Stratified Society and Segmented Schools: The Effect on Children's Learning

Abhijit Visaria* and Minseop Kim^

- Correspondence to Abhijit Visaria, Graduate Group in Demography, Population Studies Center, University of Pennsylvania, 216 McNeil Building, 3718 Locust Walk, Philadelphia, PA 19104. E-mail: avisaria@sas.upenn.edu

^ School of Social Policy and Practice, University of Pennsylvania.

As India continues to grow in global importance as an emerging economy, there is increasing focus on India's continued poor rankings on human development indicators, especially the quality of education. Elementary education remains one of the fundamental pillars of a society's development, and there is a growing concern that gains made in the past two decades, measured in terms of improved geographical access to primary schools and higher primary school enrollments, are not translating into higher learning achievement for a large majority of India's children, especially in rural areas and smaller towns.

Existing literature on elementary education in India has sought to explain historical trends in low levels of literacy and low enrollment levels, and current trends in high dropout rates, as the effect of social stratification along gender, caste and class lines manifested in limited access for a large population. (The PROBE Team, 1999; Aggarwal, 2001; Ramachandran, 2002). Society in India has historically been stratified according to the '*Varna'* or the caste system, which involved religious, social and economic norms for people belonging to different castes. The caste system has been used to justify exploitation and the withholding of economic opportunities from 'lower castes'. A person's caste identity continues to play an important role in his/her life, from opportunities for interaction with neighbors, treatment by school teachers, to access to common property resources and employment opportunities. In large parts of India, an individual's accumulation of human, social and cultural capital is influenced very significantly by his/her caste identity. Although schools in principle 'equalize' the experience for children belonging to different social strata, children belonging to scheduled castes or tribes are likely to experience schooling differently due to overt or subtle discrimination from teachers or their peers, as well as due to the effects of disadvantaged access to learning resources and environments.

There are few studies however that explore the relationship between the institutional characteristics of schools and educational outcomes. Research has noted that there is a rising perception among parents that the quality of education in private schools is better than in government schools (Kingdon, 1996, De et al, 2002). This has been accompanied by a steady rise in the number of private schools – in both urban and rural areas – that despite their fee-charging structure and higher overall costs, attract parents from better-off as well as poor households. It is not clear however whether private schools are indeed imparting a better quality of education than government school. Additionally, another characteristic of schools that is a potential predictor of learning achievement is the medium of instruction. With a sharp rise in the proportion of service-industry employment relative to manufacturing or agriculture (which has seen a decline in the past three decades), the importance of English as the language of communication across sectors and different parts of the country has risen. One indicator of this is that during the 1990s, most state governments speeded up the introduction of English as a second language, from Grade 5 to Grade 1, in schools where Hindi or other regional languages were the medium of instruction. On the other hand, an increasing number of parents, despite their own lack of fluency with English, now prefer to send their children to schools where English is the medium of instruction, for English is seen to open up learning opportunities and employment prospects (National Council for Educational Research & Training, 2006). The question however remains if children learn better if they are taught in their native language (Hindi or a regional language) compared to English.

Another major recent development in Indian schools has been the introduction of meals in schools. The Government of India started the National Program of Nutritional Support to

Primary Education in 1995, commonly known as the Mid-day Meal Scheme (MDM), and re-launched it in 2004 based on findings that fewer children remained away from schools or drop-out of schools if they were given meals at school, and that attention spans and learning levels were severely impeded if children were hungry (summarized in Dreze & Goyal, 2003). The MDM scheme is addressed both at improving the nutritional status as well as learning achievements of children in primary schools. This paper takes advantage of the data on the availability of mid-day meals at schools to see if there is indeed an effect on the learning achievement levels of children.

DATA

This study uses data from the India Human Development Survey (IHDS), conducted jointly in 2004-05 by the University of Maryland, College Park and the National Council of Applied Economic Research (NCAER), New Delhi, and funded by the National Institutes of Health (Desai et al, 2007). The survey included detailed data pertaining to demographic characteristics of households as well as short tests in reading fielded to children in the age group of 8-11 years. The sample was selected using a clustered sampling procedure designed to provide a nationally representative sample of India, and the survey was conducted with 41,554 households across 1503 villages and 971 urban locations of India, covering all 28 states and 5 of 7 union territories.

Dependent Variable

The IHDS includes learning achievement assessments in reading, writing and mathematics. For the purposes of conducting preliminary analysis for this paper, we restrict our analysis to reading achievement.

Reading skill tests in the IHDS were designed using a multi-point scale, which corresponds to the common teaching curriculum for the main language taught in Grades 1 and 2 in schools in India. Children are classified according to their ability to read in one of five ordered categories: (1) Cannot read at all; (2) Can read letters but not form words; (3) Can put letters together to read words but not read whole sentences; (4) Can read a short paragraph for 2-3 sentences but not fluent enough to read a whole page; (5) Can read a one page short story. Given the diversity in languages and dialects in India often within a state as well, there have been few standardized tests fielded across India that measure reading abilities for young children. The assessment included in the IHDS is unique in that it was developed in partnership with *Pratham*, a large non-government organization, which had field-tested the tool during its direct implementation of elementary education programs in 21 different states. The words included in categories (3), (4) and (5) were either from, or equivalent in level, to the official Grade 1 and 2 textbooks published by the education board of the state where the child was assessed. The sentences in category (4) pertained to the reading level expected of children who have successfully cleared Grade 1, and the short story that of Grade 2 children. The IHDS interviewers were trained by *Pratham* staff in building a rapport with children, helping them overcome their shyness about being assessed and in conducting the test.

Independent Variables

This analysis employs a set of school-related independent variables as well as demographic

and household background characteristics. India is a country of immense social and cultural diversity and includes a number of caste divisions in every religion with unique social and cultural histories. Caste is a critical indicator of social stratification especially in rural areas, and scheduled caste and tribal groups have historically been and continue to be marginalized and discriminated against in material terms by "higher" or "forward castes". This ranges from exploitation in physical and economic terms, limited or lack of access to pubic resources, to poor or indifferent behavior from public officials including teachers. Children in scheduled caste and scheduled tribe households are not less likely to be enrolled in schools relative to other children, but previous research has indicated that they are more likely to drop out of school at different stages (Ramachandran, 2002).

The IHDS collected data on the highest educational level of mothers and fathers in the household, which we have included in our analysis. The educational level of the mother is of interest, since they remain the primary caregivers of children and are in a position to influence the learning of children by way of revision, preparation for exams, interest in and interaction with school activities, participation in parent-teacher meetings, etc. Although fathers are seen less as care-givers, the association of their educational achievement with their children's learning may be interesting to examine and compare with that of the women in the household. This paper also factors in residence-location, in asking the question whether an urban/rural divide exists in India whereby children in urban India perform better than children in rural India. We also control for distance from the school in our analysis, using distance as a continuous variable, to see if there is any effect on learning of the distance that children travel which may be associated with their attendance and the time they spend studying.

Descriptive statistics from the data are presented in Table 1 below.

Table 1: Descriptive Statistics

Variables	Obs.	Mean/Proportion	SD
Age	17,061	9.48	1.08
Gender	17,061		
Male	8,940	52.4%	
Female	8,121	47.6%	
Education of Father	16,333	6.34	4.94
Education of Mother	16,912	3.83	4.60
Log income	17,061	10.21	1.69
Caste	14,111		
Brahmin & High castes	3,356	23.73%	
OBC	5,723	40.47%	
Scheduled caste	3,729	26.37%	
Scheduled tribe	1,333	9.43%	
Region	17,061		
Rural	12,040	70.57%	
Urban	5,021	29.43%	
Distance	15,466	1.54	2.01
School Type	15,466		
Government School	10,308	66.65%	

Private School	5,158	33.35%
Medium	15,440	
Hindi & State Lang	12,972	84.02%
English	1,873	12.13%
Other	595	3.85%
Mid-day Meal	15,306	
No	7,125	46.55%
Yes	8,181	53.45%
Reading	12,356	
Cannot read	1,128	9.13%
Letter	1,611	13.04%
Word	2,491	20.16%
Paragraph	2,703	21.88%
Story	4,423	35.80%

METHODS

Our analysis employs the continuation ratio model. Unlike other models such as cumulative logit or adjacent categories that may be used to study ordered categorical variables, the continuation ratio model can account for the ordered categories of learning achievement that represent an advancement of stages, such that the individual must pass the preliminary stages in order to get to higher stages. In this study, categories for reading (cannot read, can read letters, can read words, can read a paragraph, and can read story) clearly indicate a progression of stages.

In addition, other models might be inappropriate for these data for statistical and theoretical reasons. For example, the cumulative logit model is valid as long as the proportional odds assumption is not violated. However, given the sample size of these data, it is expected that the proportional odds assumption will be violated. In the case of the adjacent model that allows only one set of coefficients for every adjacent pair, it is unreasonable to believe that the coefficient of predictors is the same for all category pairs. Finally, although the stereotype model can deal with the problem of the adjacent model by directly estimating *phi* (Long and Freese, 2006), this model is also problematic in that the *p*-values are not invariant to the choice of the base category or the corner constraint.

In the continuation ratio model, the logit scale of the outcome is a linear and additive function of predictors. Maximum likelihood (ML) method is used to estimate the odds among subjects at level *j* of having an outcome greater than level *j* relative to being at level *j* (McGowan, 2000). If we assume that the categories on dependent variables are ordered in the sequence *j*=1….J, and if we define A_{ij} as the probability that child *i* at stage *j* advances to stage *j* + 1, the continuation ratio model can be specified as (Allison, 1999):

$$\log \frac{A_{ij}}{1 - A_{ij}} = \alpha_j + \boldsymbol{\beta X_i} \quad j = 1,2, \ldots., J - 1$$

$$\text{where } A_{ij} = \Pr(y_i > j \mid y_i \geq j) \text{ and } \boldsymbol{\beta X_i} = \beta_1 x_{i1} + \cdots + \beta_k x_{ik}$$

In order to test restrictions of the continuation ratio model that the effects of the predictors are

the same at each stage, we include interaction terms between our predictors and different levels of our dependent variable. If the interactions have significant effects, we will further estimate separate logistic models for each stage.

Finally, in this study, we use Huber-White standard error in order to address clustering of data. Given the data structure and survey design in which individuals (i.e. children) are nested within a family and a school, if we do not account for clustering of data, standard errors may be underestimated, resulting in Type I error. Huber-White standard errors will therefore account standard error underestimation.

RESULTS

Table 2 presents the odds-ratio results of the continuation ratio model, using advancement to the next stage of reading as the dependent variable. In Model 1, we find significant results for our demographic and background predictors in the model except for sex. In terms of school characteristics, school-type is not statistically significant. As one would expect, for every one-year increase in the age of the child, the child is 37% more likely to advance to the next stage of the reading achievement level. Caste categories have a statistically significant result in the expected direction. With reference to Brahmins and other higher castes, children of other backward classes are 16% less likely to advance to the next stage of the reading achievement level. The disadvantage of scheduled caste and scheduled tribe children is greater, with a 24% less likelihood of advancement to the next stage of the reading achievement level.

Distance and mid-day meals have an unexpected and yet statistically significant result. For every one-unit increase in the distance between the child's residence and school, the odds of advancement to the next stage of the reading achievement are greater by 5%. These results indicate that children who are given mid-day meals in schools have 22% lower odds of advancement to the next stage of the reading achievement.

Conceptually, the medium of instruction is more relevant for math compared to reading, given that reading achievement was measured in the same language as the medium of the school. We intend to extend this draft paper further using the math learning achievement data from the IHDS. Interestingly, we do find that the medium of instruction does have a significant result on reading achievement – children who are taught in schools where the medium of instruction is English are 12% less likely, compared to their counterparts being taught in Hindi or the state language to advance to the next stage of the reading achievement level.

Table 2: Results of the Continuation Ratio Model

	Model 1 (N=42458)	Model 2 (N=42458)
Variables	Odds Ratio	Odds Ratio
Age	1.37***	1.55***
Sex (Ref. = Male)	.97	.82*
Education of Father	1.06***	1.04***
Education of Mother	1.02***	1.12***
Log income	1.02*	1.06**

Caste (Ref. = Brahmin & high caste)		
OBC	.84***	.57***
Scheduled caste	.76***	.61***
Scheduled tribe	.76***	.56***
Region (Ref. = rural)	1.08	1.41**
Distance	1.05***	1.22**
School Type (Ref. = Government School)	1.09	1.87***
Medium (Ref. = Hindi & State Lang)		
English	.88*	1.47
Other	.89	5.53*
Mid-day Meal	.78***	.96
Reading Level (Ref. = Level 1)		
Level 2	.53***	4.63**
Level 3	.26***	.85
Level 4	.15***	4.62**
Interaction Terms with Reading Levels		
Age x Level 2		.91*
Age x Level 3		.94
Age x Level 4		.76***
Gender x Level 2		1.08
Gender x Level 3		1.30**
Gender x Level 4		1.23*
Education of Father x Level 2		.98
Education of Father x Level 3		.97*
Education of Father x Level 4		.96**
Education of Mother x Level 2		.99
Education of Mother x Level 3		.95**
Education of Mother x Level 4		.91***
Log income x Level 2		.92**
Log income x Level 3		.97
Log income x Level 4		.97
Caste1 x Level 2		1.26
Caste2 x Level 2		.95
Caste3 x Level 2		1.33
Caste1 x level3		1.49*
Caste2 x level3		1.22
Caste3 x level3		1.27
Caste1 x level4		1.72***
Caste2 x level4		1.54**
Caste3 x level4		1.47
Region x level2		.83
Region x level3		.89
Region x level4		.61***
Distance x level2		.85**
Distance x level3		.86**
Distance x level4		.84**
School type x Level2		.45***
School type x Level3		.58***
School type x Level4		.58***

Medium1 x Level2		.76
Medium1 x Level3		.58
Medium1 x Level4		.55
Medium2 x Level 2		.16**
Medium2 x Level 3		.14**
Medium2x Level 4		.12**
Meal x Level 2		.78
Meal x Level 3		.75*
Meal x Level 4		.82
Log likelihood	-15523.12	-15297.81
Pseudo R2	.10	.11

* p< .05, ** p < .01, and *** p<.001

The continuation ratio model assumes that the effects of the predictors are the same at each level. However, since it is possible for the effects of predictors to vary across levels, we test this assumption by including interactions between the level of learning and other predictors. The results of Model 2 indicate that there are statistically significant interactions with almost all predictors, which implies that each variable has different effects on advancement to the next stages at each level. In order to understand the effects of predictors at each stage more clearly, we now estimate separate logistic models for each stage. These results are presented in Table 3.

Table 3: Separate Logistic models for Reading Achievement Levels

	Advancement from "cannot read" to next stages (N=12237)	Advancement from "letter" to next stages (N=11493)	Advancement from "word" to next stages (N=10302)	Advancement from "paragraph" to "story" stages (N=8426)
Variables	Odds Ratio	Odds Ratio	Odds Ratio	Odds Ratio
Age	1.55***	1.41***	1.46***	1.18***
Gender (Ref.= male)	.82*	.88	1.07	1.01
Education of Father	1.05***	1.03**	1.02**	1.00
Education of Mother	1.12***	1.11***	1.07***	1.02**
Log income	1.06**	.98	1.02	1.03
Caste (Ref. = Brahmin & high caste)				

OBC	.57***	.71***	.85*	.98
Scheduled caste	.61***	.58***	.74***	.94
Scheduled tribe	.56***	.75*	.72**	.83
Region (Ref. = rural)	1.41**	1.17	1.25**	.86*
Distance	1.22**	1.03	1.05*	1.03*
School Type (Ref. = Government School)	1.87***	.85	1.08	1.09
Medium (Ref. = Hindi & St Lang)				
English	1.47	1.11	.86	.81*
Other	5.53*	.89	.80	.67
Mid-day Meal	.96	.75***	.72***	.79***
Log likelihood	-2472.51	-3557.07	-4557.16	-4711.06
Pseudo R2	.12	.07	.07	.02

* p< .05, ** p < .01, and *** p<.001

The results of the separate logistic models are useful in nuancing the results of the continuation ratio by helping us "view the process in terms of stages" (Allison, 1999:158). We see that age continues to have a significant effect at every stage of reading – with every one-year increase in age, there are 55% higher odds of advancement into a higher category of reading achievement compared to an inability to read. However, the effect of age is smaller at the last stage- with every one-year increase in age, there are 18% higher odds of advancement from reading a paragraph to reading a story. While all children who were administered the reading assessments are expected according to the school curriculum to be able to read short-stories, the continued positive effect of age is likely indicative of the fact that older children are able to 'revise' their earlier lessons in readings and have more opportunities to pick up reading skills. Interestingly, gender, family income and school-type have a significant effect only at the basic level - i.e girls are 18% more likely than boys to not be able to advance to the next stage from inability to read. However, they are not disadvantaged at different levels of reading itself. For every one-unit increase in the log family income, a child has 6% greater odds of being able to read (at any level) compared to not being to read. The effect of income does not continue beyond the first stage, and children in families with higher income are not more likely to advance to a higher reading level.

Along with school-type, where children in private schools are more likely to be able to read compared to not read, these results appear to be signaling the presence of a threshold effect. Being male, income and school-type have a positive effect only in being able to read compared to not read, whereas once the child is able to read at least letters, they do not further the reading achievement of the child. The medium of instruction at school does not have a significant effect except that children taught in English have 19% lower odds of advancing from the paragraph-reading level to being able to read a story. The sample for children being taught in 'other' languages is very small and the 'other' category is undefined in the data for any meaningful effect to be analyzed. Distance from school continues to have an unexpected result with children who travel longer distances to school more likely to be able to read (greater odds of 22%), be able to read sentences, paragraphs, and story compared to just words (5% more likely) and read short stories compared to paragraphs (3%).

We find significant results for parental education. Mother's education matters more than

father's education, although both mother's and father's education has a decreasing effect at higher reading levels. For every one-year increase in mother's education, the odds of a child being able to read (compared to not being to read) are 12%, and 5% greater for a one-year increase in father's education. Subsequently, children are 7% more likely in advancement from being able read a word to being able to read a paragraph or short story for every one-year increase in mother's education, and 2% for father's education.

DISCUSSION

The results of our study indicate that relative to sociodemographic factors which have a persistent effect at different levels of reading achievement, school characteristics such as school type, medium of instruction and mid-day meals explain fewer differences in reading achievement.

Our study shows that there is a clear advantage that children receive from being in private schools. Children in private schools are more likely to be able to read than children in government schools. However there are no further differences in reading achievement levels that can be attributed to school-type once the child is able to read. This indicates that private schools are significantly more successful at teaching children to read, likely due to a greater investment at early levels of a child's enrollment. However, other structural factors not related to school-type are responsible for differences at higher levels of reading.

The results for mid-day meals in schools are surprising and merit further investigation. Whereas children in schools where mid-day meals are served are not more likely to be able to read, they are less likely to be at higher reading levels. This is likely to be a function of the phenomenon where schools in distant, remote or poorer geographical areas have been a focus of the government policy for the implementation of mid-day meals. As a result, these schools may have continuing lower learning levels for their children because of other resource constraints, with mid-day meals achieving an effect that is as yet not still lower than other schools.

The positive association of distance to school with reading ability and higher reading achievement appears to be implying that distance to school is not a close proxy indicator for school availability. The greater odds of being able to read, and the much lower yet statistically significant odds of being at higher reading levels may be suggesting two things about distance: more-motivated or economically better-off children may be choosing to travel longer distances to travel to schools perceived to be better, and/or families that send their children longer distances to school may also be more invested in their children's education. We do not see meaningful effects of the medium for instruction, suggesting that the language that the child is taught in does not matter for their reading achievement.

Socio-demographic characteristics of households, especially the social group category, matter significantly in explaining differential levels of reading. Children belonging to lower caste households are at a disadvantage and consistently less likely than their higher caste counterparts to demonstrate reading skills or be at higher reading levels. This indicates that although children from disadvantaged social groups may be attending schools in larger numbers than before as national enrolment data show, they are not learning as well as many

of their peers. An urban advantage exists in education, and is present at both the level of any-reading ability and being able to read whole paragraphs or stories compared to only words. Children in urban areas are more likely to interact with a larger group of educated peers and adults, and have greater opportunities to participate in environments and situations that encourage reading in particular, indicating that children in rural areas are faced with multiple relative disadvantages in stimulating formal educational achievement.

By controlling for the age of the child, we expect to account for differences that may exist due to children of different ages being at the same reading level in school. However, in order to fully account for the effect of the child's grade on his/her learning achievement, we intend to include in our final analysis a control for the child's grade in school.

The results of this analysis also draw our attention to the utility of the continuation ratio model to the study of indicators with ordered categories, and seeing if meaningful effects of relevant predictor variables exist at different levels of a dependent variable such as reading.

REFERENCES

Aggarwal, Yash. 2001. *Progress towards Universal Access and Retention: Analytical Report*. New Delhi: National Institute of Educational Planning and Administration.

Allison, Paul D. 1999. *Logistic Regression Using SAS: Theory and Application*. Cary, N.C.: SAS Institute.

De, Anuradha, Manabi Majumdar, Claire Noronha & Meera Samson. 2002. "Private Schools and Universal Elementary Education", in R. Govinda (ed.), *India Education Report*. New Delhi: Oxford University Press.

Desai, Sonalde, Amaresh Dubey, B.L. Joshi, Mitali Sen, Abusaleh Shariff, and Reeve Vanneman. 2007. *India Human Development Survey (IHDS)*. ICPSR22626-v2. University of Maryland and National Council of Applied Economic Research, New Delhi. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Dreze, Jean & Aparajita Goyal. 2003. "Future of Mid-Day Meals". *Economic & Political Weekly*, 38(44): 4673-4683.

Kingdon, Geeta Gandhi. 1996. *Private Schooling in India: Size, Nature and Equity-Effects*, Development Studies Programme Working Paper No. 72. London: London School of Economics.

Kingdon, Geeta Gandhi, Robert Cassen, Kirsty McNay & Leela Visaria. 2004. "Education and Literacy", in Tim Dyson, Robert Cassen & Leela Visaria (eds.), *Twenty-First Century India: Population, Economy, Human Development, and the Environment*. New York: Oxford University Press.

Long, J. Scott & Jeremy Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata*. College Station: Stata Press.

McGowan M.J. 2000. "Ordinal Outcomes with the Continuation Ratio Model". Paper presented at the Proceedings of the Northeast SAS Users Group, Philadelphia. (*www.nesug.org/proceedings/nesug00/ps/ps7015.pdf)*

National Council for Educational Research and Training (NCERT). 2006. *Position Paper of the National Focus Group on Teaching of English.* New Delhi: NCERT.

The PROBE Team. 1999. *Public Report on Basic Education in India*. New Delhi: Oxford University Press.

Ramachandran, Vimala. 2002. *Gender and Social Equity in Primary Education: Hierarchies of Access*. New Delhi: Sage Publications.