India, which arguably made 'duration spent in preschool'. Hence our primary explanatory variable, 'preschool entry age', is endogenous. Estimating the impact of preschool entry age on a child's later development was challenging due to the possible *sample selection bias* and *endogeneity issue*s. We followed a unique identification strategy like Heckman-IV, where we simultaneously addressed sample selection and endogeneity issues. Our empirical analysis shows that the 'age of entry into preschool' is a significant determinant for children's later development at the primary school stage. We find that if the age of entry in preschool increases (and hence duration spent in preschool reduces), then that negatively affects the child's development at the primary school level. Therefore to ensure better development (both social and cognitive) of the child, parents should send their children to preschool without any delay in preschool entry. Given the socioeconomic inequalities and different forms of discrimination against disadvantaged children, preschool experiences can act as "levelling the field" for the future. As already mentioned, in a developing country where pre-existing socioeconomic inequalities prevail, deferred entry to preschool can only perpetuate these existing inequalities. India, being a typical example of a developing country context, our findings also advocates for early entry to

preschool. These findings contribute to the sparse body of developing country literature and hold an alternative view against the deferred entry into preschool. In the Indian context, public preschool, namely the Anganwadi centre under the ICDS scheme, represents the predominant preschool facility free of any financial burden to parents. Therefore, this study suggests a renewed campaign favouring the ICDS programme and more preschool enrollment to ensure a child's social and cognitive development.

Reference:

Angrist, J. D. and Krueger, A. B. 1991. "Does Compulsory School Attendance Affect Schooling & Earning?" *The Quarterly Journal of Economics* CVI(4): 979-1014.

Baum, C. F. and Schaffer, M. E. 2012. "IVREG2H: Stata module to perform instrumental variables estimation using heteroscedasticity-based instruments". Statistical Software Components S457555, Boston College Department of Economics., revised 07 February 2019.

Baumann, J. F. (2010). Children's ability to comprehend main ideas in content textbooks. *Reading World*, 22(4), 322-331.

Baumeister, R. and Vohs, K . (2004). *Handbook of Self-Regulation: Research, Theory, and Applications*. Guilford Press. New York, USA.

Bedard, K. and Dhuey, E. 2004. "The Persistence of Early Maturity: International Evidence of Long-Run Age Effects." *UC Irvine Economics Colloquium.* Available at < https://escholarship.org/uc/item/9c01w01k>.

Belsky, J., Bell, B., Bradley, R.H., Stallard, N., Stewart-Brown, S.L. (2007). Socioeconomic risks, parenting during the preschool years and child health age 6 years. *European Journal of Public Health*, 17(5): 508-513.

Bickel D. D., N. Zigmond, and Strayhorn, J. 1991. “Chronological Age at Entrance to First Grade: Effects on Elementary School Success.” *Early Childhood Research Quarterly* 6(2).

Blanden, J., Del Bono, E., Hansen, K., McNally, S., and Rabe, B. 2014. “Evaluating a demand-side approach to expanding free preschool education.” Retrieved from <https://www.nuffieldfoundation.org/sites/default/files/files/Childoutcomes_final.pdf.>

Blossfeld, H.-P. & Roßbach, H.-G. (Eds.). (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS). Edition ZfE* (2nd ed.). Springer VS.

Boekaerts, M., and Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review*, 54, 199–231.

Bracey, G. 1989. “Age and Achievement.” *Phi Delta Kappan* 70(9): 732.

Brent, D., May, D. and Kundert, D. 1996. “The Incidence of Delayed School Entry: A Twelve-Year Review.” *Early Education Development* 7(2): 121-135.

Brownell, C. and Kopp, C. (eds) (2007). *Socioemotional Development in the Toddler Years*. The Guilford Press: New York.

Cameron, A. C. and Trivedi, P. 2009. *Microeconometrics: Methods and applications*. New York: Cambridge University Press.

Ceka, A. and Murati, R. (2016). The role of parents in the education of children. *Journal of Education and Practice*, 7(5).

Chancel, L., Piketty, T., Saez, E. and Zucuman, G. (2021). World inequality report 2022. World Inequality Lab. UNDP.

Cmic, K. and Larnberty, G. 1994. “Reconsidering school readiness.” *Early Education and Development* 5(2): 91- 105.

Coolahan, K., Fantuzzo, J., Mendez, J., and McDermott, P. (2000). Preschool peer interactions and readiness to learn: Relationships between classroom peer play and learning behaviors and conduct. *Journal of Educational Psychology,* 92: 458–465.

Cooper, K. and Stewart, K. (2021). Does household income affect children's outcomes? A systematic review of the evidence. *Child Indicators Research,* 14, 981-1005.

Cornelissen, T. and Dustmann, C. 2019. "Early school exposure, test scores, and non-cognitive outcomes." *American Economic Journal: Economic Policy* 11(2): 35–63.

Cornelissen, T., Dustmann, C., Raute, A., and Uta Schönberg, U. 2018. "Who benefits from universal child care? Estimating marginal returns to early child care attendance." *Journal of Political Economy* 126 (6): 2356–2409.

Crosser S. L. 1991. "Summer birth date children: Kindergarten entrance age and academic achievement." *Journal of Educational Research,* 84(3):140-146.

Cunha, F. and Heckman. J. 2007. "The technology of Skill Formation." *American Economic Review* 97 (2): 31–47.

Cunha, F., Heckman, J., Lochner, L. and Masterov, D. V. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education,* eds. Hanushek, E. and Welch, F. Amsterdam: Elsevier.

Currie, J. and Almond, D. (2011). Human capital development before age five. In Card, D. and Ashenfelter, O *(eds). Handbook of Labor Economics*. Volume 4B. Elsevier.

Darren Lubotsky, D. and Kaestner, R. 2016. Do 'Skills Beget Skills'? Evidence on the effect of kindergarten entrance age on the evolution of cognitive and non-cognitive skill gaps in childhood." *Economics of Education Review,* 53: 194-206.

Datar, A. 2006. Does delaying kindergarten entrance give children a head start? *Economics of Education Review, 25:* 43-62.

Dean, J. T., and Jayachandran, S. (2020). Attending kindergarten improves cognitive development in India, but all kindergartens are not equal. Working Paper. Retrieved from https://faculty.wcas.northwestern.edu/~sjv340/kindergarten.pdf

Dobkin, C. and Ferreira, F. V. (2007). Do school entry laws affect educational attainment and labor market outcomes?" Available at <http://dx.doi.org/10.2139/ssrn.1007492>.

Elder, T. E. and Lubotsky, D. H. (2009). Kindergarten entrance age and children's achievement- Impacts of state policies, family background and peers. *The Journal of Human Resources,* 44(3): 641-683.

Engle, P.L., Black, M.M., Behrman, J.R., … Young, M.E. (2007). Strategies to avoid the loss of development potential in more than 200 million children in the developing world. *The Lancet,* 369 (9557): 229-242.

Felfe, C. and Lalive, R. (2018). Does early child care affect children's development?" *Journal of Public Economics*, 159: 33–53.

Ghosh, S (2019). Inequalities in demand and access to early childhood education in India. *International Journal of Early Childhood*, 51: 145–161.

Ghosh, S. and Dey, S. (2020). Public or private? Determinants of parents' preschool choice in India. *International Journal of Child care and Education Policy,* 14(3).

Graue, E. and DiPerna, J. (2000). Redshirting and early retention: Who gets the gift of time and what are its outcomes?" *American Educational Research Journal,* 37(2): 509-534.

Havnes, T. and Mogstad, M. (2011). No child left behind: Subsidised child care and children's long-run outcomes. *American Economic Journal: Economic Policy,* 3(2).

Havnes, T. and Mogstad, M. (2015). Is universal child care leveling the playing field? *Journal of Public Economics*, 127, 100–114.

Heckman, J. J. (2008). Schools, Skills, and Synapses. *Economic Inquiry,* 46 (3): 289–324.

Howes, C., Hamilton, C. E., and Philipsen, L. C. (1998). Stability and continuity of child-caregiver and child-peer relationships. *Child Development*, 69, 418–426.

Jones, M. M. and Mandeville, G. K. (1990). The effect of age at school entry on reading achievement scores among South Carolina students. *Remedial and Special Education*, 11(2): 56-62.

Jonsson, J. O., & Erikson, R. (2000). Understanding educational inequality: The Swedish experience L'Annee sociologique (1940/1948) Troisieme serie, Vol. 50(2), 345–382. Paris: Presses Universitaires de France.

Kagan, S. L. (1990). Readiness past, present and future: Shaping the agenda. *Young Children*, 48(1): 48-53.

Katz, L. (2000). Academic redshirting and young children. *ERIC*. Washington, DC, Office of Education Research and Improvement.

Konold, T. R., & Pianta, R. C. (2005). Empirically-derived, person-oriented patterns of school readiness in typically-developing children: Description and prediction to first-grade achievement. *Applied Developmental Science*, 9: 174–187.

Krauerz, K. (2005). Straddling early learning and early elementary school. *Journal of the National Association for the Education of Young Children*, 64(3): 50-58.

Kumar, A. (2016). Lifecycle-consistent female labor supply with nonlinear taxes: evidence from unobserved effects panel data models with censoring, selection and endogeneity. *Review of Economics of Household*, 14(1) (March): 207–229.

Ladd, G. W., Price, J. M., and Hart, C. H. (1988). Predicting preschoolers' peer status from their playground behaviors. *Child Development*, 59: 986–992.

Lee, V. E. and Burkham, D. (2002). Inequality at the starting gate: Social sociodemographic differences in achievement as children begin school. Washington, DC: Economic Policy Institute.

Lewbel, A. (2012). Using heteroskedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business and Economic Statistics,* 30: 67–80.

Lincoln, J. and Painter, G. (2006). Does the Age that Children Start Kindergarten Matter? Evidence of Long-Term Educational and Social Outcomes. *Educational Evaluation and Policy Analysis*, 28(2): 153-179.

Magnuson, K. and Waldfogel, J. (2016). Trends in income-related gaps in enrollment in early childhood education: 1968 to 2013. *AERA Open.* Available at *https://doi.org/10.1177/2332858416648933.*

Miller-Cotto, D., & Byrnes, J. P. (2020). What's the best way to characterise the relationship between working memory and achievement? An initial examination of competing theories. *Journal of Educational Psychology*, 112(5): 1074-1084.

Oxfam (January, 2022). Inequality kills: India supplement. New Delhi.

Peng, P. and Kievit, R. A. (2020). The Development of Academic Achievement and Cognitive Abilities: A Bidirectional Perspective. *Child Development Perspectives*.

Prochner, L (2002) Preschool and playway in India. *Childhood*, 9(4): 435–453.

Rana, K. and Sen, S. (2008) The ICDS programme in West Bengal: Scopes and challenges, Pratichi (India) Trust. Presented at the Regional Consultation on the Status of the Young Child on 12th & 13th March 2008 at Ranchi, Jharkhand.

Rao, N, Ranganathan, N, Kaur, R, et al. (2021). Fostering equitable access to quality preschool education in India: Challenges and opportunities. *International Journal of Child Care and Education Policy*, 15(1): 1–22.

Rao, N., Richards, B., Sun, J., Weber, A., and Sincovich, A. (2019). Early childhood education and child development in four countries in East Asia and the Pacific. *Early Childhood Research Quarterly*, 47:169–181.

Rothbart, M. K., and Jones, L. B. (1998). Temperament, self-regulation, and education. *The School Psychology Review*, 27, 479–491.

Rudasill, K. M., Gallagher, K.C., and White, J. M. (2010). Temperamental attention and activity, classroom emotional support, and academic achievement in third grade. *Journal of School Psychology*, 48: 113-134.

Schober, P. S., & Spiess, C. K. (2013). Early childhood education activities and care arrangements of disadvantaged children in Germany. *Child Indicators Research*, 6(1): 709–735.

Sigurdsen, P., Berger, P. and Heymann, J. (2011). The effects of economic crises on families caring for children: Understanding and reducing long-term consequences. *Development Policy Review*, 29(5): 547─564.

Singh, R., & Mukherjee, P. (2018). Effect of preschool education on cognitive achievement and subjective wellbeing at age 12: evidence from India. *Compare: A Journal of Comparative and International Education*.

Stipek, D. (2002). At what age should children enter kindergarten? A question for policymakers and parents. *SRCD Social Policy Report* 16(2): 3-16.

Stipek, D. (2009). School entry age. *Encyclopedia of Early Childhood Development* 2nd Ed. Stanford University School of Education, USA.

Van De Ven, W. P. M. M. and Praag, V. B. (1981). The demand for deductibles in private health insurance: A probit model with sample selection. *Journal of Econometrics*, 17(2): 229–252.

Vecchiotti, S. (2001). Kindergarten: The Overlooked School Years. *The Foundation for Child Development Working Paper Series.*

Wooldridge, J. M. (2006). *Introductory Econometrics- A Modern Approach*. South-Western CENGAGE Learning.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Zimmerman, B. J. (2002). Achieving self-regulation: The trial and triumph of adolescence. In F. Pajares & T. Urdan (Eds.), *Academic motivation of adolescents*, Vol. 2, pp. 1–27. Greenwich, CT: Information Age.

Appendices:

Appendix 1: descriptive statistics of the development indicators

Variable	Obs	Mean	SD.	Min-Max
Attention	1369	3.76	0.85	1 - 5
Spontaneous	1369	3.66	0.90	1 - 5
Assignment	1369	3.58	0.93	1 - 5
Memory	1369	3.52	0.93	1 - 5
Own Ideas	1368	3.48	0.93	1 - 5
Assessment	1369	3.74	0.92	1 - 5
Make friends	1369	3.91	0.74	1 - 5
Share	1369	3.83	0.78	1 - 5
Help Peers	1369	3.83	0.80	1 - 5
Group Activities	1369	3.81	0.81	1 - 5
Control Temper	1368	3.78	0.77	1 - 5
Compromise	1369	3.81	0.77	1 - 5

Source: Authors' calculation from primary data

Appendix 2: Estimation of the selection equation (i.e. equation 1)

Variables	Coefficients
Distance of the Preschool (Ref: more than 500 m.)	
Within 500 meters	0.192*
	(0.083)
Household Income	0.000
	(0.000)
Total Child in household	-0.005
	(0.057)
Religion (Ref: Hindu)	
Islam and others	0.655***
	(0.110)
Caste (Ref: General)	
Other Castes (S.C, S.T. & OBC)	0.112
	(0.095)
Father's Education Level (Ref. up to primary)	
Secondary	-0.202
	(0.103)
Higher Secondary and above	0.313
	(0.219)
Mother's Education Level (Ref. up to primary)	
Secondary	0.445***
	(0.101)
Higher Secondary and above	0.642**
	(0.214)
Father's Employment Status (Ref. Regular job)	
Casual or no job	-0.296**
	(0.090)
Mother's Employment Status (Ref. Casual/no job)	
Regular Job	0.211
	(0.213)
Sex of the Child (Ref. Male)	
Female Child	0.094
	(0.081)
Health Status of the child (Ref. below Average)	
Average or good	0.606***
	(0.130)
Residing District (Ref. Howrah)	
Murshidabad	-1.278***
	(0.120)
School specific class teacher-fixed effects (for 84 primary schools)	Yes
Constant	0.127
	(0.44)
Observation	1351
Pseudo R2	0.253

Standard Errors in parenthesis. ***p<0.001, **p<0.01, *p<0.05