

Does Medium of Instruction matter? A propensity score analysis
of the effect of English Medium Education on students learning
outcomes in India

by

Sandeep Shankar Murugesan

MPP Essay

Submitted to Oregon State University

In partial fulfillment of the requirements for the degree of Master of Public

Policy

Master of Public Policy essay of Sandeep Shankar Murugesan presented March 27th, 2015

Approved:

Dr. Victor Tremblay, Economics

Dr. Alison Johston, Political Science

Dr. Todd Pugatch, Economics

Sandeep Shankar Murugesan, Author

ABSTRACT

Over the past two decades there has been sharp increase in the demand for English-language skills in India. This led to a rapid expansion of English Medium Schools, more so in the private sector than the public sector. Recent studies have estimated the economic benefits of English- language skills but few studies have compared the educational performance of students from the English Medium Schools (EMS) versus Regional Medium Schools (RMS). This study seeks to address this gap in the literature. Specifically, this paper uses propensity score matching methods to examine whether students between 8 and 11 years old have higher reading, mathematics and writing scores when they attend EMS instead of RMS. The study makes use of the data from the Indian Human Development Survey (IHDS) that was collected by the National Council of Applied Economic Research (NCAER) in collaboration with University of Maryland in 2005. The estimates from the propensity score matching analyses suggest that there is no significant difference in learning outcomes between the two groups. Based on the findings, I further discuss the policy implications, limitations of the study and the potential areas for future research.

Table of Contents

1. Introduction	7
2. Background.....	9
2.1. Colonial History and Post Independence Language Policy	9
2.2. Post-Liberalization shift in priorities.....	11
2.3. Private Vs Public Schools Debate	12
3. Literature Review	13
3.1. The link between language of instruction and learning outcomes	14
3.2. Gender Discrimination in Educational Choice.....	15
3.3. Economic Class and differential access to Primary Education	16
3.4. Government Programs and Primary School Enrollment.....	17
3.5. The influence of Caste on Education	17
3.6. Schooling and the Urban Rural Divide in access to education	18
3.7. Household Composition and Choice of Education	19
3.8. Parental Education and Choice of Education.....	19
4. Empirical Framework	20
5. Data.....	24
6. Data Analysis.....	30
6.1. Model Validation.....	31
6.2. Matching procedures	32
7. Discussion.....	33
7.1. Characteristics of English Medium School Students	33
7.2. Propensity Score Model Results Discussion	37
8. Limitations.....	39
9. Policy Recommendations	41
<i>References</i>	43

List of Tables

1.	Descriptive Statistics for Regional Language Schools Vs English Medium Schools	26
2.	Basic Regression Model	30
3.	Characteristics of English Medium School Students Logit Model.....	34
4.	Net Impact of EMS on Students computed using different Matching Techniques	37
5.	Descriptive Statistics of differences in student inputs between EMS and RMS	38
6.	Net Impact of EMS on student's time allocation for Home Work	39

List of Figures

1.	Student Reading Skills	49
2.	Student Mathematics Skills.....	49
3.	Student Writing Skills.....	50
4.	Student Enrollment Status.....	50
5.	Comparison of Percentage of Students Enrolled in English Medium Schools between Public and Private Schools.....	51
6.	Student Enrollment in Different Schools.....	51
7.	Comparison of Student Enrollment in Private and Public Schools	52
8.	Balancing Property Test for Model 1.....	52
9.	Balancing Property Test for Model II.....	52
10.	Overlap Test for Model 1.....	53
11.	Overlap Test for Model II.....	54
12.	Distribution of propensity score of Matched Samples.....	54

List of Abbreviations

IHDS	Indian Human Development Survey
MI	Medium of Instruction
EMS	English Medium School
RMS	Regional Medium School
PSM	Propensity Score Matching
NN	Nearest Neighbor
EGS	Education Guarantee Scheme
VEC	Village Education Committee
SC	Scheduled Caste
ST	Scheduled Tribe
OBCs	Other Backward Classes
ASER	Annual Survey of Education Report
CBSE	Central Board of Secondary Education
NGO	Non-Governmental Organization
NCAER	National Council of Applied Economic Research
NKC	National Knowledge Commission
UPA	United Progressive Alliance
TLF	Three Language Formula

1. Introduction

For more than two decades, scholars have debated the impacts of globalization on the economy, culture, politics and society. One of the significant outcomes of globalization is the emergence of English as the lingua franca of the entire world. In recent years, scholars have concentrated on the impact of globalization on the language policy in the developing world (Chang, 2006; Her, 2007; Hornberger & Vaish, 2009; Kirkgöz, 2009). Undoubtedly, the emergence of English as the ‘global language’ poses both social and economic challenges in many parts of the world and policy makers around the world are beginning to grapple with this change.

In a global market there are significant economic benefits of learning English. In fact, many studies indicate that there is a high “wage premium” for English skills in the labor market (Azam, Chin, & Prakash, 2013; Kapur & Chakraborty, 2008; Munshi & Rosenzweig, 2003). The growing demand for learning English medium schools¹ is partly driven by the expectation of earning a “wage premium” in the labor market.² From a public policy perspective, improving the access to English education is important for two main reasons. First, English skills enhances the capability of individuals to participate in the global economy. Second, it can address the growing wage inequality which arises due to differential access to quality English training. However, the question still remains whether the government should expand public English medium schools. The argument in favor of expanding English medium schools is that it can improve the access to learning English. On the other hand, there are genuine concerns over shifting language policy in favor of promoting English medium schools. For instance, education

¹ By definition an “English Medium School” is one in which ‘English’ is the primary medium of pedagogy. The same logic applies to the Regional Medium Schools as well.

² Data collected by the National University for Education Planning and Administration indicates that those opting for English Medium Education have increased by 150% between the years 2003 to 2008.

psychology literature shows that for children may be better-off if they are taught in their native language (Abadzi, 2006). Kosonen (2005) argues that when children are offered opportunities to learn in their native language, they are more likely to attend and succeed in school. Also, some studies show that when children learn in their native language the parents are more likely to participate in their children's learning (Benson, 2002). Therefore if English medium education has shown to have adverse effects on learning outcomes, then students from disadvantaged backgrounds would be put at a "double disadvantage". Unfortunately, few studies have estimated the impact of English Medium Education on the student learning outcomes in India. This paper seeks to address this gap in the literature.

In this paper I use the nationally representative student-level dataset (Desai, Vanneman, & National Council of Applied Economic Research, 2013), to estimate the impact of English Medium Schools on student learning outcomes. Since the expansion of EMS has been predominantly in the private sector, only students between the age 8 and 11 years, who are currently enrolled in private schools,³ were included in the study sample. The challenge in estimating the impact of Medium of Instruction on learning outcomes is to overcome the endogenous factors such as student motivation and teacher skills which can affect learning outcomes. Propensity score matching analysis allows us to partly overcome this problem, although it is no "magic bullet" to solve all problems concerning endogeneity. In short, the propensity score matching method *ex post* tries to recreate the conditions of a randomized control trial. This estimates obtained by using this method are shown to be an improvement over those that are derived from traditional regression techniques (Dehejia & Wahba, 2002).

³ "Private Schools" include Private recognized and unrecognized Schools, Government Aided Schools, Convents, Madrassas and junior college.

The results from the propensity score analysis suggest that there is no significant difference in the learning outcomes between the students who attend EMS and RMS. This shows that the medium of instruction does not appear to influence learning outcomes. The results of this study can have important implications on the debate on language policy. In the later sections I discuss the limitations of the study and then, based on the findings of this paper, I discuss why a more pragmatic policy is required in dealing with the question of language and point out the possible areas for future research.

2. Background

2.1. Colonial History and Post Independence Language Policy

India's tryst with English dates back to the early nineteenth century, when the British formally introduced English education. Viswanathan (2014) argues that the British education policy was "set out to create a middle class serving as an agency of imperialist economy and administration and, through it, to initiate social change through a process of differentiation". During the British Raj, English evolved as the language of the Indian elite and it came to be associated with power and privilege. Therefore, historically speaking, English has played a huge role in shaping India's political, social and economic landscape.

The "British Raj" ended in 1947. Post Independence, the Constitution Assembly debated the role of English in free India. There was a general feeling that an Indic language should be made the official language. There were calls for Hindi, which was spoken by the majority, to be made the national language⁴ but it was not accepted by the non-Hindi speaking members of the

⁴ Many members of the Constituent Assembly, especially those from the Southern States opposed the imposition of Hindi as the national language. T.A.Ramalingam Chettiar, a member from the State of Madras, told the Assembly "you cannot use the word national language, because Hindi is no more national to us than English or any other language. We have got our own national languages."(Volume IX, Constituent Assembly Debates)

Constituent Assembly. The members reached a compromise and it was decided that both English and Hindi shall be the official languages of the Central Government, each state shall have its own official language and the communication between the Centre and States shall be in English and Hindi. The implication of this decision was that by default all bureaucrats were expected to have a working knowledge of English. Moreover, in practice, English continues to be the link language between the Hindi speaking and non-Hindi speaking population in India.

As far as the language policy with respect to schools was concerned, the government adopted a “three language formula”. And these three languages were to be taught as school subjects, regardless of the Medium of Instruction, namely: 1) mother tongue or regional language; 2) Hindi or English; 3) one modern Indian language or foreign language not covered under 1 and 2. By definition, an English medium school is one in which the core curriculum is taught in English. In addition to that, the students are taught a secondary language which is either the regional language or Hindi. Since education was included as a State subject, the “three language formula” (TLF) was not uniformly implemented across all the states. For instance, only two languages are being taught in the public schools in the State of Tamil Nadu —namely Tamil and English. According to the Seventh All India Education Survey (2002), 90.61% and 84.86% schools followed the TLF in the upper primary and secondary stage respectively. The survey also reveals that 55.05% and 54.1% schools taught English as the second language in the upper primary and secondary stage respectively.⁵

⁵ In most of the states Classes I-IV/I-V constitute ‘primary stage’; Classes V-VII/VI-VII/VI-VIII constitute ‘upper primary stage’; Classes VIII-X/IX-X constitute ‘secondary stage’; and Classes XI-XII as ‘higher secondary stage’. Source: 7th All India Education Survey

2.2. Post-Liberalization shift in priorities

India dismantled the “License Raj”⁶ and shifted to a more open economy in the early 1990’s. Over the past two decades, there has been a tremendous growth in India’s Information technology sector. It has grown from contributing around 2% of the GDP in the late nineties to 4.8% at the end of 2006 and now stands nearly around 7.5%. This growth was fueled by an English-speaking educated middle class, which is currently estimated to be around 300 million.

Since the service sector has been the main driver of India’s economic growth and generates relatively high paying jobs, the demand for English education has increased tremendously. For instance, a recent study by Azam et al (2013) reveals that the hourly wages for men who speak fluent English 34% higher than those who did not know English. This has forced the Government to rethink its language policy. In 2006, the United Progressive Alliance (UPA) appointed the National Knowledge Commission (NKC) to study the problem and make a set of recommendations. The NKC report recommended making English as a compulsory language and advised the government to expand EMS in the public sector (India, 2007). According to 7th All India Education Survey (2002), 12.98%, 18.25%, 25.84% and 33.59% schools used English as the medium of instruction at the primary, upper primary, secondary and higher secondary stage respectively. The corresponding figures from 6th Survey (1993) were 4.99%, 15.91%, 18.37% and 28.09% respectively.

The post-liberalization phase of India’s economic policy was also marked by the failure of the government to provide universal quality public education. This fueled a rapid expansion in private provision of primary schools throughout India. The state of public schools was so

⁶ Post Independence, India adopted Socialistic Economic Policies. The State curbed the growth of private sector by instituting severe licensing requirements for producing goods and services within the country.

abysmal that the PROBE Team (1999) study reported that "even among poor families and disadvantaged communities, one finds parents who make great sacrifices to send some or all of their children to private schools, so disillusioned are they with government schools". This led to a situation when parents started to send their children to fee charging private schools rather than to the free of cost public schools. In 1993 approximately 10% of children aged 6–14 were enrolled in a private school. By 2008 an around 22.5% of young school children were enrolled in a private school (Cheney, Ruzzi, & Muralidharan, 2005).

2.3. Private Vs Public Schools Debate

In the recent past, many studies have compared the performance of private and public schools in India (Muralidharan, 2006; Muralidharan & Kremer, 2006; Muralidharan & Sundararaman, 2013b) These studies reveal that despite the fact private schools in India use fewer resources, students who attend private schools outperform their counterparts in public schools on almost all parameters. A recent study revealed that the rate of teacher absenteeism in public schools is higher than that of private schools (Kremer, Chaudhury, Rogers, Muralidharan, & Hammer, 2005). Therefore "teacher absenteeism" and poor instructional quality could be among the most important factors that explain the performance gap between private and public schools in India. A randomized control trial study to evaluate the effect of teacher performance pay incentive seems to indicate that those students who are assigned to the incentive classroom performed much better than those students in the control group (Muralidharan & Sundararaman, 2009) On the other hand, private school teachers are usually less qualified and paid less than one third the salaries compared to teachers in public schools. The same study reveals that 28% of the population in rural India has access to fee-charging private primary schools in the same village, and that 16.4% of children aged 6 – 14 in rural India attend fee-charging private schools.

Moreover, private schools are more likely to have more contract teachers⁷ hired within local areas than public schools (Muralidharan & Sundararaman, 2013a). A more recent study has also made the case that there is little evidence of “cream skimming”⁸ in private schools in India (Tabarrok, 2013).

This study excludes the students who attend public schools for three main reasons 1) There is a huge performance gap between private and public schools in India due to the reasons which we have discussed in the previous section. 2) The expansion of EMS has been mainly in the private sector. With our overall sample we can observe that among those students who have enrolled in public schools only 2.66% students attend public EMS (see Figure 5) 3) The PSM method is a data intensive technique and it can only match on observable differences. Including only private school students helps to match the unobservable variables that might lead parents to choose private instead of public schools. There are fewer differences between private English Medium Schools and private Regional Medium Schools. Therefore such a comparison can produce more accurate results in the propensity score analysis.

3. Literature Review

In this section, I will first discuss the theoretical and empirical link between the student learning outcomes and the medium of instruction. After that I will discuss the variables which affect parental school choice. These variables include gender of the student, economic class of the household, caste, government incentives, the neighborhood of the household (urban or rural), parental education and household composition.

⁷ “Contract teachers” broadly refers to those teachers recruited by the community (though not always) on a contractual basis and on a fixed honorarium to overcome the problem of teacher shortage and teacher absenteeism in rural and remote areas

⁸ “Cream Skimming” refers to pattern of school enrollment wherein the best students from the public schools are absorbed into private schools.

3.1. The link between language of instruction and learning outcomes

The study examining the link between the language of instruction and learning outcomes has several antecedents. The education psychology literature shows that for children may be better-off if they are taught in their native language (Abadzi, 2006). The evidence from the literature suggests that mother tongue instruction in the early years of childhood is can avoid cognitive disadvantages (James Cummins, 1978; Jim Cummins, 1979). This significance is more for children who come from disadvantaged backgrounds. Kosonen (2005) argues that when children are offered opportunities to learn in their native language, they are more likely to attend and succeed in school. Also, some studies show that when children learn in their native language the parents are more likely to participate in their children's learning (Benson, 2002).

On the other hand, there is also evidence from the psychobiological literature which shows that younger children learn languages more easily than adolescents and adults. This is referred to as the "critical period hypothesis" by cognitive scientists (Hakuta, Bialystok, & Wiley, 2003; Johnson & Newport, 1989). The application of this theory would suggest that if children are exposed to the new language at a very young age they will acquire the language skills more easily and therefore this will not have any negative impacts on the learning outcomes of students.

According to Human Capital Theory, improvement in "knowledge stock and learning capabilities" of the population would have a significantly positive impact on the overall economy (Foray & Lundvall, 1998). Using this framework, within the economics of education literature, it there is evidence to suggest that the "medium of instruction policy in education" does have an effect on human capital formation. R. Ramachandran (2012) provides evidence from Ethiopia showing that a switch to mother tongue instruction for primary school led to a significant increase in student educational attainment. Likewise, evidence from Yoruba indicates that

“higher repetition rates, dropout rates and overall lower achievement” can be partly explained by difference in the medium of instruction (Bamgbose, 2005). In the United States, Thomas & Collier (1997) analyzed the impact of Bilingual Schools versus English Schools on language minority students over the period 1985-2001. The study revealed that language minority students in English schools performed poorer in English tests and had higher dropout rates and lower educational attainment compared their counterparts in Bilingual Schools.

In the Indian context there are not many studies which have empirically examined the link between the medium of education and student learning outcomes except for Muralidharan & Sundararaman (2013b). Although this study cannot test the accuracy of these theories in general, it can throw some light on the state of English medium schools in comparison to the regional medium schools.

RQ: Does English Medium Education have a negative impact on Student Learning Outcomes?

3.2. Gender Discrimination in Educational Choice

The phenomenon of gender discrimination in intra-household allocation of resources in India has been documented extensively (Alderman & King, 1998; Duraisamy & others, 1992). A more recent study has pointed out that “girls experience gender discrimination especially from age 10 onwards, with almost universal disadvantage in the amount of education expenditures in the group of 15-19 year olds.” (Zimmermann, 2012) Moreover, gender bias tends to be more acute in rural areas when compared to urban areas (Azam & Kingdon, 2013). In general, parents tend to spend allocate more resources for the education of the male child compared to the female child. The differential treatment of the female child can be due to social and economic reasons.

(Dreze & Sen, 2003) argue that “entrenched belief of gender division of labor” is an important factor for gender discrimination. Results from Gandhi Kingdon (2002) reveal that a plethora of factors influence girls educational attainment—namely “parental background, wealth, opinions, individual ability, age-at-marriage and quality of primary education”. Based on the findings of these studies, we can expect that girls will be less likely to attend EMS due to gender discrimination in Indian households.

3.3. Economic Class and differential access to Primary Education

The causal link between economic status and access to quality education is quite robust and there is copious literature to support the hypothesis. Geeta G. Kingdon (1996) and Filmer & Pritchett (2001) reveal that on average a “rich” child is 31 percentage points more likely to be enrolled than a “poor” child. The study goes on to add that this gap is not even in all States. For instance, the State of Kerala the gap is just 4.6% whereas Bihar is 42.6%. However, recently there has been a sharp rise in the growth of “low cost” private schools in India. Even though free government schools are available, parents tend to send their children to fee-levying private schools (Muralidharan & Kremer, 2006). In India, there is a gulf between regulation “on paper” and regulation “in practice”. Dixon & Tooley (2005) show how actual regulation on paper can stifle innovation and entrepreneurship in the education sector and that in practice the growth of low cost private schools remains “extra legal”. Muralidharan & Kremer (2006) argue that nearly 53% of private schools are “unrecognized” by the government. But the question whether even the low cost private schools is accessible to the economically and socially disadvantaged children still remains relevant. Härmä (2011) shows that low cost schools are still unaffordable to the bottom two wealth quintiles families and argue that “increased reliance on a market in education will not help to achieve equitable access to primary schooling for all”.

3.4. Government Programs and Primary School Enrollment

The Constitution of India guarantees every child, aged between 6 and 14, the right to free and compulsory education. In order to achieve universal primary education, in the past, different governments have designed various government programs aimed at achieving this goal. Besides the Central government, various State governments provide financial and material incentives for students. As per the Fifth All India Education Survey (1986), throughout India, out of the 5.29 lakh primary schools, 1.47 lakh schools provide mid-day meals, 2.48 schools provide free uniforms and 3.13 lakh schools provide free textbooks. These incentives are intended to enhance primary school enrollment. Although there are not many empirical studies examining the effectiveness of the free uniform and free text book programs in India, there are some studies which have analyzed the impact of mid-day meals (MDM) program. Afridi (2011) points out that, in Rural Areas, the mid-day meals program has been successful in “improving participation rates of girls thereby reducing gender disparity in schooling.” Likewise, a study based on Karnataka’s experience shows that the MDM program has a positive effect on enrollment, attendance, dropout rate and retention rates, and a marginal impact on student’s scholastic performance (Laxmaiah et al., 1999). Besides these educational parameters, the MDM scheme is shown to have positive impact on the nutritional factors and the general health of the students (Singh, Park, & Dercon, 2014). Based on the results of these studies and given the high demand for EMS, we would expect government incentives to have a positive impact on enrollment into EMS.

3.5. The influence of Caste on Education

The caste system plays a huge role in determining social expectations in India. Although in modern times the influence of caste is waning primarily due to globalization and urbanization, it

still places an important role in Indian society. Borooah & Iyer (2005) show that, when the parents are illiterate, the “community effects” are more pronounced on the educational outcomes of the student. In recent times, the analyses of Munshi & Rosenzweig (2003) is a seminal contribution to the understanding of influence of caste in modern Indian society. The study makes the following conclusion:

Caste continues to play a particular (gender-specific) role in shaping schooling choices in the new economy of the 1990s. But the overall increase in English schooling in recent years, and the growing mismatch in education choices and hence occupational outcomes between boys and girls in the same caste, suggest that the remarkably resilient caste system might finally be starting to disintegrate.

In a globalized world English education is being viewed as an instrument of social mobility. This seems to be the case across all caste based communities. A recent newspaper reported that the Dalits in Uttar Pradesh have constructed a temple dedicated to the “Goddess of the English Language”. Dalit activist and intellectual, Chandra Bhan Prasad remarked.

She (Goddess of English) holds a pen in her right hand which shows she is literate. She is dressed well and sports a huge hat - it's a symbol of defiance that she is rejecting the old traditional dress code. In her left hand, she holds a book which is the constitution of India which gave Dalits equal rights. She stands on top of a computer which means we will use English to rise up the ladder and become free forever. (P, News, village, & Pradesh, n.d.)

Although it is just a piece of anecdotal evidence, it is symbolic of how the socially disadvantaged communities perceive the value of English education.

3.6. Schooling and the Urban Rural Divide in access to education

Although there is a clear divide between rural and urban private school enrollment rates, Tilak (2001) argues that “the relative size of both the government and the government-aided sectors seem to be shrinking and that of the private unaided sector is increasing”. Muralidharan

& Kremer (2006) argue that approximately 28% of the rural population have access to fee charging private schools and around 16.4% of children aged 6 – 14 in rural India attend fee-charging private schools. Drèze & Kingdon (2001) show that parental education and motivation, the distance of the school, the quality of the school, work opportunities, village development, teacher postings, teacher regularity and mid-day meals, as the basic factors which determines a child's participation in school in rural India. Kochar (2004) argues that the gap between rural and urban schooling can be explained by not just conditions in the local village economy, but also the functioning and size of the relevant labor market. The study shows that among at least the landless laborers, the schooling choice reflects the possibility of employment in urban areas. Based on the results of these studies, we expect students from urban neighborhood to be more likely to attend EMS.

3.7. Household Composition and Choice of Education

India is going through a demographic transition. It has witnessed a secular decline in the fertility rates and the size of the family (Drèze & Murthi, 2001). Intra-household resource allocation depends on the composition of the household. With respect to educational choice, studies show that household size and the number of children in the household are negatively correlated with the choice of schools (Desai, Dubey, Vanneman, & Banerji, 2009). On the other hand, *ceteris paribus*, female headed household (especially from backward castes) are more likely to live in poverty, more so in rural areas and studies have shown that children from female headed household in rural areas are less likely to attend school (Ray, 2000).

3.8. Parental Education and Choice of Education

The causal link between parental education and child schooling has been widely recognized in the literature (Desai et al., 2009; Dreze & Kingdon, 2001; Duraisamy & others, 1992;

Kambhampati & Pal, 2001). There is also evidence of “same sex effects”. In other words, maternal education level has a more significant effect on the girl’s school participation and the paternal education level on the boy’s (Dreze & Kingdon, 2001; Kambhampati & Pal, 2001). The link between the demand for English Medium Education and parental education is also recognized in the literature. (Kapur & Chakraborty, 2008)

4. Empirical Framework

Ideally, a randomized experimental design is the best approach to calculate the average treatment effect of EMS. However, conducting such large scale social experiments is not always practicable. The standard non-experimental regression techniques are likely to be inaccurate since data for all relevant variables are not easy to obtain. Alternatively, a simple comparison between the average outcomes of students from EMS with those from RMS will obviously suffer from “selection bias” because the students are not randomly assigned to the schools. To overcome the problem of selection bias there are various empirical strategies that can be used to adjust the systematic differences between the treatment and control groups. In this paper I make use of “propensity score matching” (PSM) method. The advantage of using the PSM method is that it allows us to make causal inferences even in a non-experimental setting, conditional on the basic assumptions being met. The estimates obtained by this method are shown to be an improvement over other non-experimental estimates and are much closer to the estimates obtained from experimental studies (Dehejia & Wahba, 2002). However, more recent literature on this subject points out that this method can be effective only if it satisfies certain basic conditions. (Smith & Todd, 2005)

According to Rosenbaum and Rubin (1983), “the propensity score is the estimated conditional probability of assignment to a particular treatment given a vector of observed

covariates”. Based on the propensity scores of each sample, the statistical model generates a comparison group that has similar characteristics to those in the treatment group, except for the fact that they do not get the treatment. In other words, this comparison group is similar to the “control group” in an experimental design. The probability of being a part of the “treatment group” or “control group” is predicted as a function of a set of observed covariates that influence the program participation (in this case it is student enrollment in English Medium School) and outcomes. A PSM estimator pairs each program participant with a non-participant and the difference in the mean of the outcome variables of interest between both groups can be interpreted as the “average treatment on the treated” of the particular program.

The underlying basic assumptions in the PSM method are:

- a) The assignment to a treatment only depends on the observable pre-intervention variables. In other words, after controlling for the covariates, the treatment is as “good as random”. This is known as the ‘unconfoundedness’ or ‘selection on observables’ assumption.

$$(Y_0, Y_1) \perp D | X$$

Where ‘ Y_0 ’ and ‘ Y_1 ’ is the potential outcome of untreated and treated individuals, ‘ D ’ is the treatment and ‘ X ’ is the set of observed covariates.

- b) The probability of assignment to a treatment is bounded away from zero and one, otherwise known as “overlap” assumption.

$$0 < P(D = 1|X) < 1$$

If the two assumptions are not met, then the PSM estimators are likely to be biased. In practice, the “unconfoundedness” assumption can never really be tested. Therefore the choice of

variables to be included in the logit model is very important. As a rule, only those variables which affect both the program participation and the outcome must be included in the model. Therefore the choice of variables should either be based past research or should have strong theoretical backing. More recent studies stress the importance of avoiding variables which can be affected by the “anticipation of participation in the treatment”, and makes a strong case to avoid “over-parameterization” (Caliendo & Kopeinig, 2008).

This study exploits the data at the *student level* i.e. the unit of analysis is the individual. The first step of the PSM method is the calculation of the propensity scores of every sample. The propensity score is derived from a logit regression in which the outcome variable of interest is whether a student is enrolled in an EMS. The determinants of EMS enrollment can be broadly classified into four categories, namely 1) Student Characteristics 2) Household Characteristics 3) School Characteristics and 4) Socio-Religious Characteristics. Within each category there are a set of variables, which in detail will be covered in the “Data” section of this paper. The baseline logit model is depicted in the following equation:

$$Y_i = \alpha + \sum_{l=1}^5 \beta_l \text{StudentCharacteristics}_{li} + \sum_{n=1}^6 \delta_n \text{SchoolCharacteristics}_{ni} + \sum_{m=1}^9 \gamma_m \text{HouseholdCharacteristics}_{mi} + \sum_{o=1}^7 \theta_o \text{SocioReligious}_{oi} + \varepsilon_i$$

Where ‘ Y_i ’ represents the dependent variable and it is a dummy in which ‘1’ stands for EMS and ‘0’ stands for RMS. ‘ ε_i ’ represents the error term.

In this study we have three outcome variables namely mathematics, reading and writing scores. The baseline logit model will remain the same for all the three outcome variables. After

estimating the propensity scores based on the logit model, the next step is to choose a matching method. There are various types of matching methods. Depending on the algorithm of the matching method, a individual from the treatment group is compared to one or more individuals from the control group.

The “Nearest Neighbor” method every individual from the treatment group will be compulsorily matched with the nearest individual from control group, and in the next step the difference between the outcomes will be computed. And finally, the “Average Treatment on the Treated” is obtained by computing the average of these differences. The drawback in this method is that the routine matches the nearest neighbor even if the propensity score is significantly different from one another and thereby could result in poor matches. This problem can arise if, for example, there are too many individuals with high propensity scores in the treatment group and few individuals in the control group with high propensity scores; as a result the NN matches individuals with relatively low propensity scores with high propensity score individuals. In such a scenario, the NN method will not produce the most reliable estimates and therefore the researcher should choose alternative algorithms. (Caliendo & Kopeinig, 2008)

The stratification matching algorithm partitions the common support region of the propensity score into a set of intervals and computes the mean of difference in outcomes between treated and control observations to estimate the ATT of a particular program (Rosenbaum & Rubin, 1983). An advantage of this method is that the outcome of all individuals in the treatment and control group are factored in estimation of the treatment effect. The disadvantage of this method is that some strata may contain a relatively small number of individuals in the treatment group compared to control group members and vice versa.

Kernel matching algorithm compares the outcome of each individual in the treatment group to a weighted average of the outcomes of all individuals in the control group. The highest weight is attached to those with scores nearest to the treated individual. Since more information is used, one advantage of this method is that the variance is very low. This method has the same drawback as the NN matching method i.e. some observations may have poor matches (Smith & Todd, 2005).

Each matching algorithm has its own merits and demerits. Therefore selecting any one method can be a problem if different methods produce significantly different results. To overcome this problem, in this study I make use of three matching methods namely Nearest Neighbor, Kernel and Stratified Matching.

5. Data

The data used for this study was collected by the National Council of Applied Economic Research in collaboration with University of Maryland in 2005 (Desai et al., 2013). The Indian Human Development Survey (IHDS) is a multi-topic, nationally representative survey (with the exception of Lakshadweep and Andaman and Nicobar islands) covering 41,554 households from 1503 villages and 971 urban neighborhoods across India. The data was collected using stratified random sampling procedure to ensure linguistic, religious and caste based subpopulations are adequately represented. The topics covered in this survey include health, education, employment, economic status, marriage, fertility, gender relations, and social capital.

As a part of the survey exercise, children aged between 8 and 11 in the household were administered reading, writing, and arithmetic knowledge tests. The test questions were designed in collaboration with researchers from PRATHAM, one of India's leading non-governmental organizations (NGO) working in the field of primary education. Prior to administrating the test,

the surveyors were trained by Pratham. The tests were developed in 13 Indian languages and standardized to ensure that the results can be compared across different languages. The students were allowed to take the test in the language they were most comfortable in. and based on the test results, IHDS classified student's capability into different categories.

The student's reading capability was classified into five categories: (1) Cannot read at all (2) Can read letters (3) Can read words (4) Can read paragraph (5) Can read a one-page short story. Similarly, a student's arithmetic skills was classified into four categories: (1) Cannot read numbers above 10, (2) Can read numbers between 10 and 99 but cannot do more complex number manipulation, (3) Can subtract two-digit numbers, and (4) Can divide a number between 100 and 999 by another number between 1 and 9. And finally, the writing test classified students into two categories: (1) Cannot write, (2) Writes with two or fewer mistakes. Figures 1, 2 and 3 show the students reading, arithmetic and writing skills in each category level. The reading, arithmetic and writing skills of the students are the dependent variables in this study.

As discussed earlier, post 1991 India has witnessed a rapid expansion privately run schools and that the trend was not just restricted to urban areas. The IHDS survey uses the 2001 census to make the distinction between rural and urban areas. Figure 7 shows the private school enrollment rate in rural and urban areas within our study sample. We can observe that about 57% in urban students and 22% in rural students attend private school. This should not come as a surprise since many studies have confirmed this phenomenon earlier (Drèze & Kingdon, 2001; Geeta Gandhi Kingdon, 2007). Since this study is concerned with only private schools investigating the impact of private English medium schools on students learning outcome, only the test results of children between the age of 8 and 11 who are currently enrolled in a school will be included in the study sample.

The main dependent variable is a dichotomous variable which indicates whether the student has enrolled in an English medium school (EMS) or a regional medium school (RMS). By definition an “English Medium School” is one in which ‘English’ is the primary medium of pedagogy. The same logic applies to the Regional Medium Schools as well. Post Independence, Indian most states were re-organized on linguistic basis. Therefore each state has a dominant regional language. To simplify our analysis, all RMS are coded as ‘0’ and EMS are coded as ‘1’. In our overall data, out of 11,060 “currently enrolled” students only 1339 are enrolled English medium schools (See Figure 4) and out of the 1339 students who have enrolled in English medium schools only 214 are in government run English medium schools. In our study sample we consider only those students who are currently enrolled in private schools, out of a total of 3731 (69%) students 1144 (30.66%) are enrolled in EMS (see figure 5).

The independent variables are split into four broad categories: Student characteristics, Household characteristics, School characteristics and Socio-Religious characteristics. The descriptive statistics of independent variables for Regional and English Medium Schools is given below in Table 1. From the results of the means comparison test between the students who attend RLS and EMS, we can observe that, on an average, EMS students spend more time for homework and they are more likely to attend private tuition. This is consistent with the evidence from the more recent literature on the difference between student study effort in private and public schools. (Muralidharan & Sundararaman, 2013b)

Table 1: Descriptive Statistics for Regional Language Schools Vs English Medium Schools

Factor	EMS = 0	EMS= 1	p-value
N	2587	1144	
Student Characteristics			
Male	1451 (56.1%)	682 (59.6%)	0.0447

Age 9	556 (21.5%)	306 (26.7%)	0.0004
Age 10	856 (33.1%)	322 (28.1%)	0.0028
Age 11	520 (20.1%)	199 (17.4%)	0.0534
Standard (in years)	3.651 (1.665)	3.802 (1.464)	0.0081
School Characteristics			
Mid Day Meals Scheme	340 (13.5%)	37 (3.4%)	<0.0001
Free Uniform	91 (3.6%)	3 (0.3%)	<0.0001
Free Books	408 (16.0%)	30 (2.7%)	<0.0001
School Fee Paid By Government	190 (7.6%)	27 (2.4%)	<0.0001
Household Characteristics			
Household Size	6.897 (3.034)	6.103 (2.772)	<0.0001
Number of Children	3.243 (1.594)	2.555 (1.207)	<0.0001
Household Monthly Consumption (log)	6.545 (0.584)	7.088 (0.576)	<0.0001
Household Income (log)	10.614 (0.998)	11.203 (0.992)	<0.0001
Fathers Education (in years)	1.957 (1.282)	2.830 (1.098)	<0.0001
Mothers Education (in years)	1.178 (1.234)	2.299 (1.300)	<0.0001
Female head	199 (7.7%)	96 (8.4%)	0.47
Either Parent Knows English	776 (33.9%)	698 (67.7%)	<0.0001
Urban Neighborhood	1161 (44.9%)	801 (70.0%)	<0.0001
Socio Religious Characteristics			
Brahmins	215 (8.3%)	127 (11.1%)	0.0065
Other Backward Caste	863 (33.4%)	339 (29.6%)	0.0247
Schedule Caste	380 (14.7%)	125 (10.9%)	0.0020
Schedule Tribe	83 (3.2%)	56 (4.9%)	0.0121
Muslims	472 (18.2%)	147 (12.8%)	<0.0001
Sikh & Jains	79 (3.1%)	52 (4.5%)	0.0225
Christians	44 (1.7%)	35 (3.1%)	0.0079

Note: Pearson's chi-squared test for dichotomous variables & two sample t-test for the continuous variable

The IHDS classified the schools into seven different categories: Education Guarantee Scheme Schools, Government Schools, Government Aided Schools, Private Schools, Convents, Madrassas, Junior College and Other Schools. The student enrollment in each of these schools is shown in Figure 6. The Education Guarantee Scheme (EGS) is a rights-based and community participation driven program. The government guarantees that it will provide a school within 90 days, if a Panchayat (village administration unit) forwards a list of 40 children (25 to 30 in tribal

areas) with no schooling facilities within walking distance (V. Ramachandran, Bose, Mantralaya, & Bhopal, 2004). Since the EGS schools are run by the Panchayat, this study will treat the EGS Schools as “Government Schools” and therefore exclude them from our analysis. Government-aided schools are very similar to privately run schools since teacher recruitment and performance are monitored by school management using locally appropriate standards (Desai et al., 2009). Based on these considerations, Government Aided Schools, Private Schools, Convents, Madrassas, Junior College and Other Schools will be combined into one single category--“Private Schools”—and, as discussed earlier, only students who attend these schools will be a part of this study.

The household survey matched each individual in the household with a unique id. Using the id the first the record of the father and mother was identified, and then their educational attainment was retrieved. A same method was used to retrieve the data to identify if the household has a female head. The IHDS survey also tested the English proficiency of each individual of the household. Using this information a new variable was created which indicates if either one parent was proficient in English. A naïve comparison between EMS and RLS indicates that parents who send their kids to EMS school are more educated and more proficient in English than parents who send their children to RLS (see Table1).

Along with the household survey, IHDS also gathered data from one government and one private primary school from every rural or urban block (wherever possible). The data related to tuition fee and other fee—namely books, uniform, bus etc—was collected from both the schools as well as the household. In addition to the school related fee, households spend money for private tuition. The total money spent on education for a child by a household is the sum of school tuition, private tuition and other miscellaneous fee. Therefore the total education

expenditure depends on the socioeconomic status of the household and other household characteristics such as family size, education level of parents etc. In this study, the data related to the education expenses is obtained directly from the household survey to maintain consistency.

The data reveals that on an average an English Medium student spends around 2875 Rs/month for school fees, 1955 Rs/ month for Books, Uniform, Transportation and other fee and 588 Rs/month for private tuition, whereas a regional medium student spends only 858 Rs/month, 870 Rs/month and 252 Rs/month respectively. These costs are borne by the household entirely and on an average the total money spent per child for EMS is nearly two and a half times more than RMS. This means those who belong to the higher socioeconomic status should have better access to EMS compared to poorer households. The total household income is an indicator of the socioeconomic status of the household. The IDHS household survey queried over 50 different income sources and broadly classified the sources of income into 8 different categories. The total income of a household was computed by calculating the sum of all the sources of income. The total household income data was shown to be skewed towards zero and therefore in order to transform the data into a more normally distributed form, the data was converted into its natural log. A small percentage (0.26%) of the income data had zero as its value. Since log transformation of zero is not possible, a value of 1 was added to all observations before transforming the data to remove any zero values. After log transformation, the skewness statistics reduced from 3.79 to -0.62. Along, with the household income, the household consumption per capita (in rupees) is also included as one of the covariates in the logit model, because the willingness of the household to spend more money for the child's education is also an important factor which determines the choice of school.

India has a wide array of castes and religions. The IHDS survey collapsed religion and caste into one single composite variable. They divided the majority religion, Hinduism, into four commonly used caste categories namely Brahmins, High Caste, Other Backward Castes and Scheduled Castes/Scheduled Tribes. Muslims and Christians were coded separately, and the rest of the religious minorities namely Sikh, Jain, and Buddhists were collapsed into one category. As far as the preference of different communities with regards to English medium education is concerned, we can observe that Christians are more likely to be in English convent schools. This also explains the reason why states of the North East and Kerala, which have a substantial Christian population, also have a high proportion of convent schools when compared to the rest of the country.

6. Data Analysis

Table 2 shows basic regression model with only the main dependent variable. Before controlling for other confounding variables, we can observe that EMS has actually has a positive impact on student learning outcomes.

Table 2: Basic Regression Model

	Dependent variables		
	Reading Score	Maths Score	Writing Score
English Medium	0.288*	0.333*	0.125*
Constant	2.925*	1.841*	0.766*
Number of Observations	3658	3630	3620
R2	0.0128	0.0256	0.021

Note: * indicates significance at 5% or lower

The first step in the analysis is to compute the propensity score. Since the treatment is a dichotomous variable (D=1 for the EMS and D=0 for RMS), I make use of a logit model to

identify the factors determining the student enrollment in EMS. Table 3 shows the results of the logit model. Using the logit model the propensity score for each individual in the study sample was generated. After generating the propensity scores we check whether the “balancing property” and the “common support” assumptions are satisfied,

6.1. Model Validation

As discussed earlier, there is no way to test if the “unconfoundedness” assumption is met. However, we need to verify whether the propensity scores, which we estimated using the logit model, adequately balances the characteristics between individuals in the treatment and comparison group. In other words, after conditioning for the propensity score, the treatment has to be independent of the characteristics of the individual. If the propensity scores do not satisfy this condition, also known as “balancing property”, then the matching quality is likely to suffer, which in turn will produce biased results. Becker & Ichino (2002) provide STATA routines for PSM estimators, which include nearest neighbor, kernel, radius, and stratification matching and Leuven & Sianesi (2003) provide a STATA routine which can be used to graphically verify whether the balancing property is satisfied. This study makes use of these routines. To check if the balancing property is satisfied the STATA routine divides the observations into an optimal number of blocks based on the estimated propensity scores, such that within each block the difference between mean of the estimated propensity score is not statistically significant. The number of blocks can either be specified by the user or STATA automatically identifies the number of blocks based on the propensity scores. This study used the later and the final number of blocks was 7 for both logit models. Following that, the routine conducts t-tests within each block to test if the distribution of covariates is the same between both groups. Figure 8 and 9 show difference in means of each covariate in the treatment and control group before and after

matching. We can clearly see that after matching the difference in the mean of each variable is almost zero, thus indicating that the balancing property is satisfied.

The “common support” is the second underlying assumption of the PSM estimator. Unlike the “unconfoundedness” assumption, the common support assumption can be easily tested. The simplest method is to graph the distribution of the propensity score in both groups and then visually analyze if there is significant overlap between the two groups. Bryson, Dorsett, & Purdon (2002) argue that for the estimation of Average treatment effect on the treated it should suffice if each treated participant has a close neighbor who is untreated. Figure 10 and 11 shows the propensity score distribution between the treated and non-treated groups and we can observe that there significant overlap between the propensity score density functions of both groups. The distribution of the propensity score can give an initial reading of the extent of overlap between both the treatment and control groups and can assist the researcher in selecting the appropriate matching method.. It is important to note that not all units in the treatment and control group will be used for matching. Figure 12 shows the propensity score distribution of the matched samples using NN method. Both logit models used in this study satisfies the common support assumption.

6.2. Matching procedures

When it comes to choosing the matching procedures it is important to consider the “bias vs efficiency” trade-off. Each matching algorithm has its advantages and disadvantages. Therefore is certainly no one right method. In this study, in order to make sure that the findings are not driven by the selection of a particular matching strategy, the PSM coefficients are estimated using three different matching algorithms—namely Nearest Neighbor, Normal Kernel, Stratified Matching. Furthermore, the standard errors are also estimated using bootstrap methods. If the

estimators do not vary significantly based on the matching procedure, the results are considered to be robust.

The final results from the data analysis are tabulated in Table 4.

7. Discussion

In this section I shall discuss the results of the logit model and then delve into propensity score estimator results.

7.1. Characteristics of English Medium School Students

As discussed earlier the determinants of English Medium Education are broadly divided into four categories. The Student Characteristics includes variables corresponding to the age, gender and the grade. The School Characteristics includes the various government based incentives such as mid-day meals, free books, free uniform and scholarships offered through the school. The Household Characteristics include the socioeconomic status, composition of the household, educational attainment of the parents and the neighborhood of the household. And finally, the socio religious group includes dummy variables indicating the community to which the student belongs. It is important to note that we use a stratified sample and therefore these results cannot be extrapolated to the entire population.

Among the variables under the “student characteristics” category we can observe from Table 3 that being male is positively correlated with the EMS enrollment. This indicates that, *ceteris paribus*, male child will be more likely to go to EMS. This result is not surprising since many previous studies have highlighted that in developing countries, there is a tendency among parents to discriminate against female children by reserving “privileged education” only for male

children. And as Geeta Gandhi and Kingdon (2005) point out studies based on the “individual” as the unit of analysis, reveal gender discrimination.

Table 3: Characteristics of English Medium School Students Logit Model

Variables	Model I	Model II
Student Characteristics		
Male	0.250*	0.222*
Age 9	0.062	0.062
Age 10	-0.263	-0.254
Age 11	-0.308	-0.305
Standard (in years)	-0.020	-0.030
School Characteristics		
Mid Day Meals Scheme	-0.234	-0.244
Free Uniform	-0.718	-0.702
Free Books	-1.663*	-1.665*
School Fee Paid By Government	-0.249	-0.244
Household Characteristics		
Household Size	-0.032	
Number of Children		-0.155*
Household Monthly Consumption (log)	0.891*	0.837*
Household Income (log)	0.184*	0.202*
Fathers Education (in years)	0.031*	0.030*
Mothers Education (in years)	0.078*	0.074*
Female head	0.220	0.201
Either Parent Knows English	0.382*	0.395*
Urban Neighborhood	0.717*	0.698*
Socio Religious Characteristics		
Brahmins	-0.405*	-0.378*
High Caste	-0.457*	-0.460*
Schedule Caste	-0.034	-0.012
Schedule Tribe	1.165*	1.160*
Muslims	-0.106	-0.048
Sikh & Jains	-0.040	-0.054
Christians	0.175	0.136
Constant	-9.789*	-9.290*
Number of Observations	3021	3021

Pseudo R2	0.2398	0.2432
------------------	---------------	---------------

Note: * indicates significance at 5% or lower

Coming to the school characteristics, it is interesting to note that government subsidies like Free Books and the Mid-Day Meal program do not have any significant effect. This result is contrary to the general consensus in the literature, which points out that the Mid Day Meals scheme has a positive impact on schooling choice of parents. However, the results should be interpreted with caution, as this study includes only students who go to private schools. Previous studies on this subject establish a causal link between socioeconomic status and the parent's decision to send the child to a private school. Since these programs primarily cater to those sections of the society who avail public schools rather than private schools, we can only conclude that the Mid Day Meals programs has no impact on among parents who send their children to private schools. Among the government incentive programs, we can observe that the free books seem to have a significant negative impact on the parent's decision to send the child to EMS.

From Table 3 we can observe that, as expected, those students who belong to a household with a higher income and higher per capita consumption will be more likely to be enrolled in EMS. As we have discussed in the Data section, we estimated that on an average the money spent on the education for students who go to EMS was approximately two and half times more than the RMS. Therefore it does not come as a surprise that, even after controlling for all other factors, rich parents are more likely to send their children to EMS. We can also observe that parental education has an impact on educational choice. It is interesting to note that maternal education has a significant positive impact on EMS enrollment. This result is consistent with the general view that more educated women tend to be more assertive in the educational choices of

the children. In addition to that, our results show that if “either parent knows English”, the child is more likely to attend an EMS. This suggests that there is evidence of intergenerational effects on educational choice. Parents who have good English skills are more likely to send their child to an EMS.

Proximity of EMS to one’s home is bound to affect parental choices. Proximity depends not only on demand factors but also on State, Central and local government policies. In the Indian context, studies show that urban neighborhoods are more likely to have more English medium schools. The results show that, *ceteris paribus*, if a household is in an urban neighborhood, the child will be more likely to attend an EMS. From the various household characteristics, one pattern is clearly visible. Urban, rich and educated parents are more likely to send their children to EMS than others. Although there is no conclusive evidence, this is consistent with the literature that EMS engages in ‘cream-skimming’. If that is indeed the case, higher academic achievement by students who attend EMS compared to RMS could be explained by “cream skimming”, a subject that requires further scrutiny.

Other household variable that have a significant is the number of children in the household. This indicates that, *ceteris paribus*, an additional child in the household decreases the likelihood that a student will be enrolled in an EMS. This makes intuitive sense because an extra child would put additional burden on the finances of the household, making EMS education less affordable for the parents.

In India, caste and religion play an important role in people’s cultural lives and influences everyday household decisions. In our model, Other Backward Castes (OBC’s) are treated as the baseline category. Within the Hindu caste hierarchy the OBC’s are lower than the Brahmins and

High Caste Hindus. From the results, we can observe that compared to the OBC's, the Brahmins and the High Caste Hindu parents are less likely to send their children to EMS. In addition to that, we can also observe the Schedule Tribal children are more likely to attend EMS compared to OBC's. These results are interesting because communities which are lower down the caste hierarchy perceive English as a vehicle for social mobility and “empowerment” (Proctor, 2010). It is also significant to note that backward communities are relatively poorer and studies have shown that they have less access to private schooling than the Christian, Sikh and Jain communities (Desai et al., 2009).

7.2. Propensity Score Model Results Discussion

Table 4: Net Impact of EMS on Students computed using different Matching Techniques

Outcome	Nearest Neighbor		Stratification		Normal Kernel	
	Model I	Model II	Model I	Model II	Model I	Model II
Reading Score	-0.103	-0.008	0.003	0.004	-0.006	-0.006
	(0.09)	(0.08)	(0.051)	(0.079)	(0.023)	(0.042)
	[-1.154]	[-0.098]	[0.052]	[0.049]	[-0.251]	[-0.131]
Mathematics Score	0.019	0.046	0.056	0.047	0.052	0.053*
	(0.053)	(0.077)	(0.071)	(0.034)	(0.029)	(0.018)
	[0.355]	[0.595]	[0.798]	[1.386]	[1.796]	[3.009]
Writing Score	0.014	0.035	0.037	0.037*	0.036	0.036
	(0.019)	(0.024)	(0.021)	(0.013)	(0.021)	(0.026)
	[0.727]	[1.473]	[1.778]	[2.769]	[1.689]	[1.400]

Note:

Figures in parenthesis indicate bootstrapped associated clustered standard errors

Figures in square brackets indicate associated t-statistics

* indicates significance at 5% or lower

Since our dependent variables are ordinal variables it is not possible to interpret the magnitude of the effect of EMS. We can only say whether the program has positive, negative or no effect on the outcome. Table 4 shows the results of the three outcomes of interest. With the exception of the writing score using Stratification Method and Mathematics Score using Kernel Method for Model II, we observe that the results are not statistically significant. Even in the case of the writing score and mathematics score the NN matching method does not indicate that EMS schools have any positive impact on the students. This indicates that for the results derived by using Model II are not robust, since it depends on the selection of the matching algorithm. It is interesting to note that for all the three outcome variables, before matching the results, showed significant positive impact. After matching (at least for Model I) the Average Treatment Effect on all three outcomes becomes insignificant. The results from our study suggest that there is no evidence that EMS has an effect on the overall learning outcomes of students.

Table 5: Descriptive Statistics of differences in student inputs between EMS and RMS

Factor	englishmedium = 0	englishmedium = 1	p-value
N	2587	1144	
School Hours (hrs/week)	31.472 (8.204)	31.321 (9.054)	0.62
Homework Hours (hrs/week)	7.553 (5.532)	10.067 (6.080)	<0.0001
Private Tuition Hours (hrs/week)	1.683 (4.221)	3.101 (5.331)	<0.0001
School fees (Rs/month)	858.714 (1099.989)	2875.715 (2819.053)	<0.0001
Books uniform bus etc (Rs/month)	870.148 (906.861)	1955.474 (2078.386)	<0.0001
Private tuition fee (Rs/month)	252.793 (1233.158)	588.137 (1840.850)	<0.0001
Total Education Expenditure (Rs/month)	2020.802 (2355.326)	5418.402 (4775.829)	<0.0001

Note: The p-value is the result from two sample t-tests.

Muralidharan & Sundararaman, (2013b) studied the effect of school choice program and one of the findings of the study was that, on an average, the English medium schools have superior indicators of school quality including “facilities; teacher experience, qualifications, and salary; and annual fees charged per child”. Table 5 shows that besides the differences in household inputs, there are significant differences in student inputs as well. Table 6 shows the ATT of EMS on the student’s time allocation for Homework. We can observe that for both Models I & II EMS students have a significantly higher probability to spend more hours per week for Homework than their RMS counterparts. However, this result is only shows that the students enrolled in EMS generally put in more study effort, and cannot be interpreted to mean that higher study time allocation leads to higher test scores.

Table 6: Net Impact of EMS on student’s time allocation for Home Work

Outcome	Nearest Neighbor		Stratification		Normal Kernel	
	Model I	Model II	Model I	Model II	Model I	Model II
Homework Hours(hrs/week)	0.885*	1.123*	1.090*	1.021*	1.020*	1.015*
	(0.258)	(0.182)	(0.261)	(0.190)	(0.372)	(0.284)
	[2.659]	[6.169]	[4.176]	[5.363]	[2.240]	[3.573]

Note:

Figures in parenthesis indicate bootstrapped associated clustered standard errors

Figures in square brackets indicate associated t-statistics

* indicates significance at 5% or lower

8. Limitations

Firstly, the main limitation of this study stems from the pooling many state language schools into one single category called “Regional Medium Schools”. Education falls in the concurrent list in the Indian Constitution, meaning, both the Centre and the State has power to change

education policy. The Sixth All India Education Survey (1993) reveals that there are forty seven languages currently being used as medium of instructions in India. Hence there is a lot of heterogeneity amongst the different regional medium schools. Each state has its own State Education Board which is responsible for running the public schools within the state and for regulating the privately run schools). In addition to that, there is the Central Board of Secondary Education (CBSE), an agency administered by the Union Government, through which the affiliated Kendriya Vidhyalaya and other private schools (mostly English Medium) are run throughout the country. Therefore this simple dichotomous categorization, in effect, subsumes all these regional variations into the overall national picture. However, for the purpose of this study, this categorization was unavoidable since the sample size within each state was too small to study each state separately. Future studies should take into account the state-level variations or should just focus on one state instead of studying the nation as a whole.

Second, the PSM estimator requires the “selection on observables” assumption to be satisfied. In case this assumption is violated, the estimates are likely to be biased. Although there is no way one can test this assumption, I recognize that this study has not adequately captured all the relevant data. For instance, IHDS does not contain the data about the student’s native language. This can have an impact on both the choice of school as well as on the academic performance of the student. As we have discussed earlier, the PSM method is no silver bullet to all problems concerning “selection bias”. Future studies should aim at measuring pre-treatment variables and follow a difference-in-difference approach to arrive at more accurate estimates.

Third, this PSM method assumes that there is no selection bias due to unobservable characteristics. This assumption can be violated if there is evidence of ‘cream skimming’ in many private English medium schools. If that is indeed the case, then there is “program selection

bias” and therefore the Average Treatment Effect on those who attend EMS is likely to be “over-estimated”. It is worth noting, however, that the final results from our study showed that there is no significant difference between both groups.

Fourth, this study makes an implicit assumption that the students, after enrolling in one school, continued their studies in the same school until the point when this survey was administered. In reality, however, this may not be true. Some students may have changed the schools. Since IHDS does not contain data on student history, this study could not include the relevant variables pertaining to the educational history of the student. If the population of students who have shifted schools is substantial (something we do not know for sure), then the net impact of English Medium Schools is likely to be inaccurate.

9. Policy Recommendations

This study finds that there is no significant impact of EMS on student learning outcomes. If one has to take student learning outcomes as the only relevant input to policy makers, then results suggest that the medium of instruction policy does not matter. However, we already know that it the cost of obtaining education through the regional language is lesser than English. Therefore there is no conclusive evidence to accept one of the main recommendations of the National Knowledge Commission Report (2006) which is to expand English medium schools even in the public sector. The medium of instruction is a red herring and the Government should rather focus on improving the quality of existing public schools. As Kochar (2002) points out, households with a lower level of education attainment are more adversely affected by poor school quality than are better-schooled households.

As discussed in the literature, there is a “wage premium” for individuals who have the ability to communicate fluently in English. The Human Capital Theory posits that education improves the learning capability of individuals. If English language is taught at schools, it can improve the English language skills of the population. This can enhance the capability of Indians to participate in the global economy, which in turn will improve overall economic performance of the country. Therefore part of the NKC recommendation which is to take necessary steps to improve English Education even in regional language schools should be implemented. This could prevent the widening social cleavages, legitimization of the social hierarchies and inequalities associated with the contemporary model of development, in which English skills plays a crucial role. The 7th All India Education Survey (2003) reveals that 55.05% and 54.1% schools teach English as the second language in the upper primary and secondary stage respectively. The Central Government should collaborate with the States to make sure English is taught in all public schools across the country.

Some scholars have argued that privileging English over other native languages can pose a serious threat to the minority languages (Rao, 2008). In the Indian context, like many European states, language ideology plays an important role in determining the language policy of the government. Therefore the state should approach this subject with great sensitivity. Privileging English over other State languages could create social tension. According to Faust & Nagar (2001)

While education in English has been advocated as a unifying and modernizing force, it is also seen as a marker of imperialism and class privilege and a terrain of struggle among elite groups. Ruptures in such a class-divided educational system in turn shape specific debates over development, democracy and social change. Uneven empowerment that an education in English generates also has its fallout in an

increasing polarization, fracturing and violence against caste, gender and religious lines.

Future research should explore the differences of different types of schools even further and should pay more attention to the school heterogeneity. Studies should also focus in estimating the impact of English education in RMS on English language skills. Besides quantitative approach there is also scope to use qualitative methodology to gain more insight on this subject.

References

- Abadzi, H. (2006). *Efficient learning for the poor: Insights from the frontier of cognitive neuroscience*. World Bank Publications. Retrieved from http://books.google.com/books?hl=en&lr=&id=kn62phyvFpwC&oi=fnd&pg=PR3&dq=+Efficient+Learning+for+the+Poor:+Insights+from+the+Frontier+of+Cognitive+Neuroscience&ots=7sEg28Erbr&sig=EE8eyvwGA_fEofBMeHlaJdp4jVc
- Afridi, F. (2011). The impact of school meals on school participation: evidence from rural India. *Journal of Development Studies*, 47(11), 1636–1656.
- Alderman, H., & King, E. M. (1998). Gender differences in parental investment in education. *Structural Change and Economic Dynamics*, 9(4), 453–468.
- Azam, M., Chin, A., & Prakash, N. (2013). The returns to English-language skills in India. *Economic Development and Cultural Change*, 61(2), 335–367.
- Azam, M., & Kingdon, G. G. (2013). Are girls the fairer sex in India? Revisiting intra-household allocation of education expenditure. *World Development*, 42, 143–164.
- Bamgbose, A. (2005). Mother-tongue education: Lessons from the Yoruba experience. *Languages of Instruction for African Emancipation: Focus on Postcolonial Contexts and Considerations*, 210–234.
- Becker, S. O., & Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *The Stata Journal*, 2(4), 358–377.
- Benson, C. J. (2002). Real and potential benefits of bilingual programmes in developing countries. *International Journal of Bilingual Education and Bilingualism*, 5(6), 303–317.

- Borooah, V. K., & Iyer, S. (2005). Vidya, Veda, and Varna: The influence of religion and caste on education in rural India. *The Journal of Development Studies*, 41(8), 1369–1404.
- Bryson, A., Dorsett, R., & Purdon, S. (2002). The use of propensity score matching in the evaluation of active labour market policies. Retrieved from http://eprints.lse.ac.uk/4993/1/The_use_of_propensity_score_matching_in_the_evaluation_of_active_labour_market_policies.pdf
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31–72.
- Chang, J. (2006). Globalization and English in Chinese higher education. *World Englishes*, 25(3-4), 513–525.
- Cheney, G. R., Ruzzi, B. B., & Muralidharan, K. (2005). A profile of the Indian education system. *Prepared for the New Commission on the Skills of the American Workforce*. Retrieved from http://www.teindia.nic.in/files/articles/indian_education_sysytem_by_karthik_murlidharan.pdf
- Cummins, J. (1978). The Cognitive Development of Children in Immersion Programs. *Canadian Modern Language Review*, 34(5), 855–83.
- Cummins, J. (1979). Cognitive/Academic Language Proficiency, Linguistic Interdependence, the Optimum Age Question and Some Other Matters. Working Papers on Bilingualism, No. 19. Retrieved from <http://eric.ed.gov/?id=ED184334>
- Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics*, 84(1), 151–161.
- Desai, S., Dubey, A., Vanneman, R., & Banerji, R. (2009). Private schooling in India: A new educational landscape. In *India Policy Forum* (Vol. 5, pp. 1–58).
- Desai, S., Vanneman, R., & National Council of Applied Economic Research, N. D. (2013). India Human Development Survey (IHDS), 2005. Inter-university Consortium for Political and Social Research (ICPSR) [distributor]. Retrieved from <http://doi.org/10.3886/ICPSR22626.v8>
- Dixon, P., & Tooley, J. (2005). The Regulation of Private schools serving low-income families in Andhra Pradesh, India. *The Review of Austrian Economics*, 18(1), 29–54.
- Drèze, J., & Kingdon, G. G. (2001). School participation in rural India. *Review of Development Economics*, 5(1), 1–24.
- Dreze, J., & Kingdon, G. G. (2001). School participation in rural India. *Review of Development Economics*, 5(1), 1–24.

- Drèze, J., & Murthi, M. (2001). Fertility, education, and development: evidence from India. *Population and Development Review*, 27(1), 33–63.
- Dreze, J., & Sen, A. (2003). Basic education as a political issue. *Education, Society and Development: National and International Perspectives*, 3.
- Duraisamy, P., & others. (1992). *Gender, intrafamily allocation of resources and child schooling in South India*. Economic Growth Center, Yale University. Retrieved from http://www.econ.yale.edu/growth_pdf/cdp667.pdf
- Faust, D., & Nagar, R. (2001). Politics of development in postcolonial India: English-medium education and social fracturing. *Economic and Political Weekly*, 2878–2883.
- Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data—or tears: An application to educational enrollments in states of india*. *Demography*, 38(1), 115–132.
- Foray, D., & Lundvall, B. (1998). The knowledge-based economy: from the economics of knowledge to the learning economy. *The Economic Impact of Knowledge*, 115–121.
- Gandhi Kingdon, G. (2002). The gender gap in educational attainment in India: How much can be explained? *Journal of Development Studies*, 39(2), 25–53.
- Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence a test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, 14(1), 31–38.
- Härmä, J. (2011). Low cost private schooling in India: Is it pro poor and equitable? *International Journal of Educational Development*, 31(4), 350–356.
- Her, J.-H. (2007). *The globalization of English: Its impact on English language education in the tertiary education sector in Taiwan*. The University of Waikato. Retrieved from <http://researchcommons.waikato.ac.nz/handle/10289/2809>
- Hornberger, N., & Vaish, V. (2009). Multilingual language policy and school linguistic practice: globalization and English-language teaching in India, Singapore and South Africa. *Compare*, 39(3), 305–320.
- India, N. K. C. (2007). *National Knowledge Commission: Report to the Nation 2006*. Government of India.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60–99.

- Kambhampati, U. S., & Pal, S. (2001). Role of parental literacy in explaining gender difference: Evidence from child schooling in India. *The European Journal of Development Research*, 13(2), 97–119.
- Kapur, S., & Chakraborty, T. (2008). English Language Premium: Evidence from a Policy Experiment in India. *Washington University in St. Louis Unpublished Paper*. Retrieved from http://www.isid.ac.in/~pu/conference/dec_08_conf/Papers/ShilpiKapur.pdf
- Kingdon, G. G. (1996). Private schooling in India: Size, nature, and equity-effects. *Economic and Political Weekly*, 3306–3314.
- Kingdon, G. G. (2005). Where has all the bias gone? Detecting gender bias in the intrahousehold allocation of educational expenditure. *Economic Development and Cultural Change*, 53(2), 409–451.
- Kingdon, G. G. (2007). The progress of school education in India. *Oxford Review of Economic Policy*, 23(2), 168–195.
- Kirkgöz, Y. (2009). Globalization and English language policy in Turkey. *Educational Policy*, 23(5), 663–684.
- Kochar, A. (2002). Emerging Challenges for Indian Education Policy. *Economic Policy Reforms and the Indian Economy*, 303–28.
- Kochar, A. (2004). Urban influences on rural schooling in India. *Journal of Development Economics*, 74(1), 113–136.
- Kosonen, K. (2005). *Education in local languages: Policy and practice in Southeast Asia. First languages first: Community-based literacy programmes for minority language contexts in Asia*. Bangkok: UNESCO Bangkok.
- Kremer, M., Chaudhury, N., Rogers, F. H., Muralidharan, K., & Hammer, J. (2005). Teacher absence in India: A snapshot. *Journal of the European Economic Association*, 3(2-3), 658–667.
- Laxmaiah, A., Rameshwar Sarma, K. V., Rao, D. H., Reddy, G., Ravindranath, M., Rao, M. V., & Vijayaraghavan, K. (1999). Impact of mid day meal program on educational and nutritional status of school children in Karnataka. *Indian Pediatrics*, 36(12), 1221–1228.
- Leuven, E., & Sianesi, B. (2003). Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. *Statistical Software Components S*, 432001.
- Munshi, K. D., & Rosenzweig, M. R. (2003). Traditional institutions meet the modern world: Caste, gender and schooling choice in a globalizing economy. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=424620

- Muralidharan, K. (2006). Public-private partnerships for quality education in India. *Unpublished Draft*. Retrieved from <http://econweb.ucsd.edu/~kamurali/papers/Other%20Writing/public%20private%20partnerships%20for%20quality%20education%20in%20india.pdf>
- Muralidharan, K., & Kremer, M. (2006). Public and private schools in rural India. *Harvard University, Department of Economics, Cambridge, MA*.
- Muralidharan, K., & Sundararaman, V. (2009). *Teacher performance pay: Experimental evidence from India*. National Bureau of Economic Research.
- Muralidharan, K., & Sundararaman, V. (2013a). *Contract teachers: Experimental evidence from India*. National Bureau of Economic Research.
- Muralidharan, K., & Sundararaman, V. (2013b). *The Aggregate Effect of School Choice: Evidence from a Two-stage Experiment in India* (Working Paper No. 19441). National Bureau of Economic Research.
- P, G., News, ey B., village, B., & Pradesh, U. (n.d.). An “English goddess” for India’s down-trodden. Retrieved March 11, 2015, from <http://www.bbc.co.uk/news/world-south-asia-12355740>
- PROBE Team, T., Team, P., & others. (1999). *Public report on basic education in India*. Oxford University Press New Delhi. Retrieved from <http://repository.tufs.ac.jp/handle/10108/49943>
- Proctor, L. M. (2010). Discourses on language, class, gender, education, and social mobility in three schools in New Delhi, India.
- Ramachandran, R. (2012). Language use in education and primary schooling attainment: evidence from a natural experiment in Ethiopia. *Documents de Treball IEB*, (34), 1–49.
- Ramachandran, V., Bose, M. S., Mantralaya, V. B., & Bhopal, M. P. (2004). Education guarantee scheme: Community-government partnership that gets millions into school in Madhya Pradesh. In *Conference on Scaling Up Poverty Reduction: A Global Learning Process, Shanghai, May* (pp. 25–27).
- Rao, S. S. (2008). India’s language debates and Education of linguistic minorities. *Economic and Political Weekly*, 63–69.
- Ray, R. (2000). Poverty, household size and child welfare in India. *Economic and Political Weekly*, 3511–3520.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.

- Singh, A., Park, A., & Dercon, S. (2014). School meals as a safety net: an evaluation of the Midday Meal scheme in India. *Economic Development and Cultural Change*, 62(2), 275–306.
- Smith, J. A., & Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*, 125(1), 305–353.
- Tabarrok, A. (2013). Private education in India: A novel test of cream skimming. *Contemporary Economic Policy*, 31(1), 1–12.
- Thomas, W. P., & Collier, V. (1997). School Effectiveness for Language Minority Students. NCBE Resource Collection Series, No. 9. Retrieved from <http://eric.ed.gov/?id=ED436087>
- Tilak, J. B. (2001). *Private schooling in rural India*. Retrieved from <http://ideas.repec.org/p/nca/ncaerw/76.html>
- Viswanathan, G. (2014). *Masks of conquest: Literary study and British rule in India*. Columbia University Press. Retrieved from <http://books.google.com/books?hl=en&lr=&id=AIExBgAAQBAJ&oi=fnd&pg=PR9&dq=Masks+of+Conquest:+Literary+Study+and+British+Rule+in+India&ots=gQjYSqiArm&sig=L317u6ixioFVRI6VyJegJKP9Nh0>
- Zimmermann, L. (2012). Reconsidering gender bias in intrahousehold allocation in India. *Journal of Development Studies*, 48(1), 151–163.

Figure 1: Student Reading Skills

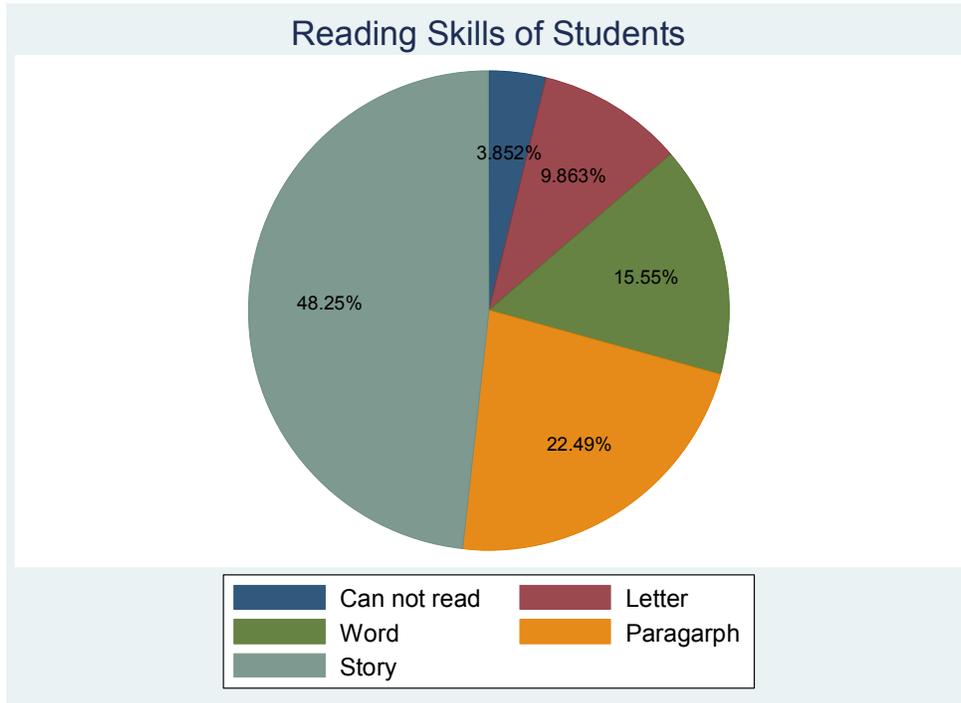


Figure 2: Student Mathematics Skills

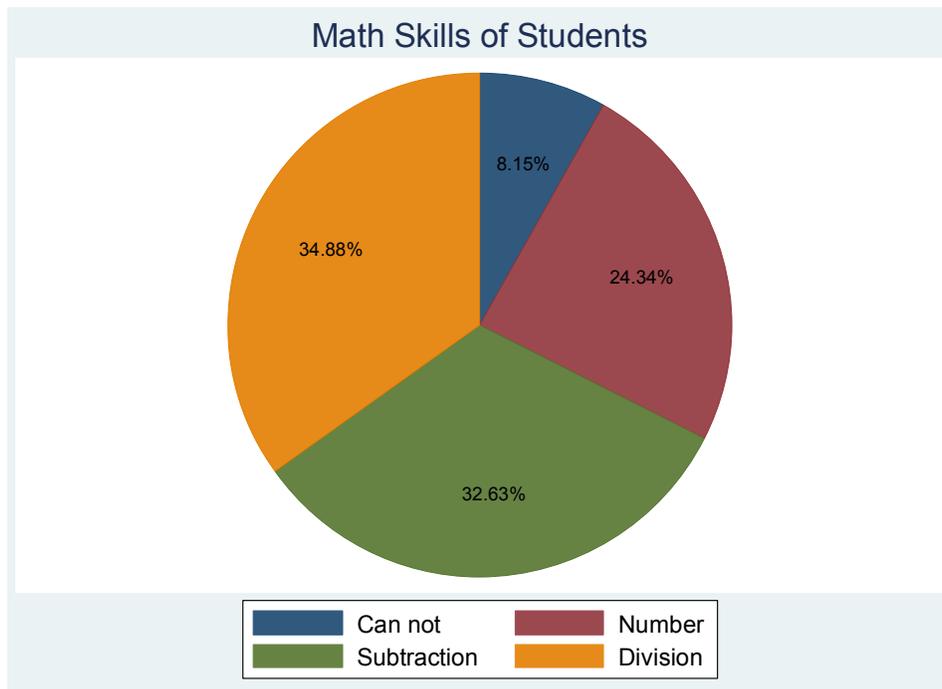


Figure 3: Student Writing Skills

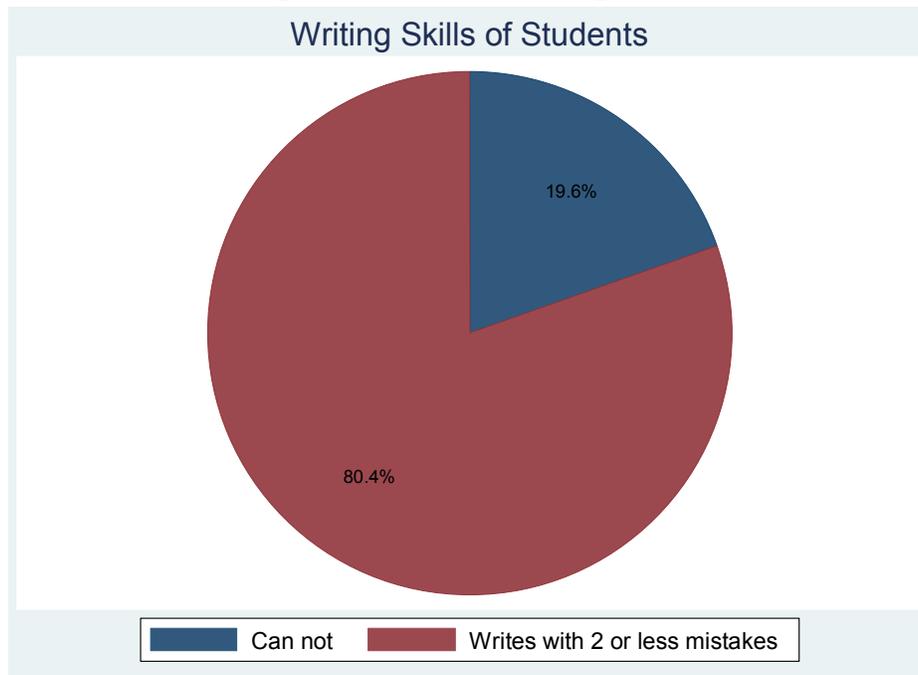


Figure 4: Student Enrollment Status

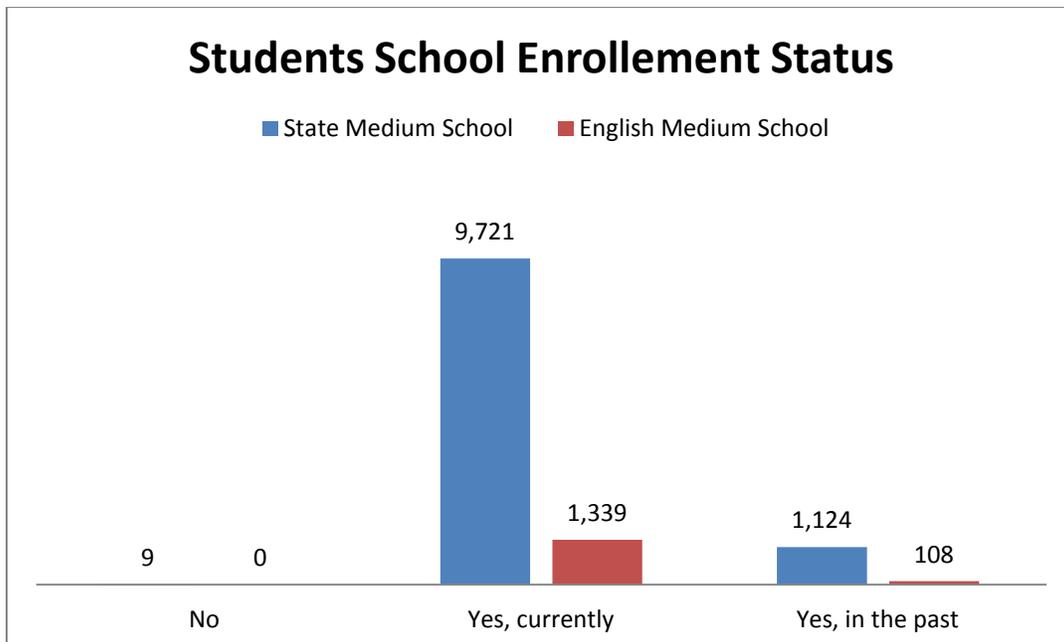


Figure 5: Comparison of Percentage of Students Enrolled in English Medium Schools between Public and Private Schools

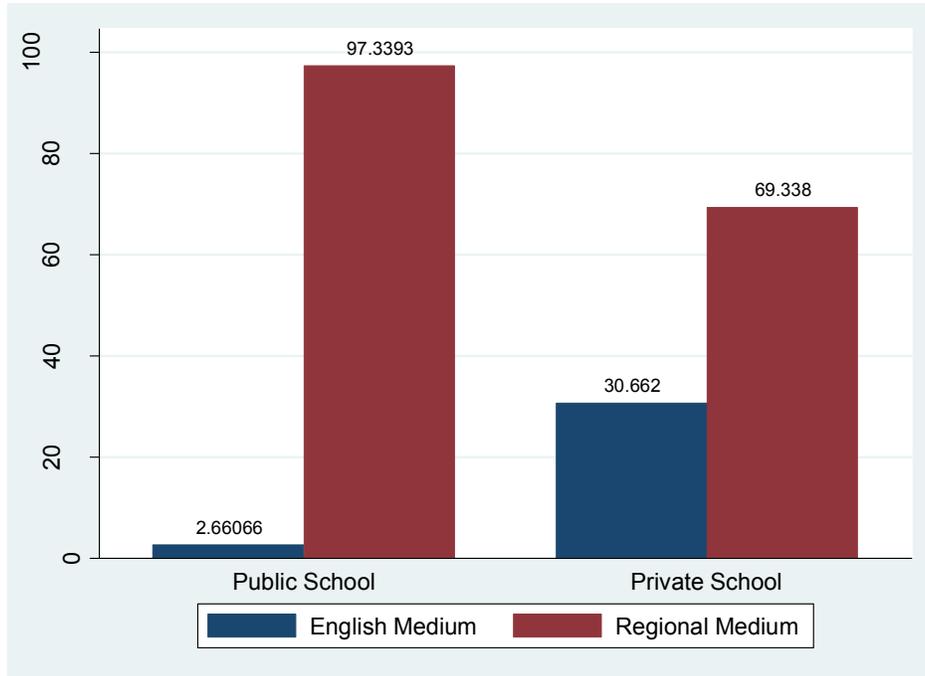


Figure 6: Student Enrollment in Different Schools

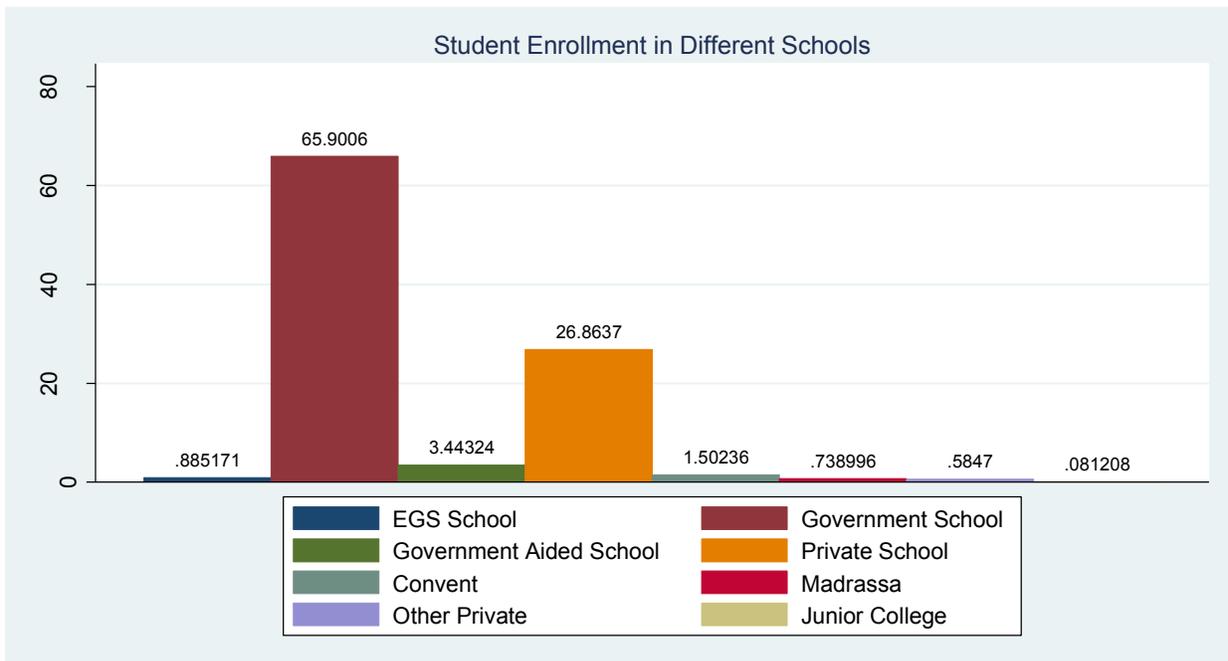


Figure 7: Comparison of Student Enrollment in Private and Public Schools

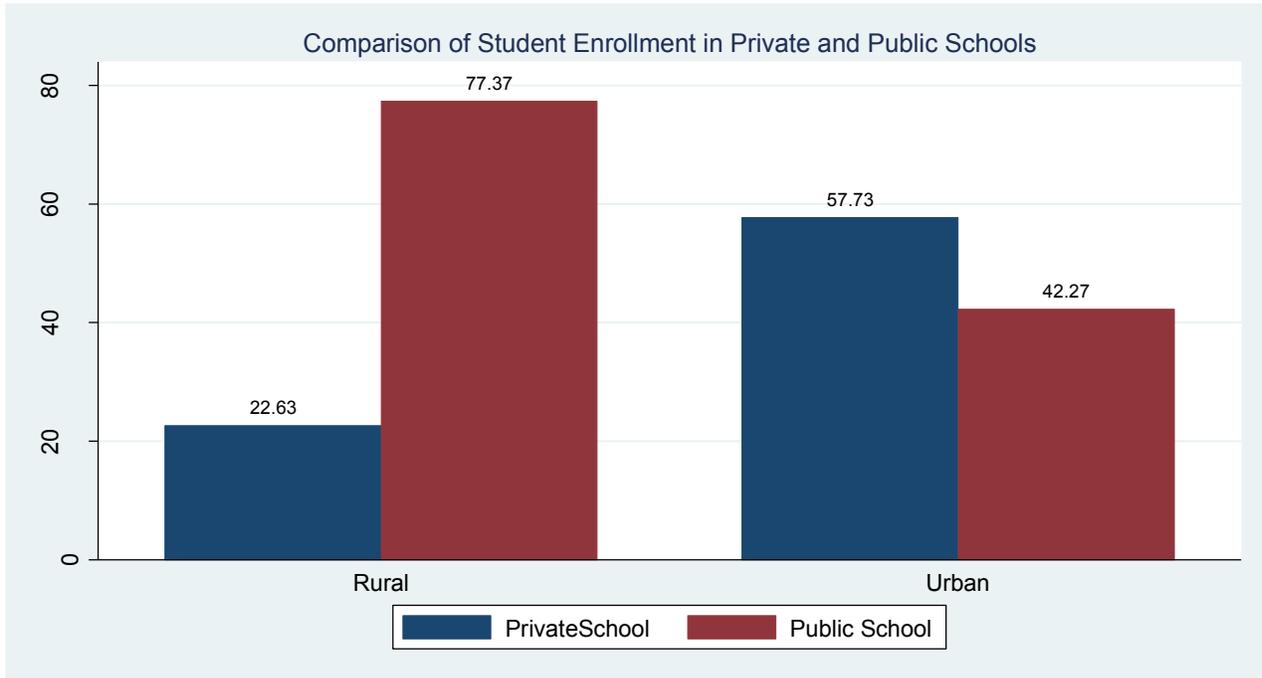


Figure 8: Balancing Property Test for Model 1

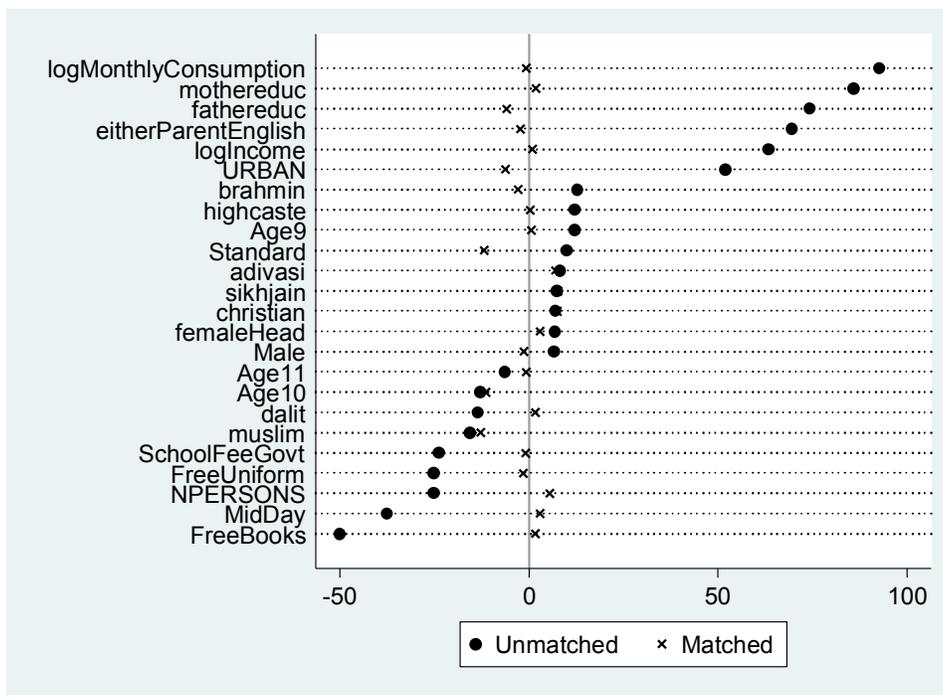


Figure 9: Balancing Property Test for Model II

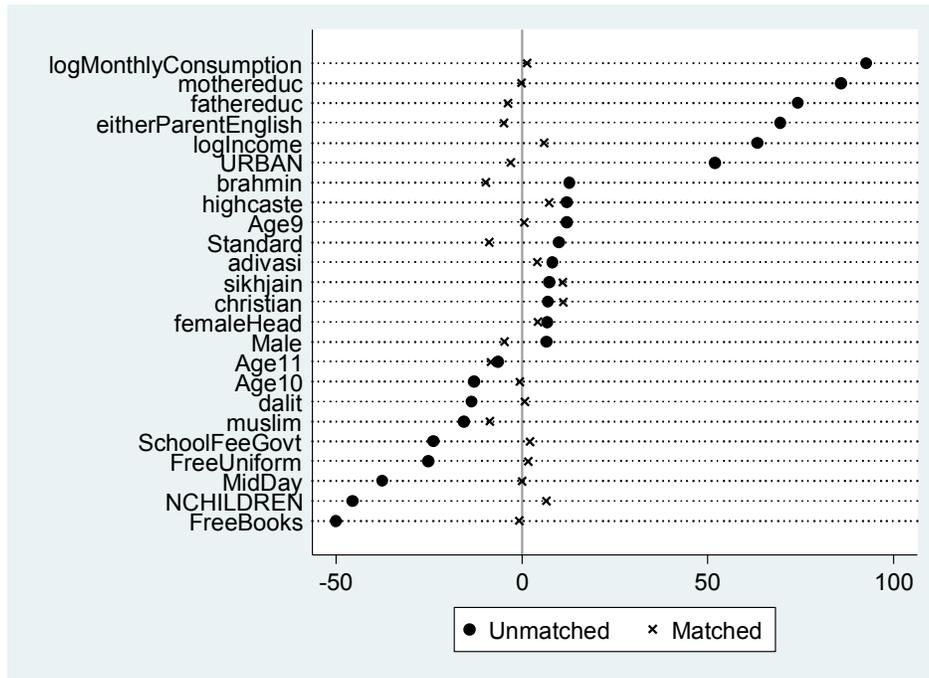


Figure 10: Overlap Test for Model 1

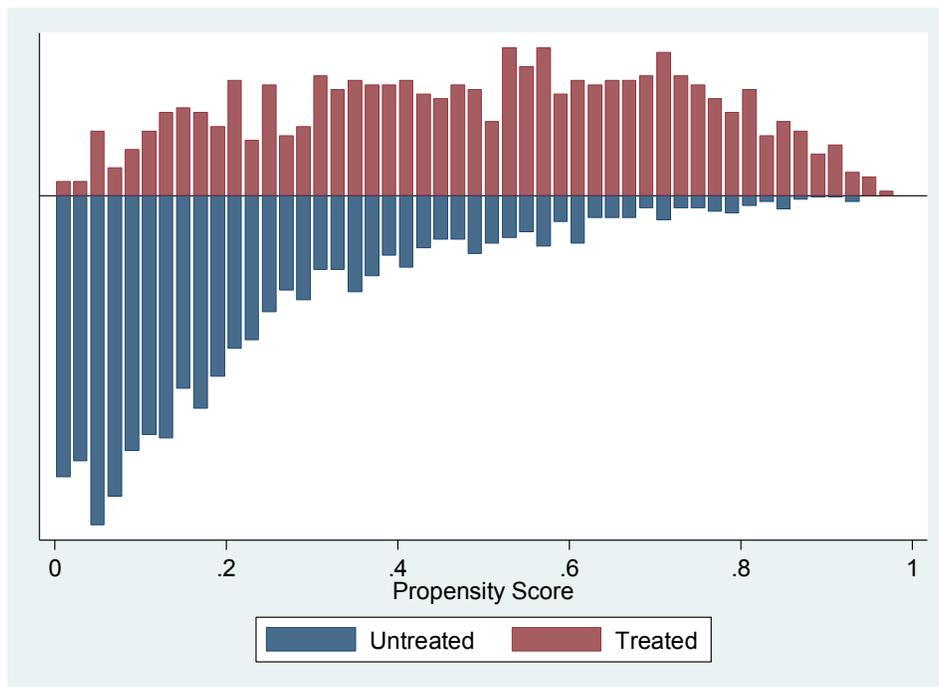


Figure 11: Overlap Test for Model II

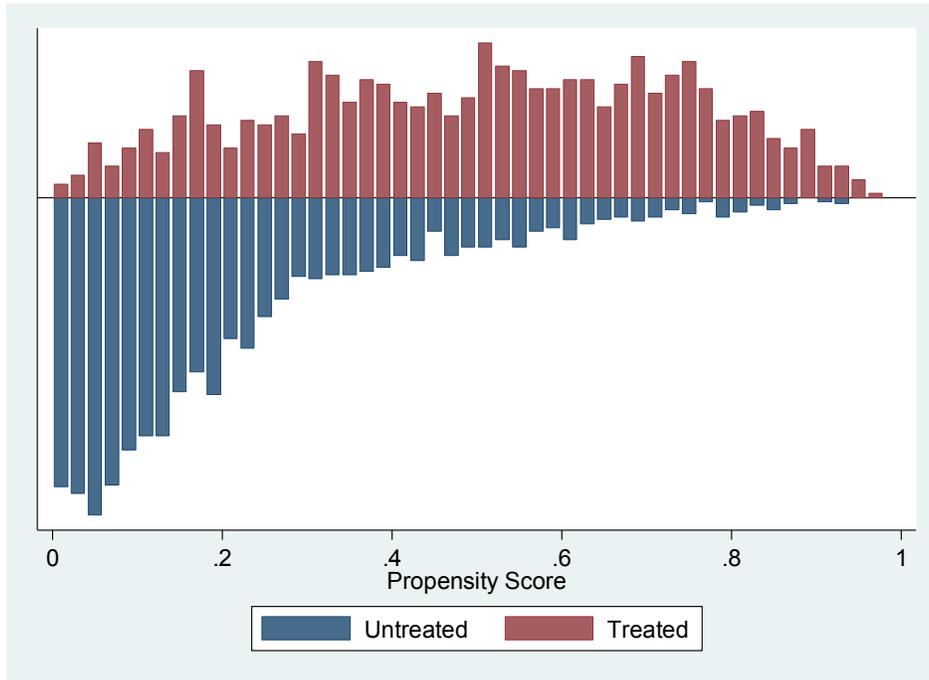


Figure 12: Distribution of propensity score of Matched Samples

