English Language Premium: Evidence From A Policy Experiment In India[+]

Tanika Chakraborty[*]

Shilpi Kapur Bakshi[**]

July 7, 2012

Abstract

A key question facing policymakers in many emerging economies is whether to promote the local language, as opposed to English, in elementary schools. In this paper, we estimate the English premium in a globalizing economy, by exploiting an exogenous language policy intervention in India. Our results indicate that a 10% increase in the probability of learning English in primary school raises weekly wages by 9%. On the average, this implies 29% higher wages for cohorts not exposed to the English abolition policy. We provide further evidence that occupational choice played a decisive role in determining the wage gap.

JEL Classifications: H4, I2, J0, O1

Keywords: English premium, triple difference, education policy, wage, occupation

[+] We thank Sukkoo Kim, Sebastian Galiani, Charles Moul, Bruce Petersen, and Robert Pollak for their invaluable advice and support, Barry Chiswick for his helpful comments and seminar participants at the 2008 Canadian Economic Conference and NEUDC conference for the discussions. We also thank Daifeng He and Michael Plotzke for their feedback. We are grateful to the Bradley Foundation for providing research support and Center for Research in Economics and Strategy (CRES), in the Olin Business School, Washington University in St. Louis, for travel grants. All errors are ours.

[*]Indian Institute of Technology Kanpur and IZA Bonn, email: tanika@iitk.ac.in

[**]TERI India and Washington University in St Louis, email: shilpikapur@gmail.com

1. Introduction

There is a longstanding interest in estimating the economic returns to the human capital embodied in language skills. The previous literature emphasizes the importance of language skills in the context of the economic assimilation of immigrants. Largely ignored however, is the importance of foreign language skills within domestic labor markets of many economies.[1] Ever since their independence, many of the former European colonies faced the dilemma of which language to encourage in educational institutions - local or colonial?[2] Often policymakers opposing foreign language training in schools argue that teaching only the native language fosters easier access to education, particularly for children from disadvantaged backgrounds, thus promoting greater equality over time.[3] Nevertheless, key changes in the economies of many developing countries have led policy makers to rethink the importance of teaching foreign language, particularly English, in schools. The argument against promoting only native language in schools is that if English is more valued in the labor market, then such a policy would make English an elite language available only at a premium. This in turn would imply an ever widening gap between the rich and the poor thus defeating the very purpose of the policy promoting native language. The debate has found renewed attention in many emerging economies like India which benefited from their pre-existing English language proficiency in an increasingly globalized world.[4] In this paper, we investigate the extent to which English language skills are rewarded, if at all, in a global labor market, in turn leaving behind those with otherwise comparable levels of education and experience but lacking English skills.

One of the major difficulties in estimating the returns to language skills, as with any other form of human capital arises because language skills are likely to be correlated with unobserved individual specific ability or family background variables that also affect labor market outcomes. We exploit a language policy intervention in India that generates plausibly exogenous variability in English skills. Until 1983, English was taught in all primary schools in the state of West Bengal, starting from first grade. Beginning in 1983, English was revoked from

[1] Few exceptions are Angrist et al (1997, 2006) and Lang and Siniver (2006), Azam et al (2010)
[2] For example, French was encouraged in the case of many African colonies and English was promoted in the case of many British colonies in Asia.
[3] Post independence, many former European colonies implemented programs to actively promote the national language at the expense of the colonial language in schools (Angrist and Lavy, 1997).
[4] For instance, Shastry (2011) finds that regions with lower costs of acquiring English skills attracted more information technology jobs post liberalization.

primary grades in all public schools in West Bengal and introduced as a part of the curriculum starting from grade 6.[5] However, cohorts who were already enrolled in school before 1983 were exempted from the policy change and continued to learn English in primary classes. Moreover, private schools were out of the purview of this policy.[6] Since individual schooling choice is endogenous, we construct district level probability measures of an individual's exposure to public school. We combine district and cohort variations generated by this exogenous language policy intervention in a two-way fixed effect model to estimate the English skill premium in India.

However, an inherent problem with this two-way-fixed-effects strategy is the possibility of confounding district trends. Districts which provided fewer English learning opportunities in schools might have experienced a greater growth of alternative English training centers in the post policy period. This will downward bias the two-way estimates. To correct for these confounding district trends we estimate a model similar in spirit to a triple difference strategy. Using other states that did not experience any change in language policy during that period as controls we are able to eliminate all factors that varied between districts for each cohort. However, West Bengal might itself have had a different economic growth compared to our control states. We include state-time interactions to account for any difference in trends between the treatment and control states. We conduct further robustness checks to confirm that our results are not driven by underlying trend differentials between the control and treatment districts.

Our estimates suggest that a 10% increase in the probability of learning English in primary school leads to a 8% increase in wages. On the average, this implies a 25% reduction in wages due to the abolition of English from public primary schools. Close examination of how the difference in wage arises, reveals that occupational choice played a decisive role in determining the wage gap. Using a multinomial logit estimation framework, we find that a lower probability of learning English significantly reduces the odds of an individual working in higher ranked or better paying occupations.[7]

[5] Few other states like Karnataka and Tamil Nadu also had similar language policy changes but in much later periods.

[6] According to the "Critical Period" hypothesis of the biological literature, there is a critical age range in which individuals learn languages more easily. If a second language is learned before age 12, the child speaks without an accent. Moreover, syntax and grammar are difficult to learn later in life (Heckman, 2007).

[7] In a later section, we define an ordinal ranking on the broad occupational categorization used in the analysis.

Angrist and Lavy (1997) use a similar policy to estimate French skill premium following the abolition of French from Moroccan primary schools. They find a positive premium associated with French writing abilities. However, since the Moroccan language policy change was a country-wide phenomenon, they could only use variations in individuals' years of schooling and cohort of birth. A serious disadvantage of using variations in years of schooling across individuals is the possible presence of education-specific cohort trends. Specifically, school premium might have gone up over time in Morocco as has happened in most countries. If this is true, it would raise the premium to years of schooling for younger cohorts relative to the older ones and hence downward bias the results. Moreover, one of the objectives behind language transition policies is to increase the accessibility of education to children from disadvantaged backgrounds making them more likely to join and stay in school. [8] If the Moroccan language policy indeed generated this type of endogenous schooling response, then individuals from younger cohorts would have lower wages than individuals with equal years of schooling from older cohorts due to their more underprivileged family backgrounds. This would upwardly bias the estimated effect of French skills in Morocco. In this paper we use district level variation in the exposure to the policy to overcome the endogeneity problems associated with using individual level years of schooling and a triple difference strategy to account for confounding trends.

Primary school language policy is relevant for many developing countries which were former American or European colonies. However, the case of India is particularly interesting in the light of its extensive linguistic diversity and the large-scale liberalization efforts undertaken in the recent decades.[9] The debate about learning English is at least a century old in India. In his writings Mohandas Karamchand Gandhi recalls that he often had private discussions about the desirability of giving children an English education. In his words, "parents who train their children to think and talk in English from their infancy betray their children and their country". These debates were later discussed in public forums where proponents of the opposite school of thought, Rabindranath Tagore being one of them, argued that preventing children from learning English would spoil their future - "if children were to learn a universal language like English

[8] In the context of India, a recent paper by Roy (2003) shows that there is not much evidence of relative improvement in school enrollment or attendance rates due to the abolition of English language learning from Primary schools in the Indian state of West Bengal.

[9] There are 22 official languages in India.

from their infancy, they would easily gain considerable advantage over others in the race of life" (Guha, 2011). However, since independence from British rule in 1947, these disagreements formed a part of the official language policy discussions and periodically resurfaced both in the national political arena and at the primary school level. While Hindi is recognized as the official national language by the Constitution of India, English has continued to be the primary medium of communication, particularly in white collar jobs. The debate over promoting indigenous languages versus English in schools was further fueled in recent times by the expansion of high-skilled export jobs following increased integration of India with the world economy. If English skills are indeed at a premium, then excluding it from public schools will reduce economic opportunities for the poor. From a public policy perspective it would mean a rethinking of previous policies which might have lost their initial relevance in the age of globalization.[10]

The rest of the paper is organized as follows. Section 2 provides a brief outline of the background of education policy in India. Section 3 discusses the possible endogeneity concerns and the identification strategy. Section 4 describes the data used in the analysis. The results of the empirical estimation are then discussed in section 5. Section 6 explores the effect of the policy on occupational choice. Section 7 draws a summary and concludes the paper.

2. Policy Background

Under the Constitution of independent India, education falls under the joint domain of both the State and Central Government of India. While general guidelines and funding is provided by the central government, policies governing the education institutions fall under the purview of the state administration. As a result in many cases, education policies in India have been influenced by respective regional political ideologies. One of the major policies the state governments have experimented with is the position of English language in the primary school syllabus. In practice, various school administrations across India have adopted two variants of language policies: use of English as medium of instruction in schools; and teaching of English as one of the subjects. The former is practiced only by a handful of private schools in the country. The second variant, teaching English as a subject, is commonly observed in private and government schools. However the grade at which English is introduced as a subject differs

[10] While a few state governments in India have repealed old policies and introduced English education to primary classes in public schools recently, these are seldom driven by any systematic evaluation of old policies.

across states and school administrations. In some states English is taught from the first grade while in some English is not taught in primary schools at all.[11] In independent India, education policy in West Bengal required the state government schools to teach English in primary school from the first grade while Bengali, and in a few cases Hindi, remained as the medium of instruction for all other subjects. However in 1983 teaching of English was abolished in primary grades of government schools.[12] Private unaided schools and government aided private schools technically remained outside the purview of the policy since they are privately managed and hence not mandated to follow managerial guidelines of the government.[13] With the new policy, English was taught as a subject only from grade 6 when students entered secondary school.[14] However, students who were already enrolled in primary school before 1983 continued to learn English as before. Thus, children entering primary school after 1983 did not learn English in primary school. Since the entry age at primary school is 6 years, this meant that children under the age of 6 in 1983, i.e. children born post 1977, were the ones affected by the policy change. Specifically, those who were born after 1977 and attended a government school did not learn English in primary grades. Children born before 1977 were not affected by the change as they would have entered primary school before 1983.

The change in 1977 was brought about by the newly elected communist government in the state who came to power for the first time that year. The purpose of the change, as pointed out by the then policymakers, was the perception that English is an elitist language from the colonial era which discouraged school participation of children from disadvantaged background. They argued that abolition of English from primary school would increase enrollment and rate of

[11] In India, primary school education typically covers grades 1-5

[12] The policy was scaled back in West Bengal in 1999 when English was reintroduced from grade 3 and was then completely repealed in 2004-05 when it began to be taught from grade 1 itself.

[13] There are three types of school in India: government (run by the government), aided (run by private management but largely government funded), and private unaided (Kingdon, 2008). We categorize schools as Public (run by the government) and Private (Aided and Unaided) to capture the difference in the adoption of the English policy. We use the terms "Public school" and "Government school" interchangeably in this paper. It is possible that some private aided schools might have been pressurized by the government to adopt the ban. However, we assume that all aided private schools continued to teach English and put them with the private unaided schools in the control group. In doing so, even if some aided schools did switch to no-English, while we treat them as teaching English, then our estimated would only be downward biased.

[14] Abolition of English could have freed up time for additional coursework. While there was no instruction from the government on how to use these hours, schools could use the extra time now on teaching extra Math instead of English, for example. However, this would only imply that our estimates provide a lower bound for the returns to English.

school completion and hence improve the educational standard of the population and reduce inequality.[15]

However, what the policymakers failed to acknowledge was the value of English skills that already existed in the domestic labor market. Indeed with liberalization, as in many of the emerging economies, English has become a lingua franca in the global as well as the domestic labor market. For example, it is widely believed that the preexisting knowledge of English has helped India emerge as the single largest destination for Information Technology Enabled Services by 2004 (Shastry, 2011). Thus investment in English skills has resurfaced as an issue of utmost importance within the domestic context of many developing countries. In India, the increase in employment probability for those with English skills has resulted in an overwhelming support from the parents to make their children get English training starting from elementary schools. A survey conducted in 2003 by the Regional Institute of English, South India (RIESI) found that more than 90% of the parents believed that learning English would help their children improve social mobility and get access to better job opportunities. It is widely believed that service sector liberalization has led to a steep rise in white-collar wages in India benefiting only the English-educated.[16] This inequality might be alleviated if individual investment in human capital responds to the changes in the labor market. However, poor households may not be able to respond to these changes to take advantage of the global opportunities. Higher returns to English skill will result in private English training to remain at a premium too. Individuals who can afford private schooling and coaching would acquire the necessary skills to find jobs requiring English skills. This in turn would exacerbate the existing inequality. India's liberalization experience provides an excellent opportunity to revisit the debate on the optimal language policy in primary schools.

3. Identification Strategy and empirical specifications

We use the exogenous education policy shift in West Bengal to identify the returns to English skills, in the backdrop of India's large scale liberalization program. Since the policy was applicable only to those children who joined the first grade after 1983 (those already in school in

[15] Roy (2003), shows that the policy failed to achieve its desired objectives in terms of greater enrollment or higher school completion rates.

[16] Munshi and Rosenzweig (2006) show that the English premium increased for both men and women from 1980s to 1990s ranging from 10% for men and 27% for women in Bombay.

1983 were unaffected), there is a variation in policy exposure across cohorts. Secondly, since the policy was implemented only in public schools, students who were more likely to go to a public school were also more likely to be affected by the policy.[17] However, individual level schooling decisions might be correlated with family background variables. Hence we construct a district (region) level probability measure of an individual's exposure to English learning opportunity as a proxy for English skills. Ideally we would want to instrument English skills of individuals by the policy change. However, it is difficult to find a comprehensive measure of English skills of individuals who are currently in the labor market. Hence we restrict our attention in this paper to the reduced form estimates of the effect of the policy on labor market outcomes. Nevertheless, the estimated coefficient from the reduced form is of interest in its own right. It contributes directly to the policy debate in school systems, across India as well as other countries, concerning the effect of introducing foreign language courses in primary school.

Furthermore, in Appendix Table 1 we provide some suggestive evidence on the effect of learning English in primary school on English skills of individuals using the India Human Development Survey (IHDS). IHDS is an India wide household survey conducted in 2005 which collected self-reported data on individual's English ability (Azam et al. 2010).

We compare English ability of children who attended government primary schools with English ability of children who attended a private school (aided or unaided) in primary grades during a period when the English ban was still effective in government schools. Since the policy was revoked in Bengal starting from 2004, we consider only children who joined the first grade before 2004. Column 1, shows that a child is 18 percentage points more likely to be able to speak in English if she attends a private school as opposed to a public school, with no English training in primary grades. Column 2 disaggregates the school types further to see if children in private aided schools have similar English skills as those in public schools, which would be the case if the aided schools also observed the English ban (as discussed in Section 2). While private unaided schools have a stronger impact on children's English ability, attending a private aided school also increases the probability of having English speaking skills by about 10 percentage points. Since we are primarily interested in the effect of learning English as an additional subject

[17] We include all privately managed government aided schools in our control group - the private school category – assuming that all those schools continued to teach English from the first grade. However, note that if some of these schools adopted the ban, our estimate would be a lower bound of the English premium in India.

in primary grades as opposed to the effect of using English as a medium of instruction we exclude in column 3 the schools with English as the medium of instruction – a very small fraction of private schools . Interestingly, private aided schools and private unaided schools that only teach English as a subject are equally effective in terms of imparting English skills as compared to public schools. In column 4 we control school hours per week and private coaching usage since children attending government schools might take up additional private English coaching in the absence of English in schools. They might also have fewer schools hours if the English ban is not substituted by additional coursework. Finally, column 5 restricts the sample to secondary school children and thus those who would have been exposed to the full effect of not learning English in primary grades if in public school. While it is difficult to infer any causal effect of the policy on English skills, these results at least provide some suggestive evidence that the not learning English in primary grades is associated with lower English skills of individuals.

Our analysis proceeds in two steps. First, we compare individuals across districts (regions) and cohorts with varying degrees of policy exposure within West Bengal. Second, we introduce the control states of Haryana and Punjab and account for differential district-cohort effects.

3.1 Intensity of Policy Exposure

We exploit the potential exposure of an individual in a specific district, or region, to public school at the time of the policy change and match that with labor market outcomes of individual in 2004. Since the new policy mandated public schools to abolish teaching of English in primary grades whereas the private primary schools were outside its purview, the probability of public school exposure proxies for the probability of learning English.

The measure of public school exposure is a probability measure of individual *i* having studied in a public school in district *d (*or region *r*) in 1983. We construct the probability of attending a public school using region level enrollment figures from National Sample Survey (*NSS*) data as follows,

$$IP_r^E = G_r^E / N_r^E$$

where, G_r^E is the number of students enrolled in public schools in region r in 1986. N_r^E is the corresponding total number of students enrolled in public and private schools. IP_r^E is the Public School Enrollment Measure – the percentage of students enrolled in public schools and hence affected by the policy change. One difficulty with this estimate is that the National sample survey is representative only at the region level, an administrative boundary bigger than a district, and thus generates very little variation in the causal variable (there are only four regions in West Bengal). Alternatively we use the data from the All India Education Survey (AIES) with information at the district level. However, AIES provides information only on the number of public and private schools but not on enrolment. Hence we construct a second measure of public school exposure, and call this the public school intensity measure,

$$\boldsymbol{IP_d^S = G_d^S / N_d^S}$$

where, G_d^S is the number of public schools in district d in 1986. N_d^E is the corresponding total number of public and private schools. IP_r^E measures the percentage of public schools in a district reflecting the potential probability of a person attending a public school. Table 1B reports the average probability of attending a public school based on these two measures. For all three states combined, the average probability of being exposed to the Language Policy change, according to the Public School Intensity measure, is 54%. According to the Enrollment measure, at the region level, it is 44%.

We construct our two public school exposure measures based on the number of public schools and school enrollment data for the year 1986-87. It is the earliest year after the policy change for which we have detailed district level school-type wise educational data available. However, since the year of data collection, 1985, is very near to the policy year, we are less concerned about the potential problem of new private schools being set up in response to meeting the increased demand for learning English. A time lag generally exists before the supply of new private schools can catch up with the increased demand. Most private schools have to be approved by the state board of education, whose members are appointed by the state government. It is unlikely that these members would allow an unfettered expansion of private schools as it would undermine the very policy of the state government. In other words, the supply of private

schools would not have responded to the demand for them in this short time (Roy, 2003). For the public school enrollment measure, we use the 42nd Educational round of *NSS* (1986-87) for similar reasons.

3.2 Two-Way Fixed effects Model

Our first estimation strategy uses the variation in treatment intensity across districts and cohorts to identify the effect of English language skills on individuals' labor market outcomes.[18] The younger cohorts are the ones deprived of English training in the primary school. Moreover, the higher the probability of attending a public school, lower is the probability of learning English. Thus, if lower English skills are associated with lower wages, the difference in average wages between the older and the younger cohorts will be negatively related to the probability of attending a public school

$$W_{icd} = \alpha_1 + \alpha_2 IP_d^S * Post + D_c + D_d + \alpha_3 X_i + e_{icd} \qquad (1)$$

Where, W_{icd} is the wage outcome of individual i born in district d and cohort c. IP_d^S is the intensity of public schools in district d at the time of the policy change. '*Post*' is a dummy indicating whether individual i is affected by the policy change. It takes a value 1 if an individual enters school in or after 1983 and 0 otherwise. Thus (IP_d^S * *Post* **)** measures the intensity of exposure to public schools for individual i of cohort c and district d. X_i includes individual level potential predictors of labor outcomes like experience, experience-squared, education and gender. e_{icd} includes unobserved determinants of the outcome variable. D_c is a cohort of birth dummy. It accounts for labor market changes that vary across cohorts and hence differences out any time trend that might have affected the pre and post-policy cohorts differently. Controlling for cohort trends reduces the likelihood of the effects of the policy change being confounded with other changes that occurred over time. D_d is the district dummy that accounts for district specific characteristics that might affect individuals in the high and low public school-intense districts differently but are time invariant. This two-way-fixed-effect model compares wage outcomes for cohorts entering school before and after the policy change and between districts

[18] This strategy is similar to Card and Krueger (1992) or Card and Thomas Lemieux (1998). More recently it has also been used by Duflo (2001) to the study the impact of school expansion on education and wages.

with a high and low probability of learning English. We cluster the standard errors at the district level. α_2 measures the impact of abolishing English education on wages. If English skills have high returns in the labor market, we would expect α_2 to be negative.

One concern is that the national household sample survey (of the NSS) from which we get the wage data does not collect information on the childhood residence of individuals. Hence, we cannot observe whether the current employment location of individuals is the same as their childhood residence where she underwent schooling. However, estimates based on the 2001 Census of India shows a very low average decadal rate of migration across districts (3.3% for West Bengal and 4% for the inter district migration for the three states combined that we use in our sample). In addition, Topalova (2005) notes that less than 0.5 percent of the population in rural and 4 percent of the population in urban areas moved for reasons of economic consideration (or employment). Thus district of current residence (or of employment) of an individual can be considered to be approximately the same as the schooling district.

3.3 District-specific time trends

The causal interpretation of α_2 in the above framework rests on the assumption that after controlling for district and cohort fixed effects, e_{icd} is independent of the interaction term. In other words, it assumes that there are no time varying district-specific factors that are correlated with our measure of policy exposure. However, the allocation of public schools across districts is likely to be influenced by the local government officials. If more efficient officials attract higher investments not only in education but also in other development areas, then districts with higher number of public schools might also experience a higher labor market growth over time which would downward bias our estimate of α_2. Indeed, Muralidharan and Kremer (2008) show that regions with higher per capita income are less likely to have private schools in India. Another confounding factor might be the growth of private coaching centers in response to the policy transition. Roy (2004) shows a considerable growth in private coaching and tuition in West Bengal after the policy change. Districts with a higher fraction of public schools, and hence fewer options of learning English in schools after the policy change, are likely to have a higher demand for private options. While growth of private schools is restricted by the government (Roy, 2004), these districts might still have a higher growth in private coaching centers. If true,

the differential growth of private coaching centers across districts will also downward bias our two-way fixed effects estimates.[19]

The estimates of α_2 might thus be threatened by the existence of district-cohort trends. As mentioned earlier, education policies are governed by state authorities and the policy under review was only implemented in West Bengal. So we use as controls other states that did not have any change in education policy at the same time as West Bengal, to control for the differential district-time trends. Specifically, we use Punjab and Haryana as the control states that continued to have English from the first grade in their public schools at the time when West Bengal experienced the change in its language policy. While many other states would qualify as control group, with no change in English teaching policy in schools around the same time as West Bengal, our choice of states is restricted to Punjab and Haryana by the limited availability of data and information regarding policy changes. As before, we compute both measures of public school exposure for these states and estimate the following regression.

$$\boldsymbol{W_{icd} = \beta_1 + \beta_2\, IP_d^S * Post * WB + D_{dc} + D_c + D_d + WB*D_c + \beta_3\, X_i + e_{icd}} \qquad \textbf{(2)}$$

In this regression $\boldsymbol{\beta_2}$ gives the causal estimate of the effect of language policy in West Bengal on wage outcomes after controlling for state, district and cohort trends and their interactions. IP_d^S , *Post*, D_d , D_c and X_i are defined as before. WB is an indicator that takes value 1 if individual *i* was born in the state of West Bengal and 0 if belongs to either of the control states: Haryana or Punjab. D_{dc} denotes the district-time trends that account for any differences in trend between the high and low public-school-intense districts apart from the English Language policy. Moreover, there might be difference in the growth pattern of West Bengal and the control states of Haryana and Punjab. Specifically, post liberalization, the higher growth of export oriented jobs in the control states of Punjab and Haryana compared to West Bengal might upward bias our estimates. Thus we include the time varying state effects, $WB*D_c$, that differences out all such state specific time varying factors.

[19] The greater growth of alternative English training centers in response to the abolition of English teaching in public schools can be thought of as an indirect impact of the policy and hence should be a part of the policy's general equilibrium effect. However, in this paper, our aim is to estimate the English skill premium using the policy as an exogenous shock, rather than evaluating the policy.

4. Data

Our data comes from two sources: The All India Educational Survey (AIES) and the National Sample Survey (NSS) provided by the Government of India. The AIES, conducted every 5-7 years, is a census of schools in India and provides district level on the number of public and private schools. The information is collected and disseminated separately for each state. The district level public school exposure measure (see Section 3.1) is constructed using the AIES 1986 round. The states of Punjab and Haryana are the only two states in the treatment group for which the state level documents were available from this period. For the region level Public School enrollment measure we use the education round of NSS (1986).

The individual level data comes from the NSS's Employment and Unemployment Survey (Schedule 10). The Employment and Unemployment rounds are 5-yearly surveys and are divided into four sub-rounds and covers both urban and rural areas. The survey includes information on household characteristics like household size, principal industry-occupation, social group and monthly per capita expenditure. It also includes detailed demographic information including age, sex, marital status, location, educational level, school attendance, occupational status, industry of occupation for those employed, as well as a daily time disposition. The survey adopts a stratified two-stage design with four sub-rounds in each survey year.[20] For this paper, we pool the data from the 55th round and the 61st round since these are the only two rounds that allow us to observe the relevant cohorts entering primary school before or after the policy change.

We restrict our sample to the working individuals in the age group 17-45 at the time of the NSS 2004 survey.[21] Individuals who are below 17 yrs in 2004 would not be in the formal labor market that requires any knowledge of English. This also excludes the possibility of child labor. In India, children begin primary schooling at the age of 6. Thus individuals born in 1976 and before would not be affected by the policy change since they would have entered primary school before 1983, the year of policy shift. Hence, the effect of the program should be felt only by those born after 1977 and hence aged 6 years or below in 1983. Individuals who are born after 1977 would be 17-22 years in 1999 (55th round of NSS) and would be 17-27 years in 2004-05.

[20] The first-stage units in the sub rounds are census villages in the rural sector and the NSSO urban frame survey (UFS) blocks in the urban sector.

[21] The results reported are not sensitive to different birth cohort windows.

These individuals who potentially joined school in the post policy period form the treatment group in our analysis. The upper cutoff age, 45 years, generates a comparable control group to our treatment group in our estimation strategy. Specifically we compare our treatment group to individuals in the age group 23-40 in 55th round (1999-00) and those in the 28-45 age group in 61st round (2004-05). Some individuals, born towards the end of the control period, could have started primary school at a later age and thus may have been exposed to the policy change biasing our estimates. However, when we repeat our analysis excluding the years of 1974-1976 from the control group, we get very similar results.

The labor market outcomes that we consider are wages and occupational choice. We deflate the weekly wages from NSS 55th and NSS 61st rounds in terms of 1982 Indian rupees using the consumer price index for industrial workers to be able to compare NSS 55th and 61st round samples. Wages are expressed in terms of total real weekly earnings.

For analyzing the occupational choices, we use the National Occupational Classification (NOC) at the one-digit level and put them into the following six broad categories following Kossoudji (1988): PROF- Professional Technical and Kindred Workers (NOC 1digit code 0-1); MNGR-Administrative, Executive and Managerial (NOC 1digit code 2); CLER- Sales and Clerical Workers (NOC 1digit code 3-4); CRAFT-Craft and Kindred Workers (NOC 1digit code 6); OPER-Production Workers and Transport Operatives (NOC 1digit code 7-8-9); SERV-Service Workers and Laborer (NOC 1digit code 5).

4.1. Descriptive Statistics: need to rewrite depending on the new table with Punjab and Haryana

Descriptive statistics are reported in Table 1A. For the treatment state, West Bengal, the average age in our sample is about 31 years with an average age at entry to school of approximately 6 years. For the control states of Haryana and Punjab that we use for the triple difference estimation, the average age in the sample is 30 years and the average age at entry to school is again approximately 6 years. Mean job experience is 8.5 years in West Bengal, while the mean job experience is about 8 years in the states of Punjab and Haryana[22]. About 25% of the sample was illiterate (or below primary educated and/or no formal schooling) in all the states. The

[22] Potential experience is calculated using the definition *job experience=minimum {age-15, age-age at highest education}*.

distributions of education and occupation are also quite similar across the sample in West Bengal and the treatment states of Haryana and Punjab. Overall, the treatment and control states are not significantly different from each other in terms of mean characteristics in 2004-2005. On the other hand, average weekly wages in 1982 Indian Rupees was 71 in West Bengal compared to 91 in Haryana and 88 in Punjab[23].

5. Results

5.2 Average Impact using English learning Probability

As discussed earlier, intensity of exposure to the English language policy varies with the concentration of public schools in a district. So we combine cohort variation with our district (region) level measure of policy exposure to identify the effect of English skills on labor market outcomes.

The results from the estimation of model (1) are reported in Tables 3 and 4. Table 3 uses the district level intensity measure while Table 4 uses the region level Enrollment measure. Since older individuals would have been in the market for a longer time and hence earn higher income than the younger cohorts by virtue of their experience, each column controls for years of work experience and a quadratic in years of experience. We also include dummies for different social groups that each individual belongs to (Schedule Caste/Tribe and others) in all our regressions. We cluster the standard errors for any within district correlations. Column 1 of Table 3 shows the results after controlling for district fixed effects and a post-treatment dummy that accounts for a possible difference in trend, apart from the policy, between the post and pre treatment cohorts. Individuals who are more likely to be affected by the policy get lower wages compared to the individuals in the control group. Specifically, an individual who is 1% less probable to learn English in primary school gets approximately 0.08% less wage.[24] Column 4 shows the results from our model in equation (1) where we control for individual birth cohorts. The results are similar after controlling for individual birth year effects instead of a post-treatment dummy, although the estimates are not precise.

[23] The current exchange rate between Rupee and Dollar is approximately 51 INR to 1 USD.
[24] Evaluated at the mean public school intensity of 32%

The estimation with our Enrollment measure can only be conducted at the region level as the survey data from which we construct the measure is representative only at the region level. Since region is an aggregation of districts, there are only four regions in West Bengal as opposed to seventeen districts. However, even with the reduced variation in the likelihood of attending a public school, we find similar results as in the case of our district regressions. The estimates reported in Column 1 and 4 of Table 4 (with a common post-treatment trend and individual birth cohort effects respectively) suggest a similar negative impact of the language policy on wages of individuals who are more likely to be affected by the policy. Again, the estimates suggest roughly a 0.08% decrease in wages due to a 1% increase in the probability of attending a public school. Overall, both at the district and the region level with different measures of the exposure to the English language policy, the estimates suggest relatively lower wages for individuals who went to primary school after the abolition of English in areas with higher intensity of public schools. These estimates imply about 2.5-3.5 % lower wages for cohorts exposed to the English abolition policy in the average district or region.[25]

5.2.1 Heterogeneity of Impact

One problem with the two-way fixed effects analysis is that younger cohorts in districts with higher private school concentration (or lower public school concentration) could be earning a higher return to human capital due to higher labor market growth in these districts. This means the two-way estimates do not truly reflect the effect of the language policy. However, better labor market conditions would affect all individuals in these districts while a language policy in school would only affect those individuals who completed some threshold level of schooling necessary for white collar jobs requiring any knowledge of English. This implies a simple check for the validity of the two-way fixed effects results. Specifically, the results should not hold for those individuals who would theoretically be unaffected by the language policy but would still be affected by any other district wide changes. Table 3 shows the estimates separately for those with less than primary schooling or no schooling and those with more than primary schooling at the district level. Columns 2-3 control for a Post Dummy while Columns 5-6 is a replication of model (1) with individual birth cohort dummies. The results in Column 3 and 6 indicate a very

[25] At the district level, we obtain estimate of average difference in wages by multiplying the average probability of not having learnt English (32%) in West Bengal by the elasticity measure of 0.08.

strong negative effect of the policy on individuals who are expected to be affected by a change in the language policy, specifically those who completed some threshold level of schooling. In this case, a 1% reduction in exposure to English language in the primary school leads to approximately a 0.35% reduction in wages. Table 4 shows the analogous results at the region level. The estimates are smaller than at the district level implying a 0.2% reduction in wages for individuals with more than primary education and exposed to the policy change.

If the two-way results were completely spurious, driven for example by differential growth in labor markets, we would expect similar results for all individuals, irrespective of their eligibility for jobs requiring English skills. The results in column 2 and 5 of Tables 3 or 4 respectively suggest otherwise. The coefficients are either very small or positive. In general the results imply a lower wage outcome only for individuals who completed more than a primary level of schooling and were exposed to the language policy change. These results are also in line with the findings of Angrist and Lavy (1997). They find no wage premium due to French skills in Morocco for having a primary school education but significant language premium for individuals with secondary schooling.

Although these results are suggestive of the negative impact of the policy on individuals who are most likely to gain from English education, they are not definitive evidence. There is always a possibility that the return to education might have declined over time due to liberalization, driving the results for the better educated individuals. Moreover, the positive coefficient on the below primary education group possibly reflects that overall wages would have grown more in the regions with greater fraction of public schools in the absence of the language policy.

5.3 Differential District Trends

While estimates from the two-way fixed model and the subsequent robustness analysis suggests that revoking English from primary school reduced wage outcomes of individuals exposed to the policy, the robustness check does not rule out the absence of time varying district specific effects correlated with the measure of policy exposure. As discussed earlier, allocation of development funds over time might be skewed towards districts that also attract higher education funds. Hence districts with higher public school concentration might have experienced a higher economic

growth. In the absence of the language policy this would imply higher wages for individuals in districts with more public schools which will underestimate the program effects. The consistency of the estimates would also be violated if growth of private English coaching centers responds more to the policy transition in districts with fewer alternatives of private schools. To see if indeed there is a differential trend across the treatment and control districts we conduct a falsification test. Table 5 reports the results of the control experiment using two types of cohorts. Column 1-2 sets the pseudo experiment on cohorts, none of whom was affected by the policy change. Individuals born between 1950 and 1974 entered school prior to the start of the language policy. Column 3-4 sets the pseudo experiment on cohorts who were always affected by the policy change. Individuals born between 1977 and 1987 entered school after the start of the language policy.[26] The results in columns 1 and 3 suggest spurious positive treatment effects. The positive significant coefficients on the interaction term imply a positive wage premium for individuals from districts with a higher concentration of public schools, in the absence of the language policy. This provides clear evidence on the presence of confounding effects that might be biasing the two-way estimates. To correct for these confounding district specific trends we compare our two-way fixed effects estimates to estimates from other states that did not experience any change in their education policies.

5.4 Controlling for District Trends

The estimates of model (2) are reported in Table 6 (district level) and 7 (region level). As before all regressions include controls for job experience, a quadratic in experience, and the social group of the individuals. The main coefficient of interest in these specifications is that of the triple interaction term (IP_d^S * *Post* * *WB*). The results indicate that controlling for district-specific time trends generates a larger impact of English skills on labor market returns. This implies that the coefficients of the two-way fixed effects model that do not account for the simultaneous positive district trends underestimate the true program effect.

The results indicate a significant negative impact of the Language Transition Policy on future returns in the labor market for any specific level of education. Individuals who went to school in West Bengal after the introduction of the Language policy in districts with a higher

[26] For the post treatment cohort the widest window we can consider is that of 10 years since 1987 born are the youngest cohorts who would be in the labor market in 2004

probability of attending public schools earned relatively lower wages. The coefficient estimate of 1.671 in table 6 suggests that a decrease in the probability of learning English by 10% lowered weekly wages, in 2004, by approximately 9% for cohorts born in West Bengal in the post policy period. The average proportion of public schools in West Bengal implies that cohorts attending primary schools in West Bengal in the post policy period have on an average a 32% lower probability of learning English. Thus on average revoking English language instruction from public primary schools lowered wages by 29%. Evaluated at the average proportion of enrollment in public schools implies a 39% English premium. For individuals with at least primary schooling the English premium is approximately 42%.

Table 7 presents the results with enrollment measures after controlling for region-specific time trends. The results are smaller in magnitude compared to the district level regressions but similar in spirit.

5.5 Sample Selection Bias

The results discussed in the previous section are based only on the sample of wage earners, who comprise approximately 43% of the individuals in our combined sample of the three states. The probability of working for a wage might depend on the ability to speak or write English. If English skills have positive influence on both employability and wages, then individuals with less exposure to English will on average have lower wage offers and a lower probability of selection into wage-earner status. As a result amongst the group of people who have less exposure to English, our sample will capture individuals with comparatively high wage offers.[27]

This implies that selecting only the wage earners is likely to violate the normality assumption on the error term with respect to the policy indicator (the interaction term). To address this selection bias, we re-estimated our model using Heckman's sample selection procedure (1976, 1979). Specifically, an indicator of whether an individual is working for a wage is regressed on the policy indicator and other controls in the first stage, and polynomials of the predicted value from this regression are used as additional controls in estimating the wage equation (1). Controlling for the probability of selection does not significantly alter our estimates

[27] This will lead to a downward bias, implying that our coefficients will be a lower bound to the estimates of English premium.

of the English Premium. Thus we do not encounter any severe selection problem by restricting the sample to wage earners.

6. Occupational Attainment Estimation

Finally, it is important to understand the channel through which the difference in wage arises between the English skilled and unskilled workers. If different remunerations accrue to workers with and without English skills within the same occupation then the gap might close over time with on-the-job training opportunities. However, if the difference is due to selection into different occupations, then it is unlikely that the difference will mitigate without policy targeting. Specifically, the ITES (Information Technology Enabled Services) sectors that emerged and grew as a result of the liberalization process is both more likely to hire English skilled workers and also are the sectors that offer relatively higher wages.[28] Thus the wage premium is possibly a result of inequality in the choice of occupations available to English-skilled and unskilled workers. In addition, lack of English knowledge may create search costs which may then change the order of occupational preferences or access to certain jobs. Occupational movement may be restricted and individuals may take up jobs for which they may be over qualified in all other aspects. Promotion and movement up the job ladder may be prevented as employers may not consider those not educated in English as trainable for higher ranked jobs.

To shed light on the mechanism responsible for the divergence in wages, we study the impact of English skill on occupational outcomes, using a multinomial model of occupational attainment. We assume that an individual's probability of attaining one occupation relative to another is independent of the presence of other possible occupations. So the multinomial logit model predicts the probability of an individual falling into one of the occupational groups relative to another group.

The empirical specification involves specification variant of the model in equation (1):

$$Log\ (P_j/P_r)_{icd} = \delta_1 + \delta_2\ IP_d^S * Post + D_c + D_d + \delta_3\ X_i + e_{icd} \qquad (3)$$

[28] Occupation of the employed individual is not included in the wage equation as it is considered a grouped variable of the wage variable. Instead both wage and occupational attainment outcome are taken as a measure of labor market outcome.

where the dependent variable measures the log odds of working in occupation category *j* relative to occupation category *r*. We construct an ordinal ranking of the occupations based on the skills they require and the average wages they pay. The ranking in descending order is: PROF, MNGR, CLER, OPER, SERV and CRAFT. IP_d^S is the district level exposure to public schools as measured by the public school intensity measure. *Post*, D_c , D_d , X_i are defined as before.

The coefficients of interest are given by δ_2 . They can be interpreted as the odds of working in one occupation relative to another as a function of the individual's exposure to English training when young.

A negative (positive) value of δ_2 implies that individuals with lower degree of policy exposure or a higher probability of learning English in school are more (less) probable to work in a higher ranked occupation. Table 8 reports the multinomial coefficients of the interaction, δ_2 , estimated from model (3). Column 1 reports the estimation results from the full sample of West Bengal, without separate education categories. Column 2 reports the coefficients for above primary-educated individuals, the group of primary interest for the purpose of this study.

When we consider all individuals, which includes illiterates and literates, most of the coefficients are negative with some of them significant at 5% level of significance. As in the wage regressions, English seems to be particularly important in deciding occupational choice for individuals with more than primary education. Specifically, for better educated individuals, greater exposure to English significantly raises the probability of joining a higher ranked occupation relative to craft. For example, row-1, column-2, shows that for individuals with more than primary schooling a 1% increase in exposure to public schools in the post policy period leads to a decrease of 4.7% in the log odds of working in a professional occupation compared to craft and kindred occupation category.

This higher (lower) likelihood of working in a higher ranked occupation as a function of higher (lower) exposure to English education shows that English language acquisition is an important determinant of occupational attainment of individuals. This suggests that the high English premium in the labor market is possibly driven to a large extent by the lack of occupational mobility for individuals with little or no English skills but otherwise similar educational attainment.

7. Conclusion

English is increasingly valued in the labor market in this era of globalization particularly with liberalization of the services sector. In this paper we estimated the returns to English skills in a globalized Indian economy by using an exogenous change in English learning opportunity. The results suggest that individuals who are more likely to have training in English earn significantly higher relative wages and better occupational outcomes even for the same level of overall education. This means that returns to specific skill sets could increase inequality further if policies are not targeted towards labor market requirements. This result is particularly relevant in the context of many developing countries which face the dilemma of whether to encourage local or global languages in primary schools. Choosing a local language might generate cultural benefits but it is generally at the cost of attaining higher economic benefits from liberalization. Moreover, discouraging global languages in public schools could aggravate inequality within developing countries by widening the gap between the elites and the poor who are unable to respond to global opportunities. More importantly, it might be inefficient to adopt such policies as they drive the economy towards a less efficient outcome. While a primary aim of teaching only local languages in primary schools is to reduce inequality by providing greater access to education, there is little evidence on higher enrollment following such intervention. Roy (2003) investigates the same policy but finds no improvement in enrollment, years of education completed or age at entry to school. Together with the results of this paper, it suggests that such regressive policies might actually increase inequality.

Interestingly, females constitute a significant proportion of the workers in the business processing industry which typically require English skills. According to NASSCOM 2004, the male-female ratio in business processing firms was 35:65. This implies that introducing English in public schools might also help females proportionately more than males, hence narrowing the male-female gap in labor force participation or wages (refer to footnote 15). As a part of future research, it would be interesting to measure whether labor market outcomes were affected disproportionately for women due to the said policy change.

Bibliography

Anandabazar Patrika, 2003. The School Education Committee recommends English only from grade 5. 1st January, (Anandabazar).

Angrist, J., Chin, A., Godoy, R., 2008. Is Spanish-only schooling responsible for the Puerto Rican language gap?. Journal of Development Economics. 85, February, 105-128.

Angrist, J., Lavy, V., 1997. The effect of a change in language of instruction on the returns to schooling in Morocco. Journal of labor Economics.

Bagchi, J., 1981. Storm over English in West Bengal: Comment. Economic and Political Weekly. May 30, pp. 993-994.

Betts, J., 1998. The impact of educational standards on the level and distribution of earnings. AER. Vol. 88, No. 1, pp. 266-275.

Carliner, G., 1981. Wage differences by language group and the market for language skills in Canada. The Journal of Human Resources. Vol 16, No. 3.

Carliner, G., 2000. The language ability of US immigrants: assimilation and cohort effects. International Migration Review. Vol. 34, No. 1 (Spring), pp. 158-182.

Chin, A., 2005. Can redistributing teachers across schools raise educational attainment? Evidence from operation blackboard in India. Journal of Development Economics. 78, 384-405.

Chiswick, B.R., Miller, P.W., 1995. The endogeneity between language and earnings: international analyses. Journal of Labor Economics. 13:246–88.

Chiswick, B.R., Miller, P.W., 2002. Immigrant earnings: language skills, linguistic concentrations and the business cycle. Journal of Population Economics.

Chiswick, B.R, Taengnoi. S., 2007. Occupational Choice of High Skilled Immigrants in the United States. IZA Discussion Paper No. 2969.

Dustmann, C., Soest Arthur van, 2002. Language and the earnings of immigrants. Industrial and Labor Relations Review. 55:3, 473-492.

Guha, R., 2011. The question of English - worldwide, English remains the choice for communication. The Telegraph, Calcutta India, November 5.

Heckman, J.J., 2007. The economics, technology, and neuroscience of human capability formation. PNAS, Vol 104, No. 33.

Kingdon, G., 2008. School-sector effects on student achievement in India, in: Chakrabarti, R. Peterson P. (eds.), School Choice International: Exploring public-private partnerships. MIT Press.

Kossoudji, A. S., 1988. English language ability and the labor market opportunities of Hispanic and East Asian immigrant men. Journal of Labor Economics. Vol. 6, No. 2, pp. 205-228.

Lang, K., Siniver E., 2009. The return to English in a non–English speaking country: Russian immigrants and native Israelis in Israel. The B.E. Journal of Economic Analysis & Policy.

Levinsohn, J., 2004. Globalization and the Returns to Speaking English in South Africa. WP.

Mehtabul, A., Chin, A., Prakash, N., 2010. The Returns to English-Language Skills in India. IZA Discussion Paper No. 4802.

Munshi. K., Rosenzweig, M., 2006. Traditional institutions meet the modern world: caste, gender and schooling choice in a globalizing economy. AER. Vol. 96, No. 4.

Muralidharan, K., Kremer, M., 2008. Public and private schools in rural India, in: Peterson, P., Chakrabarti, R. (Eds.), School Choice International: Exploring public-private partnerships. MIT Press.

Roy, J., 2003. Abolition of Teaching of English from Primary Schools in West Bengal, India: An evaluation. Princeton University.

Schmidt, P., Strauss, P., 1975. Estimation of models with jointly dependent qualitative variables: A simultaneous logic approach. Econometrica. Vol. 43, No. 4, pp.745-755.

Shastry, G. K., (Forthcoming). Human capital response to globalization: Education and information technology in India. Journal of Human Resources.

Table 1A: State level descriptive Statistics based on NSS 1999 and 2004-05

Variable	West Bengal	Haryana	Punjab
Age (years)	30.69(8.30)	29.68 (8.28)	29.87(8.38)
Age at entry at school (years)	6.36 (3.09)	5.37(3.63)	5.94(3.89)
Job Experience (years)	8.56 (9.12)	8.38(8.88)	8.18 (9.07)
Social Group			
Proportion Backward	0.368(0.482)	0.471 (0.499)	0.504(0.50)
Proportion Females	0.402(0.490)	0.384(0.486)	0.393(0.488)
Weekly Wages (deflated in 1982 Rs)	71.04(148.69)	90.60(165.38)	87.94(165.15)
Education			
Percentage Primary	12.72	12.22	11.65
Percentage Middle	19.95	12.28	11.23
Percentage Secondary	13.41	20.79	21.74
Percentage High Secondary	11.03	14.36	13.86
Percentage Graduate and above	16.06	18.47	16.23
Percentage Others (Illiterates, Below Primary, Literate with no formal schooling)	26.83	21.88	25.29
Religion			
Proportion Hindus	0.803(0.397)	0.912 (0.283)	0.450(0.497)
Occupational Distribution			
Percentage PROF	8.42	9.80	9.18
Percentage MNGR	8.24	6.76	11.64
Percentage CLER	24.93	25.29	21.78
Percentage CRAFT	11.57	10.00	9.22
Percentage OPER	43.73	44.61	43.37
Percentage SERV	3.11	3.53	4.80

Table 1B: Average Probability of attending a public school

	Percentage Public School (AIES)	Percentage enrolled in public school (NSS)
West Bengal	0.3189 (0.2190)	0.4642(0.1576)
Haryana	0.8663(0.0997)	0.4465(0.0767)
Punjab	0.8693(0.1193)	0.4162(0.0459)
Three States Combined	0.5476 (0.3268)	0.4478(0.1276)

Note: Public school (affected by the policy) refers to government run schools and Private school (not affected by the English ban) includes both government-aided privately managed schools and unaided private schools. Estimates based on 1986 round of AIES and NSS respectively

Table 3: Two-way Fixed Effect with Public School Intensity Measure (West Bengal): District Level

Dependent Variable: log of real wage						
	Control for Post			**Control for Individual Cohorts**		
	(1)	(2)	(3)	(4)	(5)	(6)
	All Individuals	Below Primary Education	Above Primary Education	All Individuals	Below Primary Education	Above Primary Education
Public School Intensity*Post Policy Dummy	-0.246*	-0.122	-1.112*	-0.186	0.0470	-1.340**
	(0.14)	(0.314)	(0.535)	(0.14)	(0.280)	(0.625)
Controls						
Experience	0.0696***	0.0159*	0.0241	0.0978***	0.0293	0.0648***
	(0.0063)	(0.0088)	(0.0246)	(0.0082)	(0.0249)	(0.0184)
Experience	-0.0022***	0.0004	-0.0015	-0.00404***	0.0006	-0.0050***
	(0.0001)	(0.0005)	(0.0011)	(0.00041)	(0.0015)	(0.0008)
SC-ST	-0.330***	-0.0669	-0.174**	-0.303***	-0.0287	-0.0752
	(-.0321)	(0.0635)	(0.0606)	(0.032)	(0.0709)	(0.0524)
Post	Yes	Yes	Yes			
Birth Cohort Dummies				Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	4.570***	4.218***	5.200***	3.734***	3.852***	4.072***
	(0.026)	(0.0388)	(0.120)	(0.13)	(0.299)	(0.237)
Observations	2766	1243	1523	2766	1243	1523
R-squared	0.236	0.142	0.194	0.29	0.232	0.371

Clustered standard errors at district level in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4: Two-way Fixed effects with Enrollment Measure (West Bengal): Region Level

	Dependent Variable : Log of real wage					
	Control for Post				**Control for Individual Cohorts**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	All Individuals	**Below Primary Education**	**Above Primary Education**	**All Individuals**	**Below Primary Education**	**Above Primary Education**
Public School	-0.187	0.164	-0.448	-0.196**	0.114	-0.568**
Enrollment * Post Policy Dummy	(0.10)	(0.087)	(0.25)	(0.055)	(0.086)	(0.16)
Controls						
Experience	0.0699***	0.0197**	0.0140	0.0990***	0.0317**	0.0089
	(0.0039)	(0.0037)	(0.0100)	(0.007)	(0.0056)	(0.0060)
Experience square	-0.0022***	-0.0004	-0.0002	-0.0041***	-0.001**	-0.0022***
	(0.0002)	(0.0002)	(0.0004)	(0.0003)	(0.0003)	(0.0004)
SC-ST	-0.345***	-0.175***	-0.325**	-0.317***	-0.170***	-0.246*
	(0.043)	(0.024)	(0.091)	(0.035)	(0.028)	(0.080)
Post	Yes	Yes	Yes			
Cohort				Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	4.645***	4.313***	5.446***	3.755***	4.208***	4.039***
	(0.013)	(0.0027)	(0.078)	(0.020)	(0.029)	(0.057)
Observations	2766	1243	1523	2766	1243	1523
R-squared	0.23	0.11	0.20	0.30	0.17	0.35

Clustered standard errors at region level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Falsification Test

Dependent Variable : log real wage				
	1950-1974 (Unaffected cohorts)		**1977-1987 (Affected cohorts)**	
	All	**Above Primary**	**All**	**Above Primary**
Public School Intensity *Post	0.458*	0.631***	1.382**	2.135**
* West Bengal	(0.23)	(0.17)	(0.57)	(0.73)
Cohort	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Observations	2670	1606	839	361
R-squared	0.31	0.32	0.24	0.34

Note: Results from control experiments using cohorts who were never affected by the language policy change (In columns 1 & 2) and those who were always affected by the language policy change (in columns 3 & 4). Standard errors at district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: District Specific Trends: District Level (Punjab, Haryana & West Bengal)

Dependent Variable : Log of real wage		
	(1)	(2)
	All States	All States Above Primary
Public School Intensity *Post	-1.671***	-1.785***
* West Bengal	(0.079)	(0.053)
West Bengal * Cohort	Yes	Yes
Above Primary*District		
Above Primary*Cohort		
District*Cohort	Yes	Yes
District Fixed Effects	Yes	Yes
Cohort Dummies	Yes	Yes
Controls	Yes	Yes
Observations	5000	2023
R-squared	0.509	0.526

Clustered standard errors at district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Region Specific Trends: Region Level (Punjab, Haryana & West Bengal)

Dependent Variable : Log of real wage		
	(1)	(2)
	All Individuals	Above Primary Education
Public School Enrollment *Post Policy * West Bengal	-0.178**	-0.502***
	(0.078)	(0.089)
Region*Cohort	Yes	Yes
West Bengal * Cohort	Yes	Yes
Region Fixed Effects	Yes	Yes
Cohort Dummies	Yes	Yes
Controls	Yes	Yes
Observations	5000	2832
R-squared	0.303	0.345

Clustered standard errors at district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Two-way Fixed Effect Estimates of Occupational Choice

	All Individuals	Above Primary Education
(PROF/CRAFT)	-1.598	-4.715***
	(1.14)	(0.832)
(MNGR/CRAFT)	-0.198	-3.484***
	(1.03)	(1.033)
(CLER/CRAFT)	-0.230	-4.028***
	(0.865)	(0.538)
(OPER/CRAFT)	1.213	-3.557***
	(0.927)	(0.619)
(SERV/CRAFT)	0.125	-4.218***
	(0.99)	(0.617)
(PROF/SERV)	-1.723**	-0.497
	(0.869)	(1.02)
(MNGR/SERV)	-0.324	0.7337
	(0.659)	(1.02)
(CLER/SERV)	-0.356	0.1903
	(0.285)	(0.552)
(OPER/SERV)	1.087*	0.6606
	(0.585)	(0.643)
(PROF/OPER)	-2.811**	-1.157
	(1.14)	(0.867)
(MNGR/OPER)	-1.412**	0.0730
	(0.662)	(0.948)
(CLER/OPER)	-1.443**	-0.4702
	(0.65)	(0.348)
(PROF/CLER)	-1.367	-0.687
	(0.889)	(0.774)
(MNGR/CLER)	0.032	0.5433
	(0.711)	(1.086)
(PROF/MNGR)	-1.399	-1.230
	(1.06)	(1.276)
Birth Cohort Dummies	Yes	Yes
District Fixed Effects	Yes	Yes
Controls	Yes	Yes
Observations	3872	2186
Pseudo R^2	0.0876	0.0997

Note: Table 8- Clustered standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Coefficients reported above are the multinomial logit coefficients of the Interaction term of Public School Intensity Measure and Post Dummy on the log-odds of working in a specified occupation relative to another.

Clustered standard errors at district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix: Table 1

Dependent Variable : English Speaking Ability					
	(1)	(2)	(3)	(4)	(5)
	All schools	All schools	No English medium schools	No English medium schools	Above Primary School Age
Private	0.189***				
	(0.015)				
Pvt-aided		0.0955***	0.106***	0.101***	0.140***
		(0.024)	(0.022)	(0.022)	(0.034)
Pvt-Unaided		0.241***	0.0887***	0.0861***	0.207***
		(0.018)	(0.021)	(0.021)	(0.044)
Age	0.0183***	0.0192***	0.0166***	0.0164***	0.0253***
	(0.0016)	(0.0016)	(0.0015)	(0.0016)	(0.0042)
School-hours				-0.00129	-0.00228
				(0.00091)	(0.0015)
Pvt-tuition				0.00265***	0.00439***
				(0.00086)	(0.0014)
Constant	-0.164***	-0.174***	-0.146***	-0.129***	-0.250***
	(0.019)	(0.019)	(0.017)	(0.022)	(0.065)
Observations	2004	2004	1891	1891	1035
R-squared	0.12	0.13	0.08	0.09	0.08

Note: Estimates based on IHDS 2004-05. Government schools are the excluded category in all columns. Columns 1 and 2 include English medium schools as well as schools that only teach English as an additional language from primary grades. The estimates in columns 3, 4 and 5 shows the difference in English skills arising from learning English only as an additional subject in primary school – they exclude the English medium schools.