Educational Attainment and Learning in India, 2004-2012

Farzana Afridi
Indian Statistical Institute
Economics and Planning Unit
New Delhi, India
Email: fafridi@isid.ac.in

Bidisha Barooah
International Initiative for Impact Evaluation, 3ie
New Delhi, India
Email: bbarooah@3ieimpact.org

Abstract

The world's population has doubled between 1960 and 2000 and is expected to rise further by more than two billion people by 2050. Asia will not only continue to be home to the largest share of world population, but it will also have the highest ratio of working to non-working population in the world in 2050. In this chapter we focus on one country—India—poised to be the largest individual contributor to the global working-age population of 15–64-year-olds over the coming three decades. The general optimism about the coming surge in working-age population is dampened by the low quality of skills of India's youth which makes it difficult to employ them productively. We analyse the educational attainment of the school-age population of the country—the additions to the future workforce - and highlight the serious concerns about the quality of skills being imparted to students at all levels and the depth of learning occurring in India's educational institutions.

JEL Classification: I21, I25

Keywords: education, quality, demographic dividend, IHDS

I Introduction

The world's population has doubled between 1960 and 2000 and is expected to rise further by more than 2 billion people by 2050 (Bloom, 2011). 97 percent of this coming population boom will occur in the least and less developed countries as their remarkable reductions in mortality rates are followed by sharp reductions in fertility (Bloom, 2011). Rising probability of infants surviving into adulthood along with a lagged decline in the number of new-borns will raise the proportion of young, working age population in most developing countries in this century. Asia will not only continue to be home to the largest share of world population, but it will also have the highest ratio of working (15-64 year olds) to non-working population (65+ years old) in the world in 2050 (Bloom, 2011). This 'demographic dividend' is expected to have a positive impact on economic growth due to several factors including the larger labor force and the higher savings rate of the working-age population relative to dependents (Bloom, Canning and Malaney, 2000).

In this chapter we focus on one country – India - poised to be the largest individual contributor to the global working age population of 15-64 year olds over the coming three decades (Aiyar and Mody, 2011). India's working age population is now 63.4 percent (Census, 2011) of the total population and expected to rise to 69 percent in 2040 (United Nations Population Division, 2014). This presents both an opportunity and a challenge for India. It is an opportunity because it is expected to give an impetus to the economic growth of India and an advantage over other developing economies such as China whose population has begun to age (United Nations Population Division, 2014).[1] Aiyar and Mody (2011) predict a large and significant impact of the rise in the working age ratio on India's economic growth, adding up to 2 percentage points per annum to the country's per capita GDP growth over the next two decades.

This optimism, of course, assumes that the additions to the working age population will be matched by employment creation and that this age-group would be able to meet the market's demand for skills. Aiyer and Mody (2011) highlighting the spatial distribution of the growth in the working age population of India, find that a third of this growth is expected to come from the poorest and least educated states of India, viz. Uttar Pradesh and Bihar. While these states may be poised to benefit the most from any demographic dividend, they also present a challenge to the country's capacity to take advantage of the changing age structure of its population. The general optimism about the coming surge in working age population is,

[1] See Figure A in the appendix.

thus, dampened by the news on two fronts in India: the slow creation of new employment opportunities, particularly in the manufacturing sector, and the low quality of skills of the youth which makes it difficult to employ them productively.[2]

In this chapter we focus on the latter constraint in harnessing India's demographic dividend. Specifically, we analyse the educational attainment of the school-age population of the country – the additions to the future workforce and the surge in the working age population of the country. Expanding and improving the quality of educational attainment remains a key challenge to taking advantage of the demographic dividend in India. The average levels of educational attainment and basic skill acquisition, including reading and writing, remain low by international standards in India (Program for International Student Assessment, PISA, 2009).[3] There exist serious concerns about the skills being imparted to students at all levels and the depth of learning occurring in India's educational institutions.

Participation or enrolment in educational institutions has been steadily expanding in India but acceptable levels of educational attainment by students have remained elusive. Educational interventions in India have been directed mostly towards increasing investments in public education by building schools, improving existing school infrastructure and training teachers. For instance, although primary school enrolment has increased from 79 percent in 2001 to 90 percent in 2007 due to public interventions such as the Sarv Shiksha Abhiyan[4], the quality of public education remains poor as reflected by high drop-out rates and low levels of learning. The drop-out rate for children progressing from grade 1 to 5 was as high as 25 percent in 2005-06 in India (Ministry of Human Resource Development Report, 2005-06). Findings from a nation-wide survey of rural primary schools show that about half of students enrolled in grade 5 cannot read texts meant for second-graders (Pratham, 2009).

In this chapter we highlight findings from the first and only nationally representative, household level panel data – the Indian Human Development Survey (IHDS), 2004-05 and 2011-12. Detailed individual level data on educational attainment were collected in this survey, including data on learning outcomes through the administration of tests of learning

[2] http://www.bloombergview.com/articles/2013-03-14/india-s-economy-leaves-job-growth-in-the-dust

[3] The Programme for International Student Assessment (PISA) is conducted by the OECD every 3 years to test15-year-old school children's performance on mathematics, science, and reading, since 2000.

[4] World Development Indicators for 2011 (http://data.worldbank.org/indicator/SE.PRM.NENR)

skills in reading and maths to 8-11 year olds currently enrolled in schools.[5] Our analysis documents the rise in the rates of participation in schooling between 2004-05 and 2011-12. However, we find that there is an alarming decline in learning levels of school age children across the country. This finding is echoed by smaller, repeated cross-sectional data from ASER which suggest that learning levels are largely stagnating or declining in rural India (ASER 2014).

In addition to the absolute decline in learning levels, our analysis suggests that while the gaps in school participation by gender, social groups and income levels have declined dramatically over the period of the study, the learning gaps have been mostly stagnant. The stimulus to economic growth from the surge in the proportion of working age population is, theoretically, expected to also arise from greater work force participation by women as their fertility levels decline. In addition, rural areas of India are more likely to contribute to the proportion of the young, working age adults since almost 70 percent of the country's population continues to reside there (Census, 2011). Thus the stagnancy in learning gaps suggests that, girls and other socially marginalised groups which have lower initial learning levels have not been able to catch up with the rest of the population in terms of their educational attainment. The concern then arises that these groups may not be able to benefit from or participate fully in economic growth or contribute to the potential demographic dividend.

Further, state and individual level fixed effects analysis suggest that low learning levels in early schooling reduces the probability of a child completing primary education, transitioning to secondary level of schooling and leads to lower overall schooling attainment. Our analysis, therefore, highlights the urgent need to redress the poor quality of educational attainment in India in order for the country to sustain high levels of economic growth in the future.

The remainder of the chapter is organised as follows. We discuss our data in section 2. Section 3 presents the results of the empirical analysis. Conclusions and policy implications are in Section 4.

2. Data

[5] The IHDS also conducted tests of writing skills in both survey rounds. The data, however, are not nuanced (0=cannot write, 1=can write with 1 or 2 mistakes, 2=can write with no mistakes), hence excluded from our analysis. However, we find do find that the proportion of 8-11 year olds who cannot write decreased by 5.63 percentage points between 2004-05 and 2011-12.

We use data from the Indian Human Development Survey (IHDS), a nationally representative survey of 41,554 households in 1503 villages and 971 urban neighbourhoods across India. The first round of the survey interviews was completed in 2004-5 while the second round was conducted in 2011-12. We use data from both survey rounds and use weights to make our numbers nationally representative.

For the purposes of this analysis we restrict our data to 5 to 21 year olds, the age group which is most likely to continue being enrolled in an educational institution either at the primary, secondary or tertiary levels. This age group would be among the significant contributors to the rise in the proportion of working age population of 15-65 year olds in the coming decades in India. The data on learning levels, however, is available only for 8-11 year olds in both rounds since only this age group was administered tests of learning. In Table 1 we summarise the full sample for repeated cross-sections of 5-21 year olds in both survey rounds.

[Table 1 about here]

We have a sample of 80,011 individuals in round 1 (2004-05) as shown in column 1 of Table 1. The mean age is 12.68 years of which more than 49 percent are females. 18 percent of the sample belongs to a 'high' caste, while over 65 percent belong to the relatively disadvantaged 'Other Backward Castes' (OBC) and scheduled caste/tribes (SC/ST) groups. Almost 28 percent of the individuals belonged to below poverty line (BPL) households while a little over 25 percent were residing in an urban area in 2004-05. Of this sample, 85 percent were ever enrolled in an educational institution but only 65 percent are currently enrolled. This suggests very high drop-out rates. The average years of schooling for this age group in 2004-05 were only 4.5 years.

Restricting the sample to the same age group in 2011-12, the demographic composition is broadly comparable with the sample in 2004-05. However, the proportion of individuals belonging to BPL households is lower at 24 percent and the urban residents rise slightly by almost 4 percentage points. More interestingly, the proportion of the population having ever attended an educational institution has increased dramatically to 93 percent and current attendance rates are more than 10 percentage points higher than in 2004-05. However, the gap between the current and ever enrolled fell modestly, suggesting continued high drop-out rates. Years of schooling rise by more than 1 year from 4.6 in 2004-05 to 5.57 years in 2011-12.

3. Empirical Analysis

A. Participation in education

We first present statistics on the participation rates and years of schooling attained by 5-21 year olds in 2004-05 and 2011-12 in India using rounds 1 and 2, respectively, of the IHDS survey data. We classify participation status into two categories – ever and currently enrolled. Ever enrolled figures give us an idea of the proportion of the 5-21 year olds who were either enrolled and dropped out or those who continue to be enrolled. The latter category is accounted for by the currently enrolled proportions. The difference between these two variables gives us a picture of the rate at which the sample is dropping-out of educational institutions.

[Table 2 about here]

In Table 2 we show the proportion of 5-21 year olds who were ever enrolled in an educational institution in 2004-05 and 2011-12. Overall, the ever enrolled rate has risen significantly by 8 percentage points. When we classify the sample into 6 age groups, we find that while the ever enrolled rates have risen across all age groups, the largest increases have been in the youngest age group of 5-7 year olds and in the later age groups of 17-21 year olds. Fewer females are ever enrolled but this gender gap declined significantly by 4.4 percentage points between 2004-05 and 2011-12, driven primarily by the gender gap declines in older age groups. In both 2004-05 and 2011-12 the ever enrolled rate is the lowest for SC/STs. However, the gap between the high castes and SC/STs declined significantly by more than 6 percentage points between 2004-05 and 2011-12. The gap between individuals belonging to below and above poverty line (APL) households fell by more than 3 percentage points. The rural-urban gap also declines significantly by almost 5 percentage points during this period.

These findings also hold for the proportion of currently enrolled population in this age group as shown in Table 3. Current enrolment rose by 11 percentage points, overall, the biggest increases coming from the older age groups of 15-21 year olds. While the gender gap fell by almost 3 percentage points, there was no significant reduction in the current enrolment gap between Hindus and non-Hindus.

Thus, overall the data indicate that the proportion of 5-21 year olds ever and currently enrolled has increased significantly during the period under study and the socio-economic gaps in participation rates have also declined. This suggests that while the participation rate

of all groups has risen, the socio-economically disadvantaged groups have improved their school participation rates significantly more than the advantaged groups.

[Table 3 about here]

How do these participation rates translate into years of schooling? Table 4 shows that there has been a significant increase of more than one year of schooling for the entire sample during this period. This increase has been driven by the large rise in educational attainment of 15-21 year olds in the sample. The gender gap in years of schooling disappeared by 2011-12 due to a lower gender gap in all age groups, and the gap moving in favour of girls in younger age group. There has been a significant decline in the socio-economic gaps in years of schooling attained, except by poverty status.

[Table 4 about here]

Overall, the empirical analysis suggests that participation rates in education have increased significantly during the period under study. This has translated in marginally significantly higher years of schooling. The marginalised socio-economic groups may have exhibited larger gains in participation rates. This suggests that access to education has risen considerably during 2004-05 and 2011-12. However, the concern that remains is that the gap between ever enrolled and currently enrolled – drop-out rates - has not declined as dramatically as the increase we observe in participation rates.

B. Learning

In the next two tables we explore the changes in reading and math skills of 8-11 year olds who were administered tests of learning ability in repeated cross-sections in 2004-05 and 2011-12.

[Table 5 about here]

Each child, who was administered the reading test, received a score between 0 and 4.[6] We create dummy variables for children who are unable to read (score 0), those who can either read letters or words (score 1 or 2) and those who can read either a paragraph or a story (score 3 or 4). The data are summarised in Table 5. The mean score of children in both rounds suggests that, on average, children's skills were between 'being able to read words (score=2) and a paragraph' (score=3). The first row, however, informs us that there has been

[6] 0=cannot read; 1=can recognize letters; 2=can read words; 3=can read a paragraph; 4=can read a story

a significant decline in the reading level of children between 2004-05 and 2011-12, by 0.062 points. Further, the proportion of children who are unable to read has increased significantly by 2 percentage points while the proportion who can read a paragraph or a story has declined significantly by over 2 percentage points.

Decomposing the data by various socio-economic groups gives us another insight – socio-economically disadvantaged groups have lower reading skills in both survey rounds and there has been no significant decline in this gap between 2004-05 and 2011-12. This holds when we look at either the overall scores or the proportion of children who are able to read a paragraph or a story. If anything, for some socio-economic groups we see a worsening of the gap – a higher proportion of SC/ST children are able to only recognize a letter or read a word relative to high castes. This also holds for the below and above poverty line and rural urban gaps as shown by the significant, positive coefficient in column 3 for 'proportion who can read a letter or a word'. We do, however, find a decline in the proportion of females who can read a letter or a word relative the males.

[Table 6 about here]

We present a similar analysis of performance on the math tests in Table 6. The math scores are coded in the range of 0-3.[7] We create dummy variables for whether a child has no math skills, if she can recognize numbers, whether she can recognize numbers and subtract and a dummy for being able to do all of the above and divide. We note that as for reading, math skills are low for the overall population under study, varying between being able to recognize numbers and perform a simple subtraction. More worryingly, there is a significant decline in the math score as well during this period. While there has been an insignificant change in the proportion of children with low math skills, the proportion of children who can subtract has increased by 3 percentage points. However, there has been more than a 5 percentage point decline in the proportion of students who can perform a simple division.

When we analyse the scores by socio-economic gaps, we find that as for reading, the disadvantaged groups perform worse than the advantaged groups in both rounds. There has been no decline in the gap in the overall score except between SC/ST and high caste children. However, the proportion of disadvantaged children (viz. SC/ST, below poverty line and rural households) who are able to perform a simple division has increased significantly in 2011-12, relative to the more advantaged groups. Besides, the socio-economic gaps in learning outcomes, Figures 1-4 show that the overall decline in the learning skills also permeate

[7] 0=no skills; 1=can recognize numbers; 2=can subtract; 3=can divide

various socio-economic groups – an absolute decline in reading and math skills for males and females, either declining or stagnant (except for 'other') for all caste categories, by poverty status and by residence.

[Figures 1-4 about here]

Our analysis suggests that almost all groups have experienced either stagnating low learning skills or even a decline in these skills during 2004-05 and 2011-12. This is in sharp contrast to the results we have from our analysis of participation rates during this period. Our observations square with those from other surveys which administer learning tests, such as ASER, but the IHDS data allow us to present the first comprehensive, nationally representative analysis of the state of educational attainment in India.

We take advantage of the panel nature of the IHDS data to analyse the relationship between a child's score in learning tests in 2004-05 and her educational attainment in 2011-12. Since children who are administered these tests are 8-11 year olds in both rounds, those whose learning scores are reported in 2004-05 would be 15-18 years of age in 2011-12. Hence, ideally they should be continuing their education in 2011-12. We analyse the effect of a child's learning level in 2004-05 on her educational attainment in 2011-12 using individual level panel data. We define our main explanatory variable in two ways – first, 'high reading (math) score' is a dummy variable that takes a value 1 if the child was at least able to read a paragraph (subtract) in 2004-05. Alternatively, 'reading (math) score above median' is a dummy variable that equals 1 if the child's reading (math) score was higher than the median score for her age cohort.

[Table 7 about here]

We first analyse the impact of learning skills during primary schooling on the probability of a child completing primary education (grade 5) and then on transitioning to secondary school in Table 7. We estimate a state fixed effects analysis, to address possible state level variations in the quality of schooling.[8] Further, we control for the child's age, gender and other household level characteristics such as caste, religion, poverty status, parental education and residence. In columns (1) and (2) our dependent variable is whether a child completed grade 5 (primary school) or not. The sample is restricted to children who had completed grade 4 or less in 2004-05. In specification (1) our main coefficients of interests are 'high reading/math score in 2004-05' when the child was between 8 to 11 years old. We find that having a high score increases the probability of completing primary school by 3.9 to

[8] Education policy and expenditures are determined by state governments in India's federal system.

4.8 percentage points. Our results are similar when we use 'reading/math score above median in 2004-05' as shown in column 2. In columns (3) and (4) we analyse whether a child's learning skill in 2004-05 affects the probability of her enrolling in secondary school. The sample is restricted to children who had completed grade 5 or less in 2004-05. Again, we find that a high reading/math score raises the probability of a child continuing schooling at secondary level by 4.8 to 8.9 percentage points as shown on the coefficients on scores in columns (3) and (4). We do not find a consistent effect of age or gender of the child on these outcomes. However, individuals belonging to Hindu households and with educated parents (relative to parents with less than primary schooling) have higher probability of completing primary education and continuing to secondary school. Children from BPL families, however, have worse outcomes relative to those from above poverty line households.
[Table 8 about here]

The links between learning levels in primary school and educational attainment later in life, as shown in Table 7, may be confounded by school, household and child level unobservables. For instance, children who are more able may have better learning skills and may also be more likely to continue their schooling into higher grades. In Table 8, therefore, we report the results of regressing years of completed schooling attained by a child in 2011-12 on her learning scores in 2004-05 in an OLS-child fixed effects model. This accounts for school, household and child level unobserved characteristics that may affect the child's educational outcome. We restrict the sample to those children who were ever enrolled in 2004-05. In column (1) are main explanatory variables are 'high reading (math) score x round 2'. The coefficient on 'Round 2' is positive and significant, indicating higher years of schooling than in 2004-05, as expected. However, the interaction term implies that years of schooling are significantly higher for those who had high learning levels in 2004-05, for both math and reading. These results hold up when are main coefficient of interest is the performance of the child relative to the age cohort median score in column 2. These analyses, thus, suggest that educational attainment is likely to be higher for children with better learning outcomes, even after we account for children's innate ability.

Our findings suggest that low learning levels are not only a concern for the present but are likely to have a significant negative effect on the future educational attainment of a child.

4. Conclusions

We present an analysis of the educational outcomes in India over the period 2004-05 to 2011-12 using the recently released, nationally representative household level panel data from the Indian Human Development Survey. Our chapter highlights the rise in participation in educational institutions during this period but accompanied by alarming declines in learning levels for the primary school going age groups. We document stagnant or declining skills in reading and math across all socio-economic subgroups. Further analysis at the individual level suggests that low learning levels raise the probability of a child dropping out of schooling and having lower total years of schooling.

Our findings hold significance given the rising proportion of working to non-working population of India. India's so-called coming 'demographic dividend' can be realised only if the young population is adequately skilled and educated to be gainfully employed and productive. These findings raise grave concerns about the quality of the human capital of India which in turn can adversely affect the country's economic growth rates.

Our chapter underscores the need to urgently improve the quality of education in India through various policy measures, including increased accountability of public school teachers, better curriculum design and reductions in student-teacher ratios (Muralidharan, 2013).

References

Aiyer, Shekhar and Ashok Mody. 2011. 'The Demographic Dividend: Evidence from the Indian States.' International Monetary Fund, working paper WP/11/38.

Bloom, David. 2011. '7 Billion and Counting.' *Science*, 333(6042):562-569.

Bloom, David, Canning, D and P Malaney. 2000. 'Demographic Change and Economic growth in Asia.' *Population and Development Review*, 26.

Ministry of Human Resource Development Report, Government of India. 2005-06

Muralidharan, Karthik. 2013. 'Priorities for Primary Education Policy in India's 12th Five-year Plan.' *India Policy Forum*.

Registrar General of India. 2011.Census of India.

Programme for International Student Assessment (PISA).

Pratham. 2009. Annual Survey of Education Report

----------2014. Annual Survey of Education Report

United Nations Population Division.
http://www.un.org/en/development/desa/population/publications/development/population-development-database-2014.shtml

Table 1: Summary Statistics

Variables	2004-05	2011-12
	N=80,011	N=67,972
Age	12.680	13.050
	(4.804)	(4.860)
Female	0.491	0.503
	(0.500)	(0.500)
High Caste	0.180	0.171
	(0.384)	(0.376)
OBC	0.350	0.353
	(0.477)	(0.478)
SC/ST	0.305	0.310
	(0.460)	(0.463)
Other minorities	0.165	0.166
	(0.371)	(0.372)
Hindu	0.759	0.754
	(0.428)	(0.431)
Below Poverty Line	0.278	0.241
(BPL)	(0.448)	(0.428)
Urban	0.253	0.293
	(0.435)	(0.455)
Ever enrolled	0.850	0.930
	(0.357)	(0.256)
Currently enrolled	0.648	0.757
	(0.478)	(0.429)
Years of schooling	4.541	5.569
	(3.859)	(4.074)

Note: Missing data for some variables in the sample. Standard deviations in parentheses. Weights used to make the averages nationally representative.

Table 2: Proportion of 5-21 year olds ever enrolled

Characteristic	2004-05	2011-12	Difference
	(1)	(2)	(2)-(1)
All	0.850	0.930	0.080***
			(0.002)
5-7 years	0.722	0.822	0.101***
			(0.007)
8-11 years	0.933	0.980	0.047***
			(0.003)
12-14 years	0.917	0.970	0.053***
			(0.003)
15-17 years	0.869	0.962	0.093***
			(0.006)
17-18 years	0.836	0.942	0.107***
			(0.007)
19-21 years	0.802	0.901	0.099***
			(0.007)
Female	0.816	0.919	0.102***
			(0.004)
Male	0.882	0.940	0.058***
			(0.003)
Female vs. Male	-0.066	-0.022	0.044***
			(0.005)
High Caste	0.932	0.968	0.035***
			(0.003)
OBC	0.863	0.939	0.076***
			(0.004)
SC/ST	0.812	0.909	0.098***
			(0.005)
SC/ST vs. high caste	-0.121	-0.059	0.062***
			(0.006)
Non-Hindu	0.793	0.906	0.113***
			(0.005)
Non-Hindu vs. Hindu	-0.076	-0.032	0.044***
			(0.006)
Below Poverty Line	0.775	0.875	0.100***
			(0.005)
BPL vs. APL	-0.104	-0.072	0.032***
			(0.006)
Rural	0.828	0.918	0.089***
			(0.003)
Rural vs. urban	-0.086	-0.041	0.045***
			(0.004)

Note: Standard errors in parentheses. ***significant at 1%.

Table 3: Proportion of 5-21 year olds currently enrolled

Characteristic	2004-05	2011-12	Difference
	(1)	(2)	(2)-(1)
All	0.648	0.757	0.109***
			(0.003)
5-7 years	0.709	0.818	0.109***
			(0.007)
8-11 years	0.912	0.972	0.060***
			(0.004)
12-14 years	0.813	0.907	0.093***
			(0.005)
15-17 years	0.577	0.772	0.195***
			(0.010)
17-18 years	0.395	0.608	0.213***
			(0.012)
19-21 years	0.395	0.608	0.161***
			(0.008)
Female	0.609	0.732	0.123***
			(0.005)
Male	0.686	0.782	0.096***
			(0.005)
Female vs. Male	-0.077	-0.050	0.027***
			(0.007)
High Caste	0.755	0.840	0.086***
			(0.006)
OBC	0.659	0.786	0.127***
			(0.006)
SC/ST	0.605	0.710	0.105***
			(0.007)
SC/ST vs. high caste	-0.150	-0.130	0.020**
			(0.009)
Non-Hindu	0.578	0.690	0.112***
			(0.007)
Non-Hindu vs. Hindu	-0.093	-0.089	0.004
			(0.008)
Below Poverty Line	0.576	0.701	0.125***
			(0.007)
BPL vs. APL	-0.100	-0.074	0.026***
			(0.008)
Rural	0.626	0.738	0.112***
			(0.004)
Rural vs. urban	-0.088	-0.064	0.024***
			(0.006)

Note: Standard errors in parentheses.***significant at 1%.

Table 4: Years of schooling attained by 5-21 year olds

Characteristic	2004-05	2011-12	Difference
	(1)	(2)	(2)-(1)
All	4.541	5.569	1.029***
			(0.030)
5-7 years	0.684	0.665	0.019***
			(0.160)
8-11 years	2.891	3.205	0.314***
			(0.029)
12-14 years	5.377	6.165	0.079***
			(0.039)
15-17 years	6.604	7.974	1.370***
			(0.064)
17-18 years	7.177	8.872	1.695***
			(0.084)
19-21 years	7.284	8.989	1.705***
			(0.092)
Female	4.334	5.575	1.241***
			(0.041)
Male	4.738	5.564	0.826***
			(0.043)
Female vs. Male	-0.404	0.011	0.416***
			(0.059)
High Caste	5.671	6.675	1.005***
			(0.066)
OBC	4.662	5.644	0.982***
			(0.057)
SC/ST	4.008	5.203	1.195***
			(0.049)
SC/ST vs. high caste	-1.663	-1.472	0.191**
			(0.082)
Non-Hindu	3.926	4.967	1.041***
			(0.050)
Non-Hindu vs. Hindu	-0.811	-0.798	0.013
			(0.061)
Below Poverty Line	3.320	4.086	0.766***
			(0.050)
BPL vs. APL	-1.690	-1.954	-0.263
			(0.061)
Rural	4.176	5.256	1.080***
			(0.037)
Rural vs. urban	-1.443	-1.069	0.374***
			(0.057)

Note: Standard errors in parentheses.***significant at 1%.

Table 5: Reading skills of 8-11 year olds

Characteristic	2004-05	2011-12	Difference
All	(1)	(2)	(2)-(1)
Overall score	2.513	2.452	-0.062**
			(0.027)
Proportion unable to read	0.109	0.129	0.020***
			(0.007)
Proportion who can read a letter or word	0.391	0.402	0.011
			(0.009)
Proportion who can read a paragraph or story	0.543	0.521	-0.022**
			(0.009)
Female vs. Male			
Overall score	-0.150	-0.066	0.084
			(0.054)
Proportion who can read a letter or word	0.038	0.010	-0.029**
			(0.014)
Proportion who can read a paragraph or story	-0.034	-0.026	0.009
			(0.019)
SC/ST vs. high caste			
Overall score	-0.742	-0.748	-0.006
			(0.067)
Proportion who can read a letter or word	0.093	0.123	0.030*
			(0.017)
Proportion who can read a paragraph or story	-0.263	-0.244	0.018
			(0.025)
Non-Hindu vs. Hindu			
Overall score	-0.232	-0.237	-0.005
			(0.056)
Proportion who can read a letter or word	0.027	0.033	0.007
			(0.015)
Proportion who can read a paragraph or story	-0.082	-0.080	0.002
			(0.020)
BPL vs. APL			
Overall score	-0.603	-0.640	-0.037
			(0.060)
Proportion who can read a letter or word	0.083	0.135	0.051***
			(0.018)
Proportion who can read a paragraph or story	-0.205	-0.190	0.015
			(0.020)
Rural vs. Urban			
Overall score	-0.500	-0.509	-0.009
			(0.046)
Proportion who can read a letter or word	0.072	0.091	0.018*
			(0.011)
Proportion who can read a paragraph or story	-0.175	-0.177	0.002
			(0.017)

Note: Reading score is in the range of 0-4. Standard errors in parentheses. Significant at *10%, **5% and ***1%.

Table 6: Math skills of 8-11 year olds

Characteristic	2004-05	2011-12	Difference
	(1)	(2)	(2)-(1)
Overall score	1.510	1.445	-0.064***
			(0.019)
Proportion with no math skill	0.189	0.176	-0.013
			(0.008)
Proportion who can ONLY recognize numbers	0.332	0.369	0.036
			(0.009)
Proportion who can recognize numbers and subtract	0.258	0.289	0.030***
			(0.008)
Proportion who can recognize numbers, subtract and divide	0.220	0.167	-0.054***
			(0.007)
Female vs. Male			
Overall score	-0.187	-0.130	0.057
			(0.038)
Proportion with no math skill	0.077	0.043	-0.034**
			(0.016)
Proportion who can recognize numbers, subtract and divide	-0.049	-0.037	0.012
			(0.014)
SC/ST vs. high caste			
Overall score	-0.630	-0.532	0.097**
			(0.049)
Proportion with no math skill	0.159	0.152	-0.008
			(0.019)
Proportion who can recognize numbers, subtract and divide	-0.218	-0.129	0.089***
			(0.020)
Non-Hindu vs. Hindu			
Overall score	-0.183	-0.200	-0.017
			(0.040)
Proportion with no math skill	0.045	0.035	0.010
			(0.017)
Proportion who can recognize numbers, subtract and divide	-0.069	-0.073	0.005
			(0.014)
BPL vs. APL			
Overall score	-0.485	-0.472	0.014
			(0.041)
Proportion with no math skill	0.144	0.155	0.011
			(0.019)
Proportion who can recognize numbers, subtract and divide	-0.131	-0.096	0.036***
			(0.014)
Rural vs. Urban			
Overall score	-0.414	-0.416	-0.002
			(0.033)
Proportion with no math skill	0.106	0.129	0.024
			(0.013)
Proportion who can recognize numbers, subtract and divide	-0.108	-0.074	0.034**
			(0.014)

Note: Math score is in the range of 0-3. Standard errors in parentheses. Significant at *10%, **5% and ***1%.

Table 7: Impact of learning scores in 2004-05 on probability of completing primary school and transitioning to secondary school by 2011-12 (OLS-FE)

	Completed primary school		Transitioned to secondary school	
Variables	(1)	(2)	(3)	(4)
Child's age	-0.005	-0.001	-0.007**	0.002
	[0.003]	[0.003]	[0.003]	[0.003]
Female child	0.008	0.008	0.008	0.008
	[0.006]	[0.006]	[0.007]	[0.007]
High reading score in 2004-05	0.039***		0.065***	
	[0.006]		[0.007]	
High math score in 2004-05	0.048***		0.089***	
	[0.007]		[0.008]	
Reading score above median in 2004-05 in		0.047***		0.069***
		[0.005]		[0.006]
Math score above median in 2004-05		0.026***		0.048***
		[0.005]		[0.006]
SC/ST household	0.007	0.005	0.008	0.004
	[0.007]	[0.007]	[0.008]	[0.008]
Hindu	0.061***	0.063***	0.079***	0.082***
	[0.009]	[0.009]	[0.010]	[0.010]
BPL	-0.020**	-0.022***	-0.034***	-0.037***
	[0.008]	[0.008]	[0.009]	[0.010]
Father's education				
Primary	0.035***	0.037***	0.074***	0.077***
	[0.011]	[0.011]	[0.012]	[0.012]
Secondary	0.069***	0.072***	0.114***	0.121***
	[0.009]	[0.009]	[0.010]	[0.011]
College	0.065***	0.070***	0.105***	0.114***
	[0.010]	[0.010]	[0.013]	[0.013]
Mother's education				
Primary	0.041***	0.043***	0.050***	0.056***
	[0.007]	[0.007]	[0.009]	[0.009]
Secondary	0.032***	0.034***	0.038***	0.046***
	[0.007]	[0.007]	[0.008]	[0.008]
College	0.026***	0.026***	0.043***	0.048***
	[0.009]	[0.009]	[0.011]	[0.011]
Urban	-0.009	-0.006	-0.009	-0.004
	[0.007]	[0.007]	[0.008]	[0.008]
Constant	0.929***	0.877***	0.870***	0.761***
	[0.053]	[0.054]	[0.054]	[0.055]
State FE	Yes	Yes	Yes	Yes
R squared	0.108	0.100	0.171	0.153
Observations	6507	6507	7645	7645

Notes: The dependent variables are dichotomous. 'High reading score' is a dummy variable that takes value 1 if the child was able to read a paragraph or a story in 2004-05 and 0 otherwise. 'High math

score' is a dummy variable that takes value 1 if the child is able to subtract and/or divide. 'Reading (Math) score above median' is a dummy variable that equals 1 if the child's reading (math) score was higher than the median score for her age cohort. The sample is restricted to children who were enrolled up to grade 4 in 2004-05 in columns 1 and 2, and up to grade 5 in columns 3 and 4. Robust standard errors reported in brackets. Significant at *10%, **5% and ***1%.

Table 8: Impact of learning scores in 2004-05 on years of schooling attained in 2011-12 (OLS-FE)

Variables	(1)	(2)
Year 2011-12 (Round 2)	0.726***	0.753***
	[0.018]	[0.018]
Child's age	0.023	0.014
	[0.019]	[0.019]
High reading score in 2004-05 x Round 2	0.524***	
	[0.049]	
High math score in 2004-05 x Round 2	0.397***	
	[0.047]	
Reading score above median in 2004-05 x Round 2		0.484***
		[0.040]
Math score above median in 2004-05 x Round 2		0.618***
		[0.039]
Constant	-1453.076***	-1506.392***
	[35.233]	[35.189]
Student Fixed Effects	Yes	Yes
R Squared	0.917	0.917
Observations	17638	17638
Number of students	8819	8819

Notes: The dependent variable is years of completed schooling in 2011-12. 'High reading score' is a dummy variable that takes a value 1 if the child was able to read a paragraph or a story in 2004-05. 'High math score' is a dummy variable that takes a value 1 if the child is able to subtract and/or divide. Reading (Math) above median is a dummy variable that equals 1 if the child's reading (math) score was higher than the median score for her age cohort. Sample is restricted to children who were ever enrolled in school. Children whose years of schooling are reported lower than 2004-05 are dropped from the sample. Robust standard errors in brackets. Significant at *10%, **5% and ***1%.

Figure1: Learning levels by gender

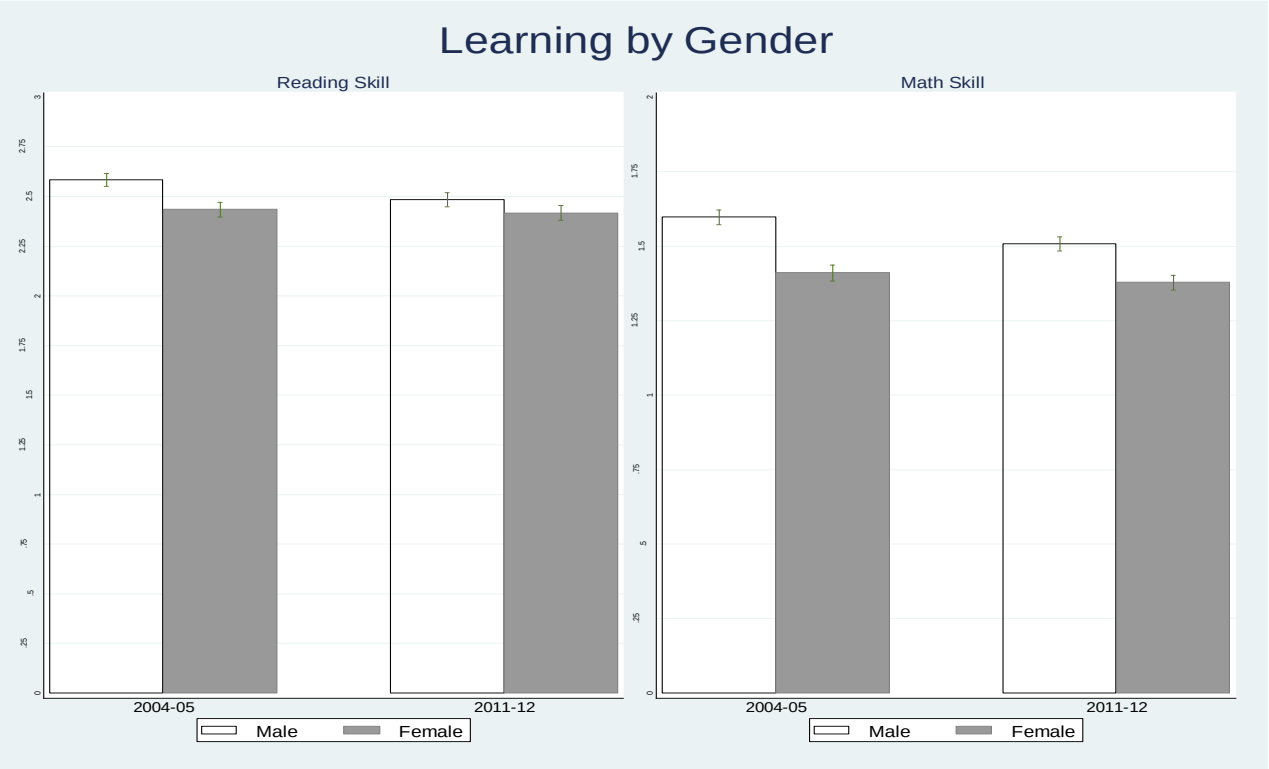


Figure 2: Learning levels by caste

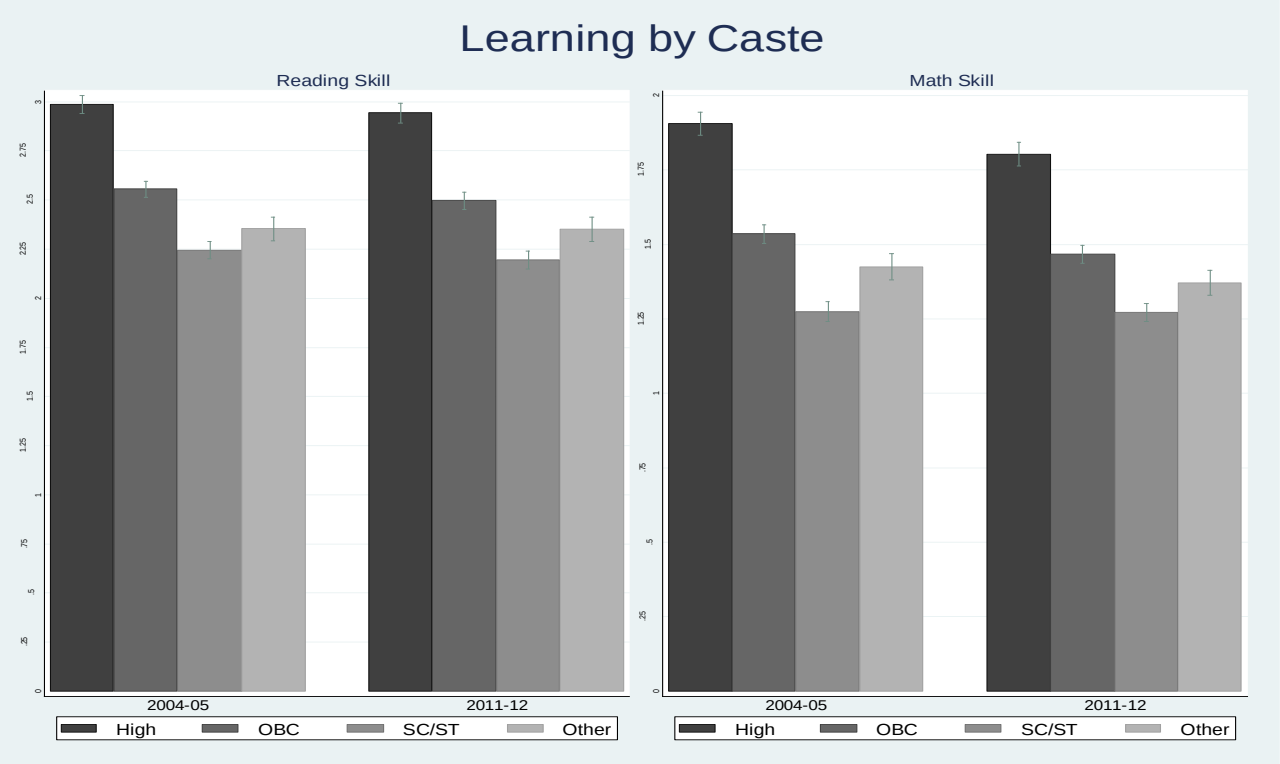


Figure 3: Learning levels by poverty status

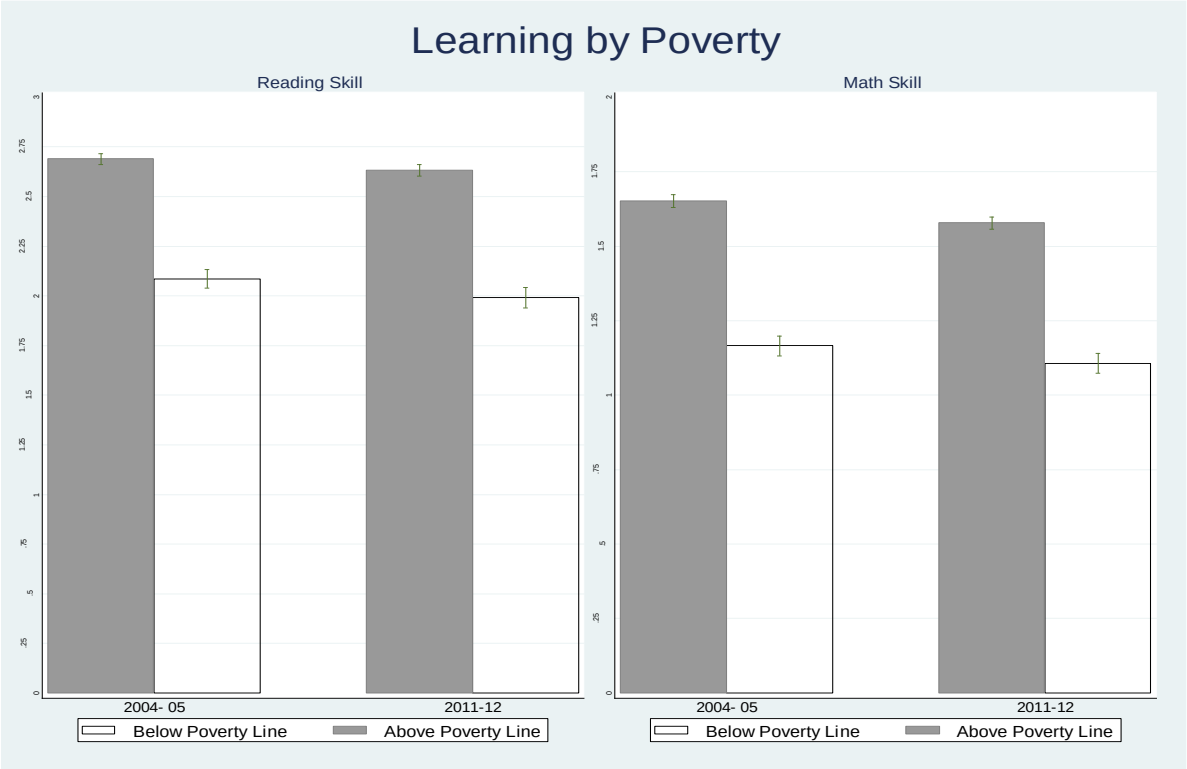
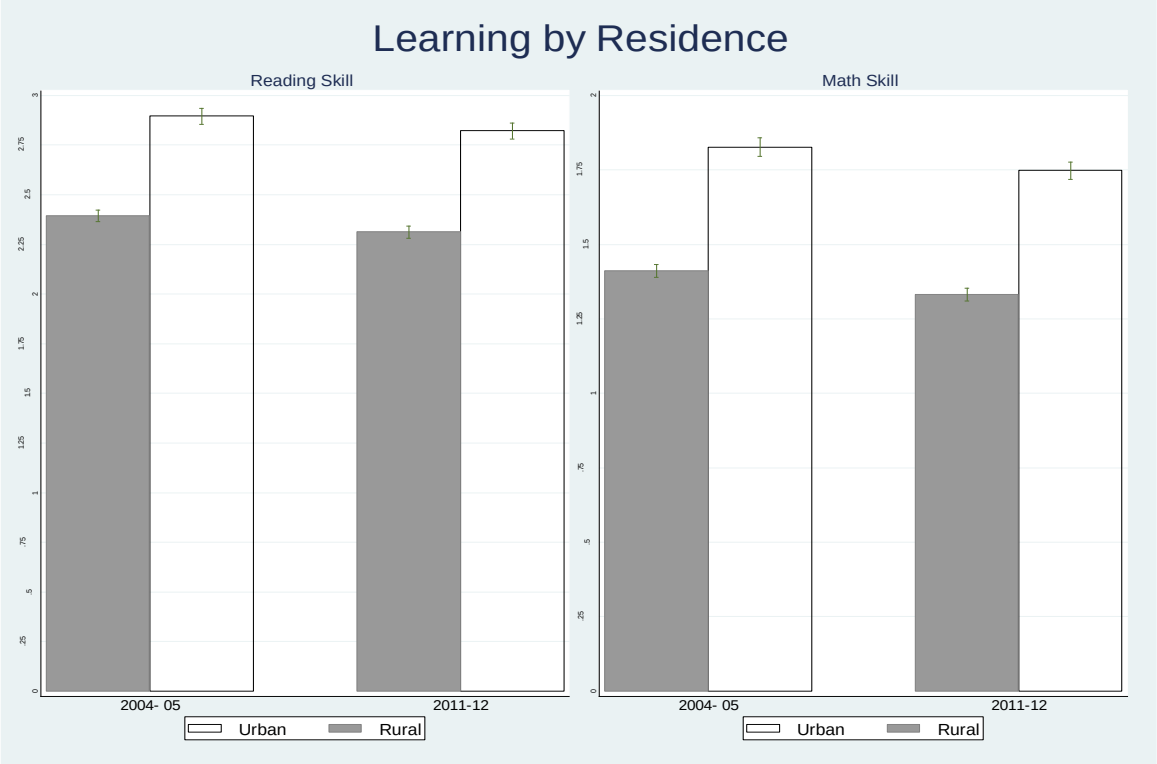
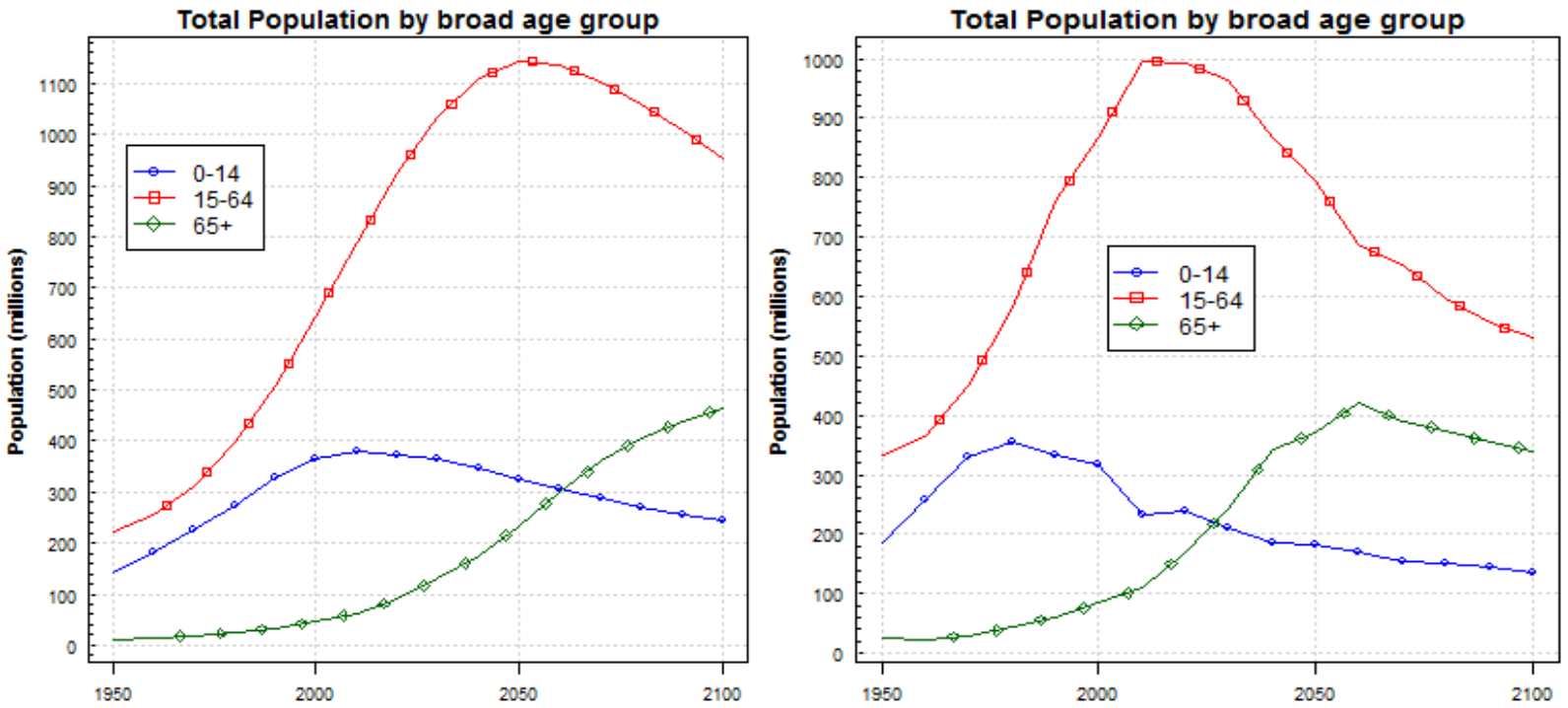


Figure 4: Learning levels by residence



Appendix

Figure A: Demographic transition



INDIA **CHINA**

Source: United Nations Population Division