Student Paper

# How Do the Effects of Private Schooling Evolve in the Long Run? How Have the Learning and Employment Outcomes of Children in the Latest Round of the Young Lives Study Changed?

Pratima Singh

Thesis submitted in part fulfilment of the requirements for the degree of MSc in Economics for Development, University of Oxford.

The data used come from Young Lives, a longitudinal study of childhood poverty that is tracking the lives of 12,000 children in Ethiopia, India (in the states of Andhra Pradesh and Telangana), Peru and Vietnam over a 15-year period. www.younglives.org.uk

Young Lives is funded by UK aid from the Department for International Development (DFID) and co-funded by Irish Aid from 2014 to 2016.

# ACKNOWLEDGEMENTS

First and foremost I want to sincerely thank my supervisor Dr. Kate Orkin. Her consistent support and guidance have been instrumental in my progress through the course and I am truly grateful to have had her as my mentor this past year. Secondly, I want to express my gratitude towards Dr. Abhijeet Singh for his feedback and for motivating me to focus my graduate research on education. His work has been truly inspiring and I could not have successfully completed my thesis without his guidance. I also want to thank Dr. Doug Gollin for being so incredibly patient and supportive and Nora Novak for constantly reminding me about submission deadlines, and for the warmest, most encouraging smiles in the hallway. Finally, I want to thank my father for encouraging me and for making my Oxford dream come true. You always have been and always will remain my biggest inspiration papa.

## ABSTRACT

In this study I attempt to identify the presence and extent of a private school premium on the cognitive skills, higher studies outcomes and psychosocial skills of children. Using a value added dynamic OLS model, I test for said premium for the younger cohort (aged 12 years), on the basis of their current and previous rounds' scores for Mathematics, English and the PPVT test. I find that private school premium for rural areas is significant for Mathematics and English. In urban areas the premium is significant for the Mathematics and PPVT tests. The magnitude of this premium has changed compared to previous studies. These differences can either be seen as a change in the way the premium for different subjects evolves as children grow older, or, as a result of more robust estimation and richer data.

For the older cohort (aged 19 years), I seek to identify the effect of attending private school in the previous round (2010) on their higher-education outcomes in the most recent round (2013). Using the linear probability model of discrete choice I find significant positive average effects of private schooling on the probability of the children attending XII$^{th}$ grade, the highest grade in secondary school, and the probability of attending post-school education – vocational, technical or degree level.

Finally, I look at the impact of private schooling on the psychosocial skills of both the cohorts. I don't find any significant impact for the younger cohort. The older cohort however, displays a significant private school premium on the agency and sense of belongingness indices. This might suggest that psychosocial skills develop later in life than cognitive skills. Therefore, even though I don't find an evidence of inequalities in this regard when the children are younger, they might grow to be significant as they enter adulthood.

# Content

# 1 Introduction

India has made considerable progress in school education in the past two decades. The findings from the tenth Annual Status of Education Report show that 96.7% of children (in the age group 6-14 years) are enrolled in school in rural India (ASER, 2014). Although near universal enrollment is a good sign, the quality of primary education has remained largely inadequate and this is reflected in the abysmally low levels of educational achievement amongst Indian school children (Kingdon 2007). Out of the sample surveyed by ASER (2014), fifty percent of class V children cannot read at even class II level and fifty percent of the children cannot do even basic arithmetic after eight years of schooling.

A growing body of literature shows that measures of school attainment (enrollment) are not as significant in explaining economic development as the level and distribution of cognitive skills (Lee and Lee, 1995, Jamison et. al., 2007). Therefore, it's the quality, and not the quantity of education that affects the economic growth of an economy. Hanushek and Woessmann (2008) show that this relationship between high performers and economic growth is especially large for poorer countries. It is imperative that the current state of school education is improved and focus is shifted to quality of education rather than just attendance, in order to make the most of India's demographic dividend.

Existing low quality of schooling has been largely attributed to the poorly resourced public schooling system in the country. Government funded, free for all public schools are characterized by teacher absenteeism, lack of basic amenities such as drinking water, proper sanitation facilities, furniture and books (Kingdon, 2007). It is perhaps as an alternative to this, that private schools have been growing rapidly in terms of both rural and urban enrollment. There has been a ten percent rise in enrollment in private schools in the past seven years (ASER, 2014). A number of studies suggest that private schools are more cost effective than public schools (Desai et al., 2008, Kingdon 1996a, 1996b, Tooley et al., 2007) and produce higher cognitive learning outcomes (Muralidharan and Kremer, 2008, PROBE, 1999, Ramachandran et al., 2004,).

If it were reasonable to believe that the expansion of private schools has come about due to their comparative advantage in producing better child outcomes then it would have extensive implications for education policies that focus on the growth of both the national income and individual earning capacities. If private schools are more efficient than public schools, it might even be prudent to restructure government schools to improve them or perhaps diversify public spending on education, which so far has focused primarily on government institutions.

An ideal way to accurately estimate the effect of private schooling and determine whether they are better at producing child outcomes than public schools would be to design an experiment where the selection into treatment (attending private school in this case) can be randomized. This would allow me to look at cognitive, psychosocial and higher study results for children in the treatment group and simply compare them to these measures for the control group (children attending public school). Getting as close as possible to the ideal method I suggest above, Muralidharan and Sundararaman (2013), generate fairly convincing experimental estimates of private school premium in the state of Andhra Pradesh. They randomize selection into private schooling by offering school vouchers through random assignment. These vouchers, given to the children in their last year of pre-school and Grade I for the entire duration up till grade 5, could be used to attend any private school in the village. They find that private school children perform better in English and Hindi and as well as public school kids in Mathematics and Telegu.

Despite all the benefits of an experimental study in terms of dealing with endogeneity, the fact remains that these vouchers are inherently artificial. They help in estimating the exact causal impact of an intervention but they don't tell us a lot about welfare implications if we don't intervene. It is extremely important to quantify exactly what is the extent of inequalities that would exist between kids if no policy changes are made. Moreover, such natural experiments are hard to come by, tricky to design, and even harder to implement.

The second best alternative is to use non-experimental panel data and try to control for any observable and unobservable differences between children that might influence selection into private schooling. The existing studies that constitute the small body of research on the role of private schools in Indian education use different quantitative techniques to deal with the possible sources of endogeneity. Most of these studies have found positive albeit varying degrees of private school premium on cognitive outcomes even after controlling for selection.

Kingdon (1996) uses data from the Lucknow district of Uttar Pradesh, India. She controls for the observed background characteristics of children, including a control for ability (using the raven's test results as a proxy). She uses models of Heckman selection to estimate private school effects on cognitive achievement. She finds a positive effect of private schooling, which attenuates but does not disappear on correcting for the selection problem. Muralidharan and Kremer (2008) use data from a nationally representative survey of rural private primary schools in India in 2003. They control for background characteristics as well as state and village fixed effects and find a large and positive private school premium, which is equal to 0.38 standard deviations even after selection is corrected for. Using three estimation techniques, namely, ordinary least squares regression, Heckman control function method based on exclusion restrictions and family fixed effects models, Desai et al. (2008) find modest but positive private school effects for reading and arithmetic skills of children. French and Kingdon, 2010 use the ASER data for 2005-2007. They try out a series of different estimations including controls for observable background characteristics of children, controlling for village fixed effects and household fixed effects to isolate village and household level confounders respectively. They find private school achievement advantage of 0.17 standard deviations when they use household fixed effects and 0.114 standard deviations when they use village level fixed effects. Chudgar and Quin, (2012) use the India Human Development Survey 2005 (IHDS) data and propensity score to deal with the selection problem. Although they find that there is a large raw private school premium when they conduct OLS, these effects disappear when they use matching estimations.

Although these studies use a host of different techniques to deal with the endogeneity problem, many concerns remain. One study that uses rich panel data containing extensive information regarding background characteristics combined with controls for shocks and patterns of time use for the children is Singh (2015). He uses longitudinal panel data for the state of Andhra Pradesh from the Young Lives study, and dynamic OLS for value added models ("VAMs") of achievement production. By generating comparable scores across rounds and cohorts using Item Response Theory ("IRT"), he finds that premium for private school children is modest and exists primarily for English for the younger cohort in his study and for Mathematics for the older cohort. The fact that using non-experimental data and dynamic VAMs, he finds estimates that seem to concur with those found in the MS study[1] boosts the robustness of his estimates and induces confidence in the methodology he employs to evaluate the private school effect.

In this paper I test for private school premium in three spheres. I test for cognitive premium for the younger cohort by using value added OLS models, following closely the methodology used in Singh (2015). I find a positive and significant premium for English at 0.35 standard deviations in rural areas but none in urban areas[2]. For Mathematics I find a private school premium of 0.18 standard deviations in both rural and urban areas. For Peabody Picture Vocabulary Test ("PPVT") I find a positive private school premium of 0.32 standard deviations in the urban areas.

I further extend the psychosocial skills analysis first conducted in Singh (2015) by creating improved indices for 'agency' and 'self-efficacy' of the children while also adding a 'sense of belongingness' index. Given the role of these skills in terms of cognitive achievement, probability of persistence in higher education, as well as simply being indicators of quality of life, any positive effect of type of schooling on these outcomes will have important implications for education policy. I find no significant effects of private schooling on any of the indices for the younger cohort. The older cohort on the other hand displays a large, positive and significant impact of private schooling in the previous period on the agency as well as the sense of belongingness indices.

Lastly, given that a large portion of the older cohort exits the schooling system in the latest round, I estimate the effect of attending private school in the previous

[1] See Singh (2015) for the explanation regarding how seemingly different results in these two studies are in fact the same and differ due to persistence through different periods and calculation methods for scores (raw scores versus IRT scores).

[2] As it is reasonable to believe that the types of private schools in these areas could be fundamentally different (Chudgar and Quin, 2012, Singh, 2015), I generate the results for rural and urban areas separately.

round, on the probability of the children attending the highest grade of secondary school and the probability of children attending higher education – vocational, technological or degree level. Given the nature of returns to education in the country, there could be large consequences of each extra year of higher education on the future earnings of the children (Kingdon, 2007). Using the linear probability model of discrete choice and controlling for background as well as a proxy for ability I find a positive and significant average private school effect of 40 percentage points on the probability of attending XIIth grade and of 48 percentage points on the probability of being attending a higher education course for the older cohort.

More specifically, I contribute to the existing literature in three ways. Firstly I improve any previous estimates of cognitive premium based on test scores by using Item Response Theory. Although Singh (2015) uses IRT to generate test scores as well, I am able to do so more robustly due to the presence of common test questions for each test in the previous and latest rounds. Singh (2015) uses different tests as a proxy for lagged test scores but due to the richness of the data in the latest round of the Young Lives study, I am able to use the previous tests taken in the same subjects as lagged scores. The results I get more or less lie between those found by Singh (2015) for the younger (aged 8 years at the time of his study) and for the older cohort (aged 15 years at the time of his study). The differences in results could be explained by the increased robustness of my estimates due to availability of richer data or perhaps by changes in premium for different subjects as children progress through elementary education. I also test for the robustness of these estimates by using an alternative empirical strategy and find that they are indeed robust. The data further allows me to test out certain assumptions necessary for the value added method to give unbiased results[3].

Secondly, the data allows me to look at how the children's later life outcomes are affected by private schooling, as the oldest cohort is 19 years old in the latest round of the Young Lives survey. The length of the panel enables me to look at post school trajectories and how they differ for private and public school children

[3] Specifically I test for unobservable factors that affect selection into private schooling and might not be controlled for by including the lagged scores on the right hand side of the regression.

if there are no policy interventions to bridge the large gap. Moreover, I can study the impact of private schooling on psychosocial skills of the children, which are believed to develop later in life, as opposed to cognitive skills, which develop fairly early on (Walsh, 2005, Grantham-McGregor et al., 2007, Cunha and Heckman, 2008).

Finally, the main aim of this study is to quantify the existing cognitive and non-cognitive inequalities between private and public school children and highlight how these affect outcomes especially as children grow older. As opposed to an experimental analysis (such as the MS study with vouchers), where the differences can be interpreted as a result of an intervention, in this study I focus entirely on the existing differences in the cognitive ability, psychosocial skills and higher study outcomes, and their consequences for child outcomes, unless policy tools to bridge these gaps are put into place.

The remainder of this essay is structured in the following way: Chapter 2 discusses the data used and descriptive statistics; Chapter 3 outlines the empirical strategy and discusses the threats to identification; Chapter 4 discusses the results, mechanisms and robustness checks; Chapter 5 concludes.

# 2 Data Description

## 2.1. Young Lives

The data used in this study is taken from the Young Lives study[4], which is a longitudinal study of child poverty. It follows two cohorts of children in four countries: Ethiopia, Andhra Pradesh state (India), Peru and Vietnam. For the

[4] The data used in this extended essay comes from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India (Andhra Pradesh and Telangana), Peru and Vietnam (www.younglives.org.uk). Young Lives is funded by UK aid from the Department for International Development (DFID), with co-funding from 2010 to 2014 by the Netherlands Ministry of Foreign Affairs, and from 2014 to 2015 by Irish Aid. The views expressed here are those of the author. They are not necessarily those of Young Lives, the University of Oxford, DFID or other funders.

purpose of this study I will use the data collected by Young Lives between 2002 and 2013, for the state of Andhra Pradesh, which comprises of four rounds of household and child surveys in the years 2002, 2007, 2009/10 and 2013 and one round of school survey in 2011 (conducted only for a sub-sample of children in the younger cohort). Administratively, the state of Andhra Pradesh is divided into districts and further split into sub-districts or 'mandals'. I will cluster standard errors at the mandal level throughout this study. The older cohort consists of 1008 children born between January 1994 and June 1995 and the younger cohort consists of 2011 children born between January 2001 and June 2002. By following the movements of children that migrated between rounds, the attrition rates in this study have been kept very low.

## 2.2. Child Characteristics

This section discusses the differences in background and time use patterns between children that attend private schools vs. children that attend public schools. Table 1 shows these differences for the children in the younger cohort.

It is clear, even more so in the case of rural areas, that there are significant and large disparities between private and government school kids when it comes to background variables such as parents' education levels, gender and so on. Time use patterns also display large gaps as children in private schools spend a lot more time on studying after school and at school itself whereas government school children spend more time on tasks such as general leisure, taking care of others and tasks on domestic farms that are unlikely to increase learning levels.

Table 1. Descriptives

| | Rural Areas | | | Urban areas | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Government | Private | Difference | Government | Private | Difference |
| Background | | | | | | |
| Mother's Education (years) | 3.28 | 5.00 | -1.72*** | 4.41 | 8.46 | -4.05*** |
| Father's Education (years) | 5.15 | 7.69 | -2.55*** | 5.50 | 9.92 | -4.42*** |
| Male | 0.49 | 0.65 | -0.16*** | 0.54 | 0.57 | -0.03 |
| First-born Child | 0.35 | 0.48 | -0.13*** | 0.35 | 0.46 | -0.11* |

| | | | | | | |
|---|---|---|---|---|---|---|
| Scheduled Caste | 0.25 | 0.08 | 0.17*** | 0.18 | 0.07 | 0.10** |
| Scheduled Tribe | 0.18 | 0.08 | 0.10*** | 0.04 | 0.01 | 0.04 |
| Other Backward Classe | 0.47 | 0.52 | -0.05 | 0.48 | 0.45 | 0.03 |
| Other castes | 0.10 | 0.32 | -0.23*** | 0.30 | 0.46 | -0.16** |
| Household size (2013) | 4.88 | 4.85 | 0.02 | 4.87 | 4.70 | 0.16 |
| Time Use | | | | | | |
| Sleeping | 8.95 | 8.86 | 0.10 | 9.12 | 8.81 | 0.30*** |
| Play/general leisure | 3.91 | 3.65 | 0.25** | 4.12 | 3.70 | 0.43** |
| Caring for others | 0.14 | 0.06 | 0.07*** | 0.11 | 0.12 | -0.02 |
| Domestic tasks and chores | 0.92 | 0.67 | 0.25*** | 0.82 | 0.67 | 0.15* |
| Studying after school | 1.97 | 2.12 | -0.16* | 1.82 | 2.04 | -0.22 |
| Tasks on family farm | 0.08 | 0.00 | 0.08*** | 0.00 | 0.00 | 0.00 |
| At school | 8.02 | 8.60 | -0.59*** | 8.01 | 8.65 | -0.64*** |
| Paid work outside household | 0.02 | 0.03 | -0.01 | 0.00 | 0.00 | 0.00 |

Note: Cells present mean of variable within school type. Asterisks indicate significant difference between mean in private and government schools within urban and rural areas. *** p<0.01, ** p<0.05, * p<0.1.

A big consideration in all the estimations used in this study is whether to control for time use or not. As is clear, there are big differences between how these children spend their time every day. However, it might be possible that the extra time spent by private school kids on studying inside and outside school is a direct result of the treatment, i.e. being enrolled in a private school. Therefore, controlling for it might bias our estimates downwards as some of the positive effects of private schooling that translate into say more hours of out of school studying will be taken out of the private school premium. However, it might be the case that certain types of kids engage in certain types of activities and therefore it might be correlated with the probability of getting treatment, i.e. private schooling (Edmonds, 2007). Thus, controlling for time use will give us a lower bound for our estimates whereas not controlling for time spent on domestic tasks etc. might depend on the social and economic background conditions, which might bias our estimates upwards. It is for this reason that I add time use as a control in the end of sequential controls. This way estimates before and after adding the time use controls can be compared.[5]

[5] On comparing these estimates for the cognitive learning outcomes it is clear that including time use as a control does not alter my estimates significantly and therefore the bias I point out here is minimal.

## 2.3. Test Scores

All test scores in this study (except for the ravens test scores) were calculated based on the cognitive tests taken by the younger cohort in round 2, round 3, the school survey and round 4, using Item Response Theory models. IRT models are more efficient than raw scores or percentage correct scores as they account for the relative difficulty of each question on a test by assuming a mathematical relationship between the latent ability of children and the probability of them answering a question correctly, which varies depending on the difficulty of each question. Since the tests taken in the past few rounds of the Young Lives study have common questions, which are repeated each year, this allows for creation of scores from different tests on the same scale and facilitates comparability across rounds. This implies that I can study the growth of child learning/achievement as the younger cohort progresses through elementary school and understand the long-term effects of private education on the cognitive outcomes. I use a three-parameter logistic ("3PL") IRT model and calculate scores based on maximum likelihood with the openirt command in Stata written by Tristan Zajonc. Das and Zajonc (2010) provide a detailed discussion of the 3PL IRT model. Due to the richness of the Young Lives data, especially the reasonably comprehensive test measures used, coupled with IRT, I am able to truly capture the variation across the ability spectrum of the children which might be lost if raw scores are used[6].

# 3. Methodology and Estimation

## 3.1. Value Added Models

Value added models of achievement production have been used extensively in multiple studies focusing on the performance of different schools and teachers in terms of cognitive learning outcomes. In this study, following closely the methodology used in Singh (2015) and Todd and Wolpin (2003,2007), I estimate value added models of achievement production using the dynamic OLS estimation, to assess whether the higher cognitive outcomes in private schools are

[6] All IRT test scores were normalized to have a mean of 0 and a standard deviation of 1.

a direct effect of private schooling or can be attributed solely to differences in the socio-economic backgrounds of children.

The basic structure of the value added models is that of cumulative effects. Todd and Wolpin give the general form of the achievement production model as follows:

$$y^{*}_{ist} = F[X_{i}(t), S_{i}(t), \mu_{is0}, \epsilon_{ist}] \qquad \{1\}$$

Where $y^{*}_{ist}$ is the achievement of child *i* in school *s* at time *t*. Achievement is a function of a vector of complete background variables $X_{i}(t)$, school based inputs $S_{i}(t)$, student endowments $\mu_{is0}$, and a time varying error term $\epsilon_{ist}$.

In this study I will estimate the lagged value added models of achievement production derived by Singh (2015) following the model specified above in conjunction with the model specified in the Andrabi et al. (2011) study. The dynamic OLS model Singh derives is as follows:

$$y^{*}_{ist} = \alpha'_{1}.x_{it} + \beta y^{*}_{i,t-1} + \mu_{it} \qquad \{2\}$$

The lagged score on the right hand side is added to capture the effect of all inputs, unobservable endowments (such as ability) and shocks in previous periods. I will discuss how addition of this control might threaten identification in section 3.3 below.

## 3.2. Main Specification

Following this, the main specification in this study is as follows:

$$Y_{it} = \alpha + \beta_{1}.\text{Private}_{it} + \beta_{2}.\text{site}_{i} + \beta_{3}.X_{i,t-1} + \beta_{4}.Y_{i,t-1} + \beta_{5}.\text{timeuse}_{it} + \varepsilon_{it} \qquad \{3\}$$

Where, $Y_{it}$ is the outcome variable that represents either cognitive outcome (or non-cognitive outcome in the next specification). $Private_{it}$ is the dummy for whether the child went to private school in 2013/14 (for the younger cohort) or in 2009/10 (for the older cohort), and enrollment in a government school is the base category; $site_{i}$ is a vector of senital site ('mandal' or sub-district) dummy variables; $X$ is a vector of background characteristics including caste dummies, wealth index, maternal and paternal years of schooling, household size, the dummy for the sex of the child and a dummy for whether she/he is the eldest child in the family; $Y_{i,\ t-1}$ is the lagged test score [7] (different in different specifications); $timeuse_{it}$ is the number of hours spent on a typical day in various activities. In most specifications, under time use I control specifically for play time /general leisure, time spent on caring for others, domestic tasks and chores, studying outside of school, tasks on the family farm, at school (including commuting to and from school) and paid work outside of home. In all specifications, the controls are added sequentially so that estimates for the private school effects using different controls can be observed. The lagged and latest scores used in this study are as follows:

| | | | |
|---|---|---|---|
| Recent Score | Math - Round 4 (12 years) | English - Round 4 (12 years) | PPVT - Round 4 (12 years) |
| Lagged Score | Math - School Survey (9 years) | English - School Survey (9 years) | PPVT - Round 3 (8 years) |

## 3.3. Threats to identification and sources of potential bias

The main source of potential bias in this study arises from the use of lagged scores as a control. Firstly, any measurement error in test scores can attenuate the beta coefficient on the lagged test scores and possibly bias the input parameters as well. Secondly, due to the nature of the relationship between $Y_{it}$ and unobservable characteristics captured by the error term, it is likely that $Y_{i,t-1}$ is also correlated with the unobservable characteristics that determine cognitive achievement. This would cause endogeneity and bias the estimates of private school premium. For instance, if students' unobservable endowments such as ability results in children with higher ability to learn faster every year, the

[7] Lagged test scores are taken only from the previous period. The reason why more lags are not added is that Singh (2015) tests whether the results change significantly on adding more lags to the equation and finds that the estimates remain the same. This implies that the latest set of lagged score results capture all the information that addition of more lags could provide. I also see this in the specification where I test for effects of adding lags from the second round of the survey to the main specification (Table A). I too find no significant effect of adding more than one lag in the estimation.

coefficient $Y_{i,t-1}$ will be biased upwards due to the positive correlation between lagged scores and ability.

This specification is used despite the above-mentioned concern because in this particular case, I argue that the lagged score can sufficiently account for ability and previous investments into achievement production (Singh, 2015, Todd and Wolpin, 2003). Thus ability does not enter the error term and bias our estimates.

Singh (2015) tests for this by using the raven's test as a proxy for ability for the older cohort. To test this in my data I use the Cognitive Development Assessment ("CDA") and PPVT scores from the first round of cognitive tests taken by the younger cohort as proxy measure of initial ability as the children were 5 years of age at the time of these tests and it predates school entry for most of them. I add these tests to the main specification for the cognitive effects of private schooling. The premise being that addition of these test scores as controls should not have any significant effect on achievement at 12 years of age (latest round) when controlling for achievement at the age of 8-9 years (previous round). I find that this condition holds. The effect of the ability proxies on private school premium is marginal and barely significant (See table A in the appendix). Therefore I can say that conditioning on the lagged scores does not lead to bias from the channel discussed above.

However, any other type of sorting that is not controlled for by the lagged achievement measures and background plus time controls could possibly be a source of bias for estimates from my main specification. For instance, if parental assessment of child performance plays a role in deciding whether the children are enrolled into private or public schooling, this might bias my estimates of the effect of private schooling. The analysis of a field experiment in Malawi shows that even though parents try to match the investments into education to their perceived academic level of the child, they are often wrong or biased in these assessments. Poorer households are more likely to inaccurately assess the child's level and therefore under-invest in education, thereby increasing the rich-poor gap in education further (Dizon-Ross, 2015)[8]. Singh (2015) also suggests that there are many ways in which parental assessment might differ from information

[8] This is a working paper by Rebecca Dizon-Ross titled, 'Parents' Perceptions and Children's Education: Experimental Evidence from Malawi', 2015. Find pdf at: http://faculty.chicagobooth.edu/rebecca.dizon-ross/research/papers/perceptions.pdf

contained in achievement data and this could in turn cause selection based on ability that is not controlled for by simply conditioning on lagged scores. To test for this he uses proxies for parental assessment of performance as well as parental aspirations from measures in the early rounds of young lives data and finds that including these measures in the estimation did not change the coefficients on the private school dummy, thereby implying that this is not a source of bias in the case of this specific study at least. Since I use the same data source I do not reproduce his test results here but follow the same reason to reject this source of bias.

Despite the concerns of possible endogeneity in VAMs, they are being increasingly used as the most efficient out of the available empirical strategies, to capture the effect of school/teacher inputs into education. Another way of looking at VAMs is as a treatment effects estimator for identifying policy effects. In this regard, the lag can be understood as a control to soak up unobserved differences, much like a matching estimator (Guarino et al., 2015). Dynamic OLS value-added models have been shown to have minimal (and statistically insignificant) forecast bias as compared to models that control for previously unobserved parental characteristics and those models that study quasi-experimental results based on teacher transfers (Chetty et al. (2014). Andrabi et al., 2011, look at value added models for panel data from Pakistan and discuss the strengths of using value added school estimates in a developing country setting. They also talk about persistence of learning between grades and state that assuming perfect persistence leads to downward estimates of the private school premium in test scores.[9] To test for robustness of the estimates I find using Young Lives data, I will estimate a propensity-score-matching model as well and compare its estimates to those found using dynamic OLS.

## 3.4. Psychosocial Skills

In the latest round of household survey, conducted in the year 2013, the Young Lives team collected data on a wide series of attitudinal items for both the

[9] I do not look at persistence in this study. However, Singh (2015) calculates the implied per year effect from the three year estimates of private school premium, using the persistence coefficient in Andrabi et al., 2011. These are lower than his three-year estimates as is expected in case of more than zero persistence.

younger and the older cohort. Using the responses to these items and following methodology specified in child psychology literature, I create three indices of psychosocial skills to determine the existence and extent of private school premium in this aspect of education (See Table C in the appendix for the list of items included in each index). These non-cognitive skills have been shown to significantly impact production of cognitive skills as well as later life outcomes of children (Cunha and Heckman, 2008; Cunha et al., 2010, Robbins et al., 2004). Given the range of items in the 'feelings' section of the latest child survey questionnaire, I construct three indices, namely the agency index, the self-efficacy index and the sense of belongingness index.[10] I will outline the theoretical basis used to construct these indices later in this section. The private school effect on these skills is measured using the core estimation equation (eq. 3), where the outcome variable $Y_{it}$, is now the index of psychosocial skills, $Y_{i,t-1}$ is the lagged score from the first tests ever taken by the children (CDA and PPVT for the younger cohort and raven's test for the older cohort)[11]. Time use is no longer controlled for because as discussed previously, it poses the threat of endogeneity and there is evidence that shows time use patterns don't affect non-cognitive outcomes whereas they do affect cognitive achievement production (Fiorini and Keane, 2014). It is for this reason that I control for them in the main specification but not in this estimation.

A large body of child education and psychology literature suggests that not only do psychosocial skills impact the production of cognitive skills but they also play an important role in determining the probability of persistence in higher education (Tinto, 1993, Bean, 1985, Pascarella and Terenzini, 1983). Pascarella and Chapman (1983) find that motivation and social integration are in fact the primary predictors of persistence in education. Therefore understanding the role of agency, self-efficacy and sense of belongingness in terms of secondary and post-secondary education outcomes is paramount.

[10] All indices are standardized to have mean 0 and standard deviation 1.

[11] It is reasonable to believe that unlike test scores, psychosocial outcomes are affected significantly by motivation in the early years of schooling and this would be reflected in any tests taken after school entry. Therefore, controlling for any of the recent tests would bias the effect of private schooling on these effects downwards (Singh, 2015). The CDA, PPVT and raven's tests on the other hand predate school entry for most of the children in both cohorts and therefore make for better controls of cognitive ability in this estimation.

### 3.4.1. Agency

Agency or locus of control is a measure of the extent to which individuals feel in control of their future outcomes (Findley and Cooper, 1983). Studies suggest that people who are 'internals', i.e. those who feel in control of their future outcomes, are more likely to achieve goals than 'externals', i.e. people who feel that external forces are responsible for their outcomes. Ducette and Wolk (1972) suggest that this is because externals exhibit lesser persistence at tasks whereas Bialer (1961) shows that there is a positive relationship between internality and a "willingness to delay rewards in order to maximize them". Therefore, for apparent reasons, if private schooling affects the locus of control, there could be important consequences for future child outcomes.

### 3.4.2. Self-Efficacy

Schwarzer (1992) describes 'Perceived Self-Efficacy' as "an optimistic self-belief". Self-efficacy is expected to positively affect goal setting, investment of effort and responses to difficult situations and setbacks. The index of self-efficacy constructed for the purpose of this study has been derived following the methodology given in Schwarzer and Jerusalem (1995). This scale has certain advantages over other measures of self-efficacy and has been used extensively with success as a predictor of adaption after life changes or simply as an indicator of quality of life.

### 3.4.3. Sense of Belongingness

The level of engagement in school and the general sense of belongingness are important outcomes of schooling in themselves. Studies suggest that they may predict mental and physical health as well as performance (Cueto et al., 2010, Juvonen, 2006, Faircloth and Hamm, 2005). The sense of belongingness index used in this study is constructed following the methodology specified in PISA (Willms, 2003).

### 3.5. Higher study outcomes

Secondary and post-secondary enrollment rates have been historically low in India but have been rising recently (World Bank, 2006). Kingdon (2007) attributes this growth in enrollment rates partly to rise in private schooling. Various studies show that the cause for low levels of secondary enrollment is anything but lack of returns to higher education. Kingdon (1998) shows that the education-wage relationship in India is u-shaped. That is, returns to secondary and higher education are significantly larger than returns to primary and middle levels of education.[12] It is therefore of great importance to determine whether private schools are more successful than public schools in terms of progression to the final stage of secondary education, i.e. XIIth grade and post-secondary education, i.e. degree, technological or vocational education. In this final estimation, using the linear probability model of discrete ordered choice, I follow the core estimation mentioned previously where the outcome variable $Y_{it}$ is the probability of the child attending XIIth grade or the probability of the child attending any of the three types of higher education. I continue to control for background, mandal fixed effects and time use. The lagged scores used as a measure of cognitive ability are once again taken from the first cognitive tests taken by the children rather than the most recent round of lagged scores to fully capture the private school effects in this estimation.

## 4. Results and Mechanisms

### 4.1. Cognitive Effects

In order to compare the improved estimates from IRT generated scores, I first present the results from the estimation based on only standardized raw scores in table 2(a) for the rural areas and table 2(b) for the urban areas. Table 3(a) shows the results from the first specification, i.e. the effect of private schooling on the cognitive ability of the younger cohort in the three tests in Mathematics, English

[12] Also see Kingdon and Unni (2001) and Kingdon (2007). Kingdon (2007) also states that estimations of the wage function using National Sample Survey data shows that the returns to each additional years of schooling are positive and significant for almost all states in India and for both genders.

and PPVT taken in the latest round of the Young Lives survey, for the rural areas. Each level of controls is implemented sequentially. Table 3(b) shows the corresponding estimates for urban areas. As can be seen, raw scores fail to capture the private school effects fully and in the case of urban areas they find no significant effects at all. IRT test scores capture the variation in ability of students and therefore use the information in test scores more efficiently than raw scores that do not take difficulty and other parameters of the questions administered into account.

Table 2(a): Raw scores: Rural

| VARIABLES | English | Mathematics | PPVT | English | Mathematics | PPVT |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Private | 0.29** | 0.32** | 0.17* | 0.24** | 0.24* | 0.10 |
| | (0.11) | (0.12) | (0.087) | (0.11) | (0.11) | (0.12) |
| Controls: | | | | | | |
| Background | Y | Y | Y | Y | Y | Y |
| Time use | N | N | N | Y | Y | Y |

Table 2(b): Raw scores: Urban

| VARIABLES | English | Math | PPVT | English | Math | PPVT |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Private | 0.21 | 0.31 | 0.41 | 0.18 | 0.30 | 0.33 |
| | (0.20) | (0.21) | (0.35) | (0.21) | (0.24) | (0.27) |
| Controls: | | | | | | |
| Background | Y | Y | Y | Y | Y | Y |
| Time Use | N | N | N | Y | Y | Y |

Note: All scores are non-IRT raw scores. Site fixed effects are controlled for. Standard errors are clustered at the sub-district level. Controls are defined as in the main regression. ***p<0.01, **p<0.05, *p<0.1.

Table 3(a) Cognitive Effects: Rural

| Tests | Math | English | PPVT | Math | English | PPVT |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Private School | 0.22*** | 0.38*** | -0.042 | 0.18** | 0.35*** | -0.082 |
| | (0.065) | (0.099) | (0.074) | (0.070) | (0.094) | (0.062) |
| Controls: | | | | | | |
| Background | Y | Y | Y | Y | Y | Y |
| Time Use | N | N | N | Y | Y | Y |
| Lagged Scores | Y | Y | Y | Y | Y | Y |
| Constant | -0.23** | -0.089 | -0.60*** | -0.72* | -0.74** | -2.06*** |

| | | | | | | |
|---|---|---|---|---|---|---|
| | (0.079) | (0.18) | (0.17) | (0.38) | (0.35) | (0.60) |
| Observations | 659 | 619 | 1,146 | 659 | 619 | 1,146 |
| R-squared | 0.428 | 0.392 | 0.310 | 0.446 | 0.405 | 0.324 |

Note: The controls are as specified in the main specification. All scores are created using IRT models. Site fixed effects are controlled for. *** p<0.01, ** p<0.05, * p<0.1

Table 3(b) Cognitive Effects: Urban

| Tests | Mathematics | English | PPVT | Mathematics | English | PPVT |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Private School | 0.19** | 0.013 | 0.42** | 0.18** | 0.018 | 0.32* |
| | -0.064 | -0.26 | -0.11 | -0.062 | -0.26 | -0.14 |
| Controls: | | | | | | |
| Background | Y | Y | Y | Y | Y | Y |
| Time Use | N | N | N | Y | Y | Y |
| Lagged Scores | Y | Y | Y | Y | Y | Y |
| Constant | -0.62** | 0.36 | -0.37 | -0.48 | -2.20* | -4.15* |
| | -0.21 | -0.24 | -0.58 | -0.8 | -0.81 | -1.67 |
| Observations | 141 | 139 | 252 | 141 | 139 | 252 |
| R-squared | 0.463 | 0.425 | 0.173 | 0.481 | 0.464 | 0.209 |

Note: The controls are as specified in the main specification. Site fixed effects are controlled for. All scores are created using IRT models. *** p<0.01, ** p<0.05, * p<0.1.

As can be seen from table 3(a) and 3(b), there is significant and positive private school premium for English of 0.35 standard deviations and for Mathematics of 0.18 standard deviations in rural areas. In urban areas the premium is equivalent to 0.18 standard deviations for Mathematics and 0.32 standard deviations for PPVT.

### 4.1.1. *Can private school premium change as children progress through elementary school?*

Comparing these results to those found by Singh (2015) shows that the effect of private schooling on mathematics found by him for children who were at the time 15 years of age (i.e. 0.20 standard deviations) is roughly the same as that found for 12 year old children in this study (0.18 standard deviations) and more than those found at 8 and 9 years old (for which he found no significant premium in mathematics) children in his study. This could mean that the private school premium for mathematics consistently grows larger with age, thus indicating increasing inequality in the mathematical skills levels of private and public school children. Since he did not study the effects of private school premium in English for the older cohort (15 years old at the time of his study), I can only compare the estimates from this study to those of 9-year-old children in his study. He found a larger private school premium in English (i.e. 0.66 standard deviations) than I do for 12 year olds in my sample (i.e. 0.35 standard deviations). This could possibly imply that perhaps students get better at self teaching themselves English as they grow older and therefore the gaps between the private and public school children grows smaller. Although Singh (2015) does not find any significant private school effects for urban areas, I find there is significant premium in mathematics and PPVT.

Since I find premium even in areas and subjects that he does not, I believe that a part of the differences are a result of the fact that I find more robust estimates due to richer data available that allows me to make efficient use of the IRT technique and improve upon any prior estimation[13].

### 4.1.2. Robustness Check

As a measure of robustness, I also do propensity score matching for this estimation. Propensity score matching has the advantage that it does not assume a linear relationship and provides a certain level of flexibility in the functional form of the relationship between the covariates and test scores.[14] I matched children on the basis of the level of education of the mother and father, wealth index of the family, gender and lagged scores using nearest neighbor propensity score matching. The results from that regression are given in table B in the appendix. These results are comparable with the results in the main specification, before time use patterns are controlled for. As is evident, results for both rural and urban areas are still the same in direction when I match children on the basis of propensity scores even though magnitudes are very slightly and not very

[13] I also have better data for patterns of time use that allows me to control for different types of work the child spends time on and may have an impact on the way these estimates have changed between Singh (2015) and my study.

[14] This improves on my main estimation, which assumes a linear relationship. However, functional form assumptions are still required to calculate the propensity score.

significantly altered in different directions for different subjects. Such close estimates using this technique boosts the confidence in estimates derived from the dynamic OLS estimation.

## 4.2. Psychosocial Skills:

Table 4(a) Psychosocial Skills: Younger Cohort

| VARIABLES | Agency | | | Self-Efficacy | | | Sense of belongingness | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Private | 0.13** | 0.11** | 0.092* | 0.16** | 0.093 | 0.077 | 0.0012 | -0.017 | -0.012 |
| | (0.045) | (0.049) | (0.047) | (0.063) | (0.057) | (0.060) | (0.061) | (0.076) | (0.079) |
| *Controls:* | | | | | | | | | |
| Background | N | Y | Y | N | Y | Y | N | Y | Y |
| Ability Proxy | N | N | Y | N | N | Y | N | N | Y |
| | | | | - | - | - | | - | - |
| Constant | -0.029** | -0.16 | -0.12 | 0.072*** | 0.48*** | 0.35** | 0.019 | 0.53** | 0.61*** |
| | (0.011) | (0.16) | (0.15) | (0.016) | (0.15) | (0.15) | (0.015) | (0.21) | (0.21) |
| Observations | 1,258 | 1,257 | 1,225 | 1,258 | 1,257 | 1,225 | 1,258 | 1,257 | 1,225 |
| R-squared | 0.013 | 0.023 | 0.025 | 0.073 | 0.084 | 0.086 | 0.024 | 0.043 | 0.048 |

Note: All the controls are defined as specified in the main specification. Standard errors are clustered at the sub-district level. *** p<0.01, ** p<0.05, * p<0.1

Table 4(a) shows that there is no significant private school premium in terms of psychosocial skills when it comes to the younger cohort. This is similar to what Singh (2015) finds for the younger cohort from their answers to the attitudinal items in the school survey in 2009. These results could imply that these children have not gained from private schooling in terms of psychosocial outcomes since the last round.

For the older cohort however, I find significant and positive private school premium on the agency and sense of belongingness indices. Table 4(b) shows that the private school premium for the older cohort on the agency index is 0.41 standard deviations and on the sense of belongingness index is 0.29 standard deviations, that too after controlling for background characteristics and lagged scores from 2007.

It is also plausible that the reason that I find a positive and large private school premium only for the older cohort is that these skills develop only at later stages in life. Walsh 2005 shows that non-cognitive or psychosocial skills are more malleable till later ages than cognitive skills. This malleability in later life is associated with the slow development of the prefrontal cortex, which is responsible for the formation of non-cognitive skills. Therefore even though there are no significant effects of private schooling on the non-cognitive outcomes of the younger cohort, these may grow very large, as they get older.

These results become even more important given the age group these children fall in due to the previously discussed literature suggesting the impact of psychosocial skills on higher study outcomes.

Table 4(b): Psychosocial Skills: Older Cohort

| VARIABLES | Agency | | | Self-Efficacy | | | Self of Belongingness | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Private | 0.61*** | 0.42*** | 0.41*** | 0.25** | 0.18 | 0.16 | 0.34*** | 0.28*** | 0.29*** |
| | (0.10) | (0.11) | (0.10) | (0.088) | (0.10) | (0.099) | (0.070) | (0.091) | (0.096) |
| *Controls:* | | | | | | | | | |
| Background | N | Y | Y | N | Y | Y | N | Y | Y |
| Ability (Proxy) | N | N | Y | N | N | Y | N | N | Y |
| | - | | | - | | | - | | |
| Constant | 0.18*** | -0.56** | -0.86** | 0.094*** | -0.13 | -0.53* | 0.082*** | -0.44* | -0.29 |
| | (0.019) | (0.22) | (0.35) | (0.017) | (0.25) | (0.27) | (0.013) | (0.23) | (0.34) |
| Observations | 716 | 716 | 716 | 715 | 715 | 715 | 716 | 716 | 716 |
| R-squared | 0.095 | 0.160 | 0.163 | 0.041 | 0.084 | 0.091 | 0.050 | 0.064 | 0.065 |

Note: Standard errors in parentheses are clustered at the sub-district level. Regression includes sub-district fixed effects. Results are from the rural sample of the 2013 data collection.

## 4.3. Higher Study Outcomes

Table 5(a) shows the average effect of private schooling in the previous round on the probability of the child attending XII$^{th}$ grade. There is a significant effect of gender and the distribution of time on different activities on probability of progressing to the highest grade in school. Here, addition of time use as a control might cause endogeneity. However I add it sequentially and the effects before adding time controls are as seen in column (3)[15]. The effect of private schooling in the previous round on likelihood of entering the highest grade in school is 40 percentage points before controlling for time use and 12 percentage points after controlling for time use.

Table 5(a) Probability of progressing to highest grade in school

| VARIABLES | XIIth Grade | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Private (2009) | 0.52*** | 0.41*** | 0.40*** | 0.12** | 0.12*** |
| | (0.032) | (0.033) | (0.035) | (0.043) | (0.039) |
| Mother's education level (years) | | 0.0034* | 0.0030 | -0.00068 | -0.0019 |
| | | (0.0019) | (0.0020) | (0.0023) | (0.0026) |
| Father's education level (years) | | 0.0055* | 0.0049 | 0.0032 | 0.0040 |
| | | (0.0030) | (0.0030) | (0.0034) | (0.0030) |
| Male | | 0.21*** | 0.21*** | 0.15*** | 0.12*** |
| | | (0.039) | (0.039) | (0.030) | (0.035) |
| Eldest child in the household | | 0.11** | 0.10** | 0.073* | 0.072* |
| | | (0.041) | (0.041) | (0.036) | (0.034) |
| Scheduled Caste | | 0.025 | 0.028 | 0.082 | 0.073 |
| | | (0.090) | (0.090) | (0.068) | (0.064) |
| Scheduled Tribe | | 0.092 | 0.10 | 0.096 | 0.080 |
| | | (0.081) | (0.080) | (0.069) | (0.062) |
| Other Backward Classes | | -0.046 | -0.039 | 0.014 | 0.026 |
| | | (0.068) | (0.069) | (0.060) | (0.056) |
| Wealth index | | 0.39*** | 0.37*** | 0.13 | 0.073 |
| | | (0.13) | (0.12) | (0.12) | (0.12) |
| Household size | | -0.019** | -0.020** | -0.0057 | -0.0027 |
| | | (0.0082) | (0.0084) | (0.0061) | (0.0056) |
| Leisure | | | | -0.038*** | -0.033*** |
| | | | | (0.0077) | (0.0071) |
| Care for others | | | | -0.078*** | -0.065*** |
| | | | | (0.0074) | (0.0089) |
| Domestic tasks | | | | -0.046*** | -0.037*** |
| | | | | (0.012) | (0.011) |
| Studying at home | | | | 0.016 | 0.014 |
| | | | | (0.011) | (0.011) |
| Domestic Tasks | | | | -0.064*** | -0.058*** |
| | | | | (0.0071) | (0.0072) |
| Activities for pay outside the hh | | | | -0.064*** | -0.058*** |

[15] I tested for collinearity in this regression and found none. Results of the vif test are not attached in this document but are available on request.

| | | | | | |
|---|---|---|---|---|---|
| | | | | (0.0035) | (0.0036) |
| Raven's Test Score | | | 0.0074 | 0.0026 | 0.0017 |
| | | | (0.0047) | (0.0041) | (0.0036) |
| Agency index | | | | | 0.089*** |
| | | | | | (0.015) |
| Sense of Belongingness index | | | | | 0.0024 |
| | | | | | (0.013) |
| Constant | 0.45*** | 0.18 | 0.017 | 0.86*** | 0.85*** |
| | (0.035) | (0.12) | (0.15) | (0.15) | (0.14) |
| Observations | 528 | 528 | 528 | 528 | 528 |
| R-squared | 0.173 | 0.254 | 0.260 | 0.553 | 0.583 |

Note: Standard errors are clustered at the sub-district level. ***p<0.01, **p<0.05, *p<0.1.

Table 5(b) shows the average effect of private schooling on the probability of the child progressing to a higher level of education, be it technological, vocational or degree level. Column (3) shows the results from just controlling for background effects and the ability proxy (raven's test scores from 2002). Once again, the effect of gender and time use patterns is significant.

Table 5(b) Probability of progressing to higher education:

| VARIABLES | Higher Education | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Private (2009) | 0.57*** | 0.48*** | 0.48*** | 0.14** | 0.13*** |
| | (0.041) | (0.050) | (0.050) | (0.046) | (0.044) |
| Mother's education level (years) | | 0.0045 | 0.0043 | 0.00031 | -0.00024 |
| | | (0.0028) | (0.0028) | (0.0012) | (0.0014) |
| Father's education level (years) | | 0.0028 | 0.0024 | 0.00015 | 0.00048 |
| | | (0.0037) | (0.0038) | (0.0027) | (0.0028) |
| Male | | 0.20*** | 0.20*** | 0.14*** | 0.13*** |
| | | (0.029) | (0.028) | (0.017) | (0.020) |
| Eldest child in the household | | 0.065* | 0.064 | 0.025 | 0.024 |
| | | (0.036) | (0.036) | (0.027) | (0.027) |
| Scheduled Caste | | 0.038 | 0.040 | 0.094* | 0.090** |
| | | (0.070) | (0.071) | (0.045) | (0.041) |
| Scheduled Tribe | | 0.052 | 0.059 | 0.037 | 0.030 |
| | | (0.072) | (0.071) | (0.054) | (0.051) |
| Other Backward Classes | | -0.020 | -0.015 | 0.039 | 0.043 |
| | | (0.051) | (0.052) | (0.039) | (0.037) |
| Wealth index | | 0.30** | 0.29** | -0.025 | -0.050 |
| | | (0.11) | (0.12) | (0.093) | (0.093) |
| Household size | | -0.015 | -0.016 | 0.0031 | 0.0045 |

| | | | | | |
|---|---|---|---|---|---|
| | | (0.0093) | (0.0093) | (0.0061) | (0.0058) |
| Leisure | | | | -0.056*** | -0.053*** |
| | | | | (0.0074) | (0.0072) |
| Care for others | | | | -0.075*** | -0.069*** |
| | | | | (0.0086) | (0.0083) |
| Domestic tasks | | | | -0.049*** | -0.045*** |
| | | | | (0.0084) | (0.0077) |
| Studying at home | | | | 0.037 | 0.036 |
| | | | | (0.023) | (0.022) |
| Domestic Tasks | | | | -0.078*** | -0.076*** |
| | | | | (0.0068) | (0.0068) |
| Activities for pay outside the hh | | | | -0.071*** | -0.068*** |
| | | | | (0.0068) | (0.0065) |
| Raven's Test Score | | | 0.0050 | -0.0010 | -0.0014 |
| | | | (0.0032) | (0.0033) | (0.0033) |
| Agency index | | | | | 0.038** |
| | | | | | (0.013) |
| Sense of Belongingness index | | | | | 0.0058 |
| | | | | | (0.012) |
| Constant | 0.31*** | 0.084 | -0.023 | 0.99*** | 0.98*** |
| | (0.037) | (0.086) | (0.11) | (0.14) | (0.14) |
| Observations | 528 | 528 | 528 | 528 | 528 |
| R-squared | 0.206 | 0.264 | 0.267 | 0.688 | 0.694 |

Note: Standard errors are clustered at the sub-district level. ***p<0.01, **p<0.05, *p<0.1.

#### 4.3.1. *Can higher psychosocial skills outcomes lead to better higher study outcomes for private school children?*

Taken together with the previous specification, where for the older cohort I find high and significant positive effects of private schooling on psychosocial outcomes, and the previously mentioned literature discussing the impact of non-cognitive skills on persistence through higher education, it would be interesting to see if these results can actually be correlated in the data. I run the specification for probability of higher education after controlling for the agency and sense of belongingness indices[16] and the results from that regression are shown in column 5

[16] It is important to acknowledge here that although it would be better to control for these skills in the previous period, the data on them is not available for the older cohort in any other period that the latest round. Inclusion of these controls from the same period might result in reverse causality and bias the estimates. As mentioned above I am unable to say anything conclusive

in both the tables. Including psychosocial indices as controls increases the significance of the results and the agency index has significant, albeit modest, effects on the likelihood of higher education. However the effects aren't large and in this study I am unable to really test the relationship between psychosocial skills and higher study outcomes but it is nonetheless of importance to further research the causes of such a strong impact of private schooling on the probability of higher educational outcomes.

## 5. Conclusion

The ministry of Human Resource Development ("MHRD"), Government of India has announced that it will soon introduce the New Education Policy ("NEP") under the tag line *'Educate Encourage Enlighten'*.[17] Under this policy the government seeks to address the issue of low quality of primary and secondary education and promote innovation and research in institutions of higher study.

Given the failure of the public school system in meeting the need for quality education, understanding and disaggregating the effects of private schooling on the diverse measures of child development – be it cognitive achievement, psychosocial skill enhancement or prolonged persistence in higher education, is more important than ever. It is also becoming increasingly pertinent to understand the relationships between these different measures to be able to formulate policy that is complimentary and works to improve the various important spheres of child outcomes. The need of the hour is to understand if and how any positive effects seen in private schools can be extended to children throughout the public and private school systems to improve future outcomes for all children. In this study I take a small step in that direction by attempting to estimate the true effect of private school premium on these outcomes.

I do so by controlling for background characteristics, time use patterns, site fixed effects and lagged scores. I argue that the lagged scores act as a proxy for innate

about the relationship between psychosocial skills and higher study outcomes in my analysis but the importance of the same cannot be denied and should be pursued in further research.

[17] See the press release document by MHRD at:
http://mhrd.gov.in/sites/upload_files/mhrd/files/nep/press-releases.pdf.

ability, shocks and many other unobservable factors that selection might be based on. Value added models are being increasingly used to identify the effects of different inputs in the achievement production function. Specially in the case of estimating addition in terms of cognitive achievement, they are being accepted as being more efficient than alternative empirical strategies. However, even though I have sought to address as many causes for selection by using the uniquely rich Young Lives data, there might still be remaining channels that could cause endogeneity, as is the case with any other non-experimental study. The main benefit of the estimation strategy used in this study is that it allows me to determine the existing inequalities that arise out of the comparative advantage of private schools in creating better child outcomes, in the most efficient way possible, given non-experimental data. I argue moreover that experimental data in a way tests more for the effects of an intervention rather than the magnitude of the gap in the lack of such policy measures or interventions.

On estimating the effect of private schooling on additional cognitive achievement since the last round, I find a positive and significant premium for English in rural areas, for Mathematics in both urban and rural areas and PPVT in urban areas. I also estimate these effects of private schooling on cognitive achievement using propensity score matching as it allows for more flexibility in the functional form of the relationship between the dependent variable and the covariates than the dynamic OLS I use (dynamic OLS assumes this relationship is linear). The results from this estimation are almost the same as those in my main regression and thereby boost the confidence in my estimates. Any differences in the results from those of earlier studies can be interpreted as differences in learning trajectories for different subjects combined with increased robustness of my estimates by improving upon previous estimations and using richer data.

I also find a large, positive and significant effect of private schooling on agency and sense of belongingness indices of the older cohort. Given the importance of these measures in predicting future cognitive learning, persistence in education and earning levels of the children, these higher returns to private schooling could have a significant impact on the lives of these children. I discuss how the insignificant results for the younger cohort might simply indicate that non-cognitive skills grow later in life than cognitive skills and therefore the positive effect of private schooling on the same can only be seen for the older cohort. This means that as children grow older the inequalities in terms of psychosocial skills between public and private school children will only increase unless measures are taken to bridge this gap.

I then use a linear probability model of discrete choice to estimate the average effect of private schooling on the probability of the children in the older cohort progressing to the highest grade of secondary school education or to post-secondary, i.e. higher education. I find a significant and positive effect of private schooling on these outcomes. Although given that this is a measure of the 'average effect' these results must be taken as only an indicator of a probability of a premium from attending private schooling, given all else is equal. I also try to determine whether a part of this positive premium for higher study outcomes comes from the fact that private school children seem to have better psychosocial skills when compared to their public school counterparts even after controlling for background and ability (proxy). I find that agency is a significant, albeit modest predictor of persistence through these higher levels of school and post-school education. Future research should aim to accurately disentangle the effects of these different measures of achievement so that education policy focuses not only on full enrollment or even measures of cognitive achievement but further extends to other important aspects of school learning that determine the returns to schooling and the future outcomes of children.

# APPENDIX

Table A: Lagged Score and Ability:

| VARIABLES | Mathematics | | | English | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Private School | 0.22*** | 0.18** | 0.17** | 0.38*** | 0.35*** | 0.36*** |
| | (0.065) | (0.070) | (0.070) | (0.099) | (0.094) | (0.091) |
| Ability (Proxy) | | | 0.038** | | | 0.062* |
| | | | (0.017) | | | (0.031) |
| Controls: | | | | | | |
| Background | Y | Y | Y | Y | Y | Y |
| Time use | Y | Y | Y | Y | Y | Y |
| Lagged Score | N | Y | Y | N | Y | Y |
| Observations | 659 | 659 | 656 | 619 | 619 | 610 |
| R-squared | 0.428 | 0.446 | 0.452 | 0.392 | 0.405 | 0.410 |

Note: All controls are as mentioned in the main specification. Site fixed effects have been controlled for. All scores are generated using IRT. *** p<0.01, ** p<0.05, * p<0.1.

Table B. Cognitive Ability: Propensity Score Matching

| | Rural Areas | | | Urban Areas | | |
|---|---|---|---|---|---|---|
| VARIABLES | Math | English | PPVT | Math | English | PPVT |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Private | 0.20** | 0.34*** | -0.066 | 0.23 | 0.37 | 0.48* |
| | (0.085) | (0.12) | (0.11) | (0.16) | (0.31) | (0.26) |
| Observations | 659 | 619 | 1,146 | 141 | 139 | 252 |

Note: Propensity scores generated on the basis of background variables. Time use patterns not controlled for. Scores are generated using IRT as before. *** p<0.01, ** p<0.05, * p<0.1.

Table C:

| | |
|---|---|
| Agency | If I study hard at school I will be rewarded by a better job in future. |
| | If I try hard, I can improve my situation in life. |
| | I like to make plans for my future studies and work. |
| Self - Efficacy | I can always manage to solve difficult problems if I try hard enough. |
| | If someone opposes me, I can find the means and ways to get what I want. |
| | It is easy for me to stick to my aims and accomplish my goals. |
| | I am confident that I could deal efficiently with unexpected events. |
| | Thanks to my resourcefulness, I know how to handle unforeseen situations. |
| | I can solve most problems if I invest the necessary effort. |
| | I can remain calm when facing difficulties because I can rely on my coping abilities. |
| | When I am confronted with a problem, I can usually find several solutions. |
| | If I am in trouble, I can usually think of a solution. |
| | I can usually handle whatever comes my way. |
| Sense-of-Belongingness | I make friends easily. |
| | I am popular with kids of my own age. |
| | Most other kids like me. |
| | Other kids want me to be their friend. |
| | I have more friends than most other kids. |
| | I have lots of friends. |
| | I get along with other kids easily. |

## Bibliography


Andrabi, Tahir et al. "Do Value-Added Estimates Add Value? Accounting For Learning Dynamics". American Economic Journal: Applied Economics 3.3 (2011): 29-54. Web.

Annual Status of Education Report (ASER). 2014. Pratham Resource Center.

Bandura, Albert. "Perceived Self-Efficacy In Cognitive Development And Functioning". *Educational Psychologist* 28.2 (1993): 117-148.

Bandura, Albert. "Self-Efficacy Mechanism In Human Agency." *American Psychologist* 37.2 (1982): 122-147.

Bean, J. P. (1985). Interaction effects based on class level in an explanatory model of college student dropout syndrome. *American Educational Research Journal, 22*, 35–64.

Bialer, I. Conceptualization of success and failure in mentally retarded and normal children. *Journal of Personality*, 1961, 29, 303-320

Chetty, R., Friedman, J. and Rockoff, J. 2014, 'Measuring the Impact of Teachers I: Evaluating Bias in Teacher Value-Added Estimates', *American Economic Review*, 104(9): 2593-2632.

Chudgar, A., Quin, E., 2012. Relationship between private schooling and achievement: results from rural and urban India. *Econ. Educ. Rev.* 31 (4), 376–390.

Cueto, S., 2007. National and international evaluations of school achievement in Peru: Balance and Perspectives. In: Group for the Analysis of Development (Ed.) Research Politics and Development in Peru. GRADE, Lima, pp. 405–456.

Cunha et al. "Estimating The Technology Of Cognitive And Noncognitive Skill Formation". *Econometrica* 78.3 (2010): 883-931. Web.

Cunha, Flavio, and James J. Heckman. "Formulating, Identifying And Estimating The Technology Of Cognitive And Noncognitive Skill Formation". *Journal of Human Resources* 43.4 (2008): 738-782.

De, Anuradha., Claire Noronha and Meera Samson. 2001. "India: Private Schools and Universal Elementary Education." South Asia Education Sector, Technical Working Paper No. 3. World Bank.

Desai, Sonalde, Amaresh Dubey, Reeve Vanneman and Rukmini Banerji (2009), "Private schooling in India: A new educational landscape", *India Policy Forum*, Vol. 5, No. 1, pp. 1-38, April.

Ducette, J., & Wblk, S. Locus of control and extreme behavior. *Journal of Consulting and Clinical Psychology*, 1972, 39, 253-258.

Edmonds, Eric V. "Child labor." Handbook of development economics 4 (2007): 3607-3709.

Evans, David R. "Public Report On Basic Education In India: The PROBE Team, In Association With The Centre For Development Economics; Oxford University Press, New Delhi, 1999, 156 Pages, ISBN 0-19-564870-6.

Faircloth, B.S., Hamm, J.V., 2005. Sense of belonging among high school students representing four ethnic groups. Journal of Youth and Adolescence 34 (4), 293–309.

Findley, M.J., Cooper, H.M., 1983. Locus of control and academic achievement: a literature review. J. Pers. Soc. Psychol. 44 (2), 419.

Fiorini, Mario and Michael P. Keane. "How The Allocation Of Children's Time Affects Cognitive And Noncognitive Development". *Journal of Labor Economics* 32.4 (2014): 787-836. Web.

French, R and GG Kingdon (2010), 'The relative effectiveness of private and government schools in Rural India: Evidence from ASER data', Department of

Quantitative Social Science, Institute of Education, University College London, No. 10-03.

Grantham-McGregor, Sally et al. "Developmental Potential In The First 5 Years For Children In Developing Countries". *The Lancet* 369.9555 (2007): 60-70. Web.

Guarino, Cassandra et al. "A Comparison Of Student Growth Percentile And Value-Added Models Of Teacher Performance". *Statistics and Public Policy* 2.1 (2015): 1-11. Web.

Hanushek, Eric A. and Ludger Woessmann. "Do Better Schools Lead To More Growth? Cognitive Skills, Economic Outcomes, And Causation". *Journal of Economic Growth* 17.4 (2012): 267-321. Web.

Jamison, Eliot A., Dean T. Jamison, and Eric A. Hanushek. "The Effects Of Education Quality On Income Growth And Mortality Decline". *Economics of Education Review* 26.6 (2007): 771-788. Web.

Juvonen, J., 2006. Sense of belonging, social bonds, and school functioning. In: Alexander, P., Winne, P. (Eds.), *Handbook of Educational Psychology.* second ed. Macmillan, New York, pp. 655–674.

Kingdon, G. (1996a) 'Private Schooling in India: Size, Nature and Equity Effects', *Economic and Political Weekly* 31.51: 3306–14.

Kingdon, G. (1996b) 'The Quality and Efficiency of Public and Private Schools: A Case Study of Urban India', *Oxford Bulletin of Economics and Statistics* 58.1: 55–80.

Kingdon, G. G. "The Progress Of School Education In India". *Oxford Review of Economic Policy* 23.2 (2007): 168-195.

Lee, Doo Won and Tong Hun Lee. "Human Capital And Economic Growth Tests Based On The International Evaluation Of Educational Achievement". Economics Letters 47.2 (1995): 219-225. Web.

Mehta, Arun. 2005. “Elementary Education in unrecognized schools in India: A study of Punjab based on DISE 2005 data” NIEPA, New Delhi.

Muralidharan, K. and M. Kremer (2008) ‘Public and Private Schools in Rural India’, in P. Peterson and R. Chakrabarti (eds) School Choice International, Cambridge, MA: MIT Press.

Muralidharan, Karthik, and Venkatesh Sundararaman. "The Aggregate Effect Of School Choice: Evidence From A Two-Stage Experiment In India". *The Quarterly Journal of Economics* 130.3 (2015): 1011-1066.

Pascarella, E. T., & Chapman, D. (1983). A multi-institutional, path analytic validation of Tinto's model of college withdrawal. *American Educational Research Journal, 20*, 87–102.

Pascarella, E. T., & Terenzini, P. T. (1983). Predicting voluntary freshman year persistence/withdrawal behavior in a residential university: A path analytic validation of Tinto's model. *Journal of Educational Psychology, 75*, 215–226.

PROBE (1999) Public Report on Basic Education in India, Delhi: Oxford University Press.

Ramachandran, V., K. Jandhalay, N. Mehrotra, L. Krishnamurthy, V. Periodi and A. Siahjee (2004) Snakes and Ladders: Factors Influencing Successful Primary School Completion for Children in Poverty Contexts, South Asian Human Development Sector Report No. 6, New Delhi: World Bank.

Robbins, Steven B. et al. "Do Psychosocial And Study Skill Factors Predict College Outcomes? A Meta-Analysis.". *Psychological Bulletin* 130.2 (2004): 261-288. Web.

Schwarzer, R. (Ed.) (1992). Self-efficacy: Thought control of action. Washington, DC: Hemisphere.

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston, Measures in health psychology: A user’s portfolio. Causal and control beliefs (pp. 35- 37). Windsor, England: NFER-NELSON.

Singh, Abhijeet. "Private School Effects In Urban And Rural India: Panel Estimates At Primary And Secondary School Ages". *Journal of Development Economics* 113 (2015): 16-32.

Tinto, V. (1993). Leaving college: Rethinking the cause and cures of student attrition (2nd ed.). Chicago: University of Chicago.

Tooley, J., P. Dixon and S.V. Gomathi (2007) 'Private Schools and the Millennium Development Goal of Universal Primary Education: A Census and Comparative Survey in Hyderabad, India', Oxford, *Review of Education* 33.5: 539–60.

Tooley, James and Pauline Dixon. 2003. "Private Schools for the Poor: A Case Study from India", CfBT Research and Development.

Van der Linden, W. J. & Hambleton, R. K. (1997). *Handbook of item response theory.* New York, Springer-Verlag.

Walsh, D. A. (2005). Why Do They Act That Way? A Survival Guide to the Adolescent Brain for You and Your Teen (1 ed.). New York: Free Press.

Willms, J.D., 2003. Student Engagement at school. A sense of Belonging and participation Results From PISA 2000, OECD. From <http://www.oecd.org/dataoecd/42/35/33689437.pdf>.

World Bank (2006), 'Secondary Education in India: Investing in the Future', Human Development Unit, South Asia Region, World Bank, draft.