English Language Premium: Evidence from a policy experiment in India[1]

Tanika Chakraborty[2]
Washington University in St Louis

Shilpi Kapur[3]
Washington University in St Louis

September 2, 2009

Abstract

A key question facing education policymakers in many emerging economies is whether to promote the local language, as opposed to English, in elementary schools. The dilemma is particularly strong in countries that underwent rapid globalization making English a lingua franca for international as well as domestic exchange. In this paper, we estimate the English premium in globalization globalizing economy, by exploiting an exogenous language policy intervention in India. English training was revoked from the primary grades of all public schools in the state of West Bengal. In a two-way fixed effects model we combine differences across birth cohorts and districts in the exposure to English education, to estimate the effect of the language policy on wage premium. In addition, since the policy was introduced only in the state of West Bengal, we combine other states with no such intervention to address the potential threat of differential district trends confounding our two-way estimates. Our results indicate a remarkably high English skill premium in the labor market. A 1% increase in the probability of learning English raises weekly wages by 1.6%. On the average this implies a 68% reduction in wages for those that do not learn English due to the change in language policy. We provide further evidence that occupational choice played a decisive role in determining the wage gap.

JEL Classifications: H4, I2, J0, O1

[1] 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.
[2] Department of Economics, Washington University in St Louis, email: tchakrab@gmail.com
[3] Department of Economics, Washington University in St Louis, email: skapur@gmail.com

1. Introduction

There is a longstanding interest in estimating the economic returns to the human capital embodied in language skills. While previous literature focused primarily on the economic assimilation of immigrants, largely ignored was the importance of foreign language skills within domestic labor markets of many economies.[4] For instance, ever since their independence, many of the former European colonies faced the dilemma of which language to encourage in educational institutions - local or colonial?[5] The argument put forward by policymakers supporting native language training is that it fosters easier access to education, particularly for children from disadvantaged backgrounds, thus promoting greater equality over time.[6] Nevertheless, key changes in the economies of many developing countries have led policy makers to rethink the importance of teaching foreign language in schools. The debate has found renewed attention in many emerging economies which benefited from their pre-existing English language proficiency in an increasingly globalized world. 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 native language promoting policy. In this paper, we investigate whether foreign language skills are indeed rewarded in a global labor market, in turn leaving behind those with otherwise comparable levels of education and experience.

One of the major difficulties in estimating the returns to language skills, as with any other form of human capital, arises when there is correlation with unobserved individual specific ability or family background variables. Econometrically, a way to disentangle the effects of English Language skills is to find experimental set ups that generate exogenous English learning opportunities. We exploit a language policy intervention in India to address this endogeneity. Until 1983, English was taught in all primary schools starting from the first grade. Beginning in 1983, English was revoked from primary grades in all public schools in West Bengal and introduced as a part of the curriculum starting from grade 6.[7] However, cohorts who were already enrolled in school before 1983 were exempted from the policy change. Moreover, private schools were out of the purview of this policy.

[4] Few exceptions are Angrist et al (1997, 2006) and Lang & Siniver (2006)
[5] For example, French in the case of many African colonies or English in the case of many British colonies in Asia.
[6] Post independence, many former European colonies implemented programs to actively promote the national language at the expense of the colonial language in schools (Angrist & Lavy, 1997).
[7] Few other states like Karnataka and Tamil Nadu also had similar language policy changes but in much later periods.

Thus the extent to which an individual is exposed to the policy is determined both by his cohort of birth and the type of schooling - public or private. We use this exogenous shift in English learning opportunity of individuals at an early age, to estimate the English skill premium in India.[8]

Using the district and cohort variations in a two-way-fixed-effects model, we find that knowledge of English has a positive premium in the domestic labor market. Moreover, as expected, the premium is higher for individuals having primary or higher levels of schooling, since jobs requiring English skills would also require a threshold level of education. 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 data from other states that did not experience any change in language policy. Controlling for district trends strengthens our two-way estimates suggesting a significantly high English skill premium in the labor market. Our estimates suggest that a 1% increase in the probability of learning English in primary school leads to a 1.6% increase in wages. On the average this implies a 68% 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 exposure to English significantly reduces the odds of an individual working in higher ranked or better paying occupations.[9]

Angrist and Lavy, 1997, use a similar policy to estimate French skill premium following the abolition of French from Moroccan primary schools. However, the Moroccan language policy change was a country-wide phenomenon, unlike the policy change in our case. They find a positive premium associated with French writing abilities by using variations in individuals' years of schooling and cohort of birth. However, a serious disadvantage of using variations in years of schooling across individuals is the 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

[8] 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).

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

and hence bias the results. Moreover, since one of the objectives behind language transition policies is to increase the accessibility of education to children from disadvantaged backgrounds, these policies could have increased educational enrollment in Morocco[10]. If the Moroccan language policy indeed generated an 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 the language policy 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.

Although our analysis is based on the Indian experience, the 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.[11] Since its independence from British rule in 1947, the disagreement over the ideal language policy has 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 in most 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 increasing 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.[12]

The remainder 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 endogeniety 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.

---

[10] In the context of India, a recent paper by Joydeep 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.

[11] There are 22 official languages in India.

[12] 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.

2. Policy Background

Under the Constitution of independent India, education fell under the purview of the state administration. As a result in many cases, education policies have been affected by regional political ideologies. One of the major policies the state governments have experimented with is how English fits into the primary school syllabus. This in turn has generated inter-state differences in the starting grade of English education. While in some states English is taught from the first grade, others do not teach English in primary schools at all. West Bengal, housing the capital of colonial India, taught English in primary schools from the first grade. However in 1977, with the election of a communist government to power, English education was perceived to be elitist giving an unfair advantage to a small section of the society, in turn perpetuating inequality. Teaching of English at the primary school level was considered "unscientific" and with their concerted efforts to overturn the inherent biases and inequities of the colonial system of education, English was completely wiped out from public primary schools in West Bengal for almost 25 years starting from 1983[13]. Thus, children entering primary school after 1983 (i.e. under 6 years in 1983) did not learn English in primary school. Children born before 1977 were not affected by the change as they would have entered primary school before 1983. Proponents of the policy argued that abolition of a foreign language from primary school would increase enrollment and rate of school completion and hence improve the educational standard of the population thereby reducing inequality[14]. However, what the policymakers failed to acknowledge was the value of English skills in the labor market. Individuals who could afford private schooling and coaching would acquire the necessary skills to find jobs requiring English skills. This in turn would further increase the inequality.

Indeed with liberalization in many of the emerging economies, English has become an unofficial link accelerating the process of global integration. It is widely believed that the preexisting knowledge of English has helped India emerge as a commercial hub in the world market for Information Technology Enabled Services. For example, while liberalization of the services and telecom sectors to private and foreign investment has led to remarkable growth in export of services in many South Asian countries, India turned out to be the single largest destination for IT services by 2004 (Shastry, 2007). Thus investment in English skills has resurfaced as an issue of utmost

[13] The policy was scaled back 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.

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

importance within many of these countries, possibly affecting many more individuals than the world immigrant population combined. The increase in employment probability for those with English skills has resulted in an overwhelming support from the parents for English training starting from elementary schools in India. A RIESI survey in India conducted in 2003 found that more than 90% of the parents believed that learning English would help their children improve social mobility and access better job opportunities. Many people agree that service sector liberalization has led to a steep rise in white-collar wages in India benefiting only the urban-English-educated[15]. This inequality might be alleviated if 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. English skill being more remunerative, private training of English is likely to remain at a premium too. Since primary education is a public good, the onus then lies on the education policymakers to provide appropriate skills targeted towards the labor market. India's liberalization experience provides an excellent opportunity to revisit the debate on the optimal language policy in primary schools.

In this paper, we use the education policy shift in West Bengal to estimate whether English skills affected individuals' labor market outcomes, in the backdrop of India's large scale liberalization program.

3. Identification Strategy and empirical specifications

Identifying the causal effect of language ability on wages is confounded by many unobserved variables like individual ability, family background etc. 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. Moreover, since most of the jobs in the liberalized service industry require skills of spoken English, self reported data on English speaking ability would mostly be subjective.

Hence we use the exogenous policy change to proxy for English skills. Theoretically, there are two sources of variation in the exposure to English learning opportunity. First, since the policy

[15] 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.

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. Secondly, since the policy was applicable only to those children who joined the first grade after 1983 (those already in school in 1983 were unaffected), there is a variation across cohorts. We employ a two-way fixed effects error correction model to estimate English language premium in the labor market using these 2 sources of variation.

Thus students in districts where the probability of attending a public school is higher are more likely to be affected by the policy change. Moreover, within those districts, cohorts who join the first grade at, and after, the year of policy change are exposed to the policy. Cohorts who attend primary school before the change in policy serve as control groups. Thus the interaction of district level exposure to public schools with the cohort of birth provides a measure of the exposure of an individual to English education.

3.1 Intensity of Policy Exposure

The National Sample Survey (*NSS*) is the only household survey in India that collects wage and occupational information of individuals. However, it lacks any information on the type of institution attended, public or private, for individuals on whom the wage survey is based. Hence we do not observe wage and relevant education details of the same individual. To deal with this problem, we compute the probability of an individual, in a specific district, attending public school at the time of the policy change from earlier rounds of National Education Census (National Sample Survey for region level). We then combine our individual level data with these district (region) estimates to have a complete set of required information on each individual in the labor market. Since the policy implied that public schools were mandated to abolish teaching of English language 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* 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. And IP_r^E is the Public School Enrollment Measure – the percentage of students enrolled in public schools and hence affected by the policy change. However, one difficulty with this estimate is that the survey is representative 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). Hence, we use a second measure of public school exposure at the district level using data from All India Education Survey (*AIES*) data on the number of public and private schools in a district.

$$IP_d^S = G_d^S / N_d^S$$

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. And IP_r^E is the Public School Intensity Measure – percentage of public schools in a district reflecting the potential probability of a person attending a public school. Table 1C reports the average probability of attending a public school based on these two measures. The two measures are very close for all three states combined. According to the Public School Intensity measure, the average probability of being exposed to the Language Policy change in a district (*i.e.* attending a public school) is 43%. According to the Enrollment measure, at the region level, it is 45%.

We construct both of our intensity measures based on public school and 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 definite 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, which 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). In regards to 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 strategy uses the variation in treatment intensity across districts and cohorts to identify the relationship between English language skills and individual labor market outcomes[28]. The younger cohorts are the ones deprived of English training in the primary school. Moreover, the higher the probability of attending a public school suggests a 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 learning English (or the probability of attending a public school)

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

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 that the effects of the policy change are 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 with a high and low probability of learning English. We cluster the standard errors at the district level. The negative value of the coefficient a_2 can then be interpreted as the estimate of the impact of English education on wages. Thus if English skills have high returns in the labor market, we expect α_2 to be negative.

[28] 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.

One concern is that the national household sample survey 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) is approximately the same as 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 exposure to the policy change. 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 better labor market which would downward bias our estimate of α_2. 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, might observe a higher growth in private English tuition. If true, the differential growth of private coaching centers across districts will also downward bias our two-way fixed effects estimates.

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, and difference out the 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[29]. Similar to the previous strategy, we compute both measures of public school exposure for these states and estimate the following regression.

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

In this regression β_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* belongs to 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, the higher growth of export oriented jobs in the control states of Punjab and Haryana compared to West Bengal might upward bias our estimates. However, the inclusion of the state time trends differences out all such state specific time varying factors.

4. Data

We use data mainly from two sources: All India Educational Survey (*AIES*) and National Sample Survey (*NSS*) provided by the Government of India. The *AIES* is a census of schools in India and provides district level and state level data on the number of public and private schools[30]. This survey is conducted every 5-7 years and provides an important source of school level data. The two district and region level public school exposure measures (see Section 3.1) used in our identification strategy are constructed using the *AIES* (1986) and education round of *NSS* (1986).

The individual level data comes from the *NSS*'s Employment and Unemployment Survey (Schedule 10) and the Educational round of NSS (round 42nd, 1986, and round 52nd, 1996).[31] The Employment and Unemployment rounds are 5-yearly surveys and are divided into four sub-rounds

[29] Even though many other states would qualify as control group we are restricted to Punjab and Haryana because School intensity data was unavailable for other states.

[30] To construct the region level measure of public school exposure, we aggregate the district level figures. NSS gives the composition of different regions.

[31] For the sub-state level of analysis, the individual and household level data is appropriately aggregated to arrive at measures of factors that influence segregation indices.

and covers both urban[32] 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[33]. 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 cohorts entering primary school both before and after the policy change. The educational rounds are conducted once every few years and provide detailed data about the education and demographic particulars of the sampled individuals at the time of the survey.

We restrict our sample to the working individuals in the age group 17-45 at the time of the *NSS* survey in 2004.[34] 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 aged 6 years and 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. These individuals then form the treatment group in our analysis as they went to primary school after 1983 and would be potentially affected by the policy. 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). There could be some individuals, born in the fag years of the control period, who started primary school at a later age and thus may have been exposed to the policy change

[32] An urban area is defined as a town if it met the following conditions: (1) a density of not less than 1000 per square mile, (2) population of at least 5000, (3) three-fourths of the occupations of the working population should be outside of agriculture, and (4) at the discretion of the Superintendent of the State, the place should have a few pronounced urban characteristics and amenities such as newly founded industrial areas, large housing settlements, or places of tourist importance, and other civic amenities. See Bose (1973). Our data set assigns a sector value to all observations

[33] 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. In 1993-94 (Round 50 of NSS), the survey covered more than 69000 rural and 46000 urban households.

[34] The results reported are not sensitive to different birth cohorts being included or age cutoffs considered as participating in the labor market.

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 one-digit level of classification of occupations and categorize the occupations into the following six broad categories similar to Kossoudji A. S (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

Descriptive statistics are reported in Table 1A and 1B. For the treatment state, West Bengal, the average age in our sample is 30 years with an average age at entry to school of 6.36 years. The average potential experience calculated using the definition *job experience=minimum {age-15, age-age at highest education}* is 8 years and average age at highest education is 16 years. About 27% of the sample were illiterates, 13% had just primary schooling, 20 % of the sample were educated up to the middle school level, 13% had secondary education and 11.4% had higher secondary schooling. 15% of the sample were graduates and above. Average deflated weekly wages in 1982 Indian Rupees was 53.55 in our treatment sample.[35] For the group of combined treatment and control states that we use for the triple difference estimation, the average age is again 30 years, the average age at entry to school is 5.96 years, the average job experience is 7 years and the average age at highest education level is 16 years. The average real weekly wage of the treatment and control states combined is 62.33 Indian rupees. The educational distribution and the occupational distribution of the control states are very similar to that of West Bengal.

5. Results

[35] The current exchange rate between Rupee and Dollar is approximately 40 INR to 1 USD.

5.1 Basic Results

We start by estimating a basic specification where log wages for individuals in the state of West Bengal are regressed on an indicator of whether the individual went to primary school attendance in the period after the policy change, suggesting exposure to the policy change. We account for controls such as job experience, age and education. The underlying assumption is that private and public schooling are equivalent in terms of English education. This method compares our potential treatment group individuals (those who were affected by the language policy change) and potential control group (those who went to primary school before language policy change and thus were unaffected by this policy change). The results are enumerated in Table 2. The negatively significant coefficient on the Post Dummy indicates that individuals in the control group, with greater exposure to English, have a wage premium over the treatment group. However, the identification in this regression comes only from variation over time. It is possible that the lower wages for younger cohorts are a result of trends in the labor market, completely unrelated to the language policy change. If that is true then these estimates would be spurious. In order to control for time trends we combine the cohort variation with cross section variation in exposure to the language policy. Particularly, we exploit the differences in the concentration of public and private schools to introduce inter district (region) variation in the exposure to English language in West Bengal.

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 public school 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 group that each individual belongs to (Schedule Caste/Tribe and others) in all our regressions. We cluster the standard errors for any within district correlations across individuals. 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 0.2% less wage. Column 4 shows the results from our model in equation (1). The results are similar after controlling for individual birth year effects instead of a post-treatment dummy although the estimates are not precise.

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 regression. The estimates reported in Column 1 and 4 of Table 4 (with a common post-treatment trend and individual 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 0.2% 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 reform, the estimates suggest relatively lower wages for individuals who went to primary school in the post policy period in areas with higher intensity of public schools and thus higher potential exposure to the policy.

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 were unaffected by the language policy but were still affected by any other district wide changes. Table 4 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). The results in Column 3 and 6 indicate a very strong negative effect of the policy on individuals who are expected to be affected by the 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 1% reduction in wages. Table 4 shows the analogous results at the region level. The estimates are smaller than at the district level but still are much bigger than for the overall population.

If the two-way results were completely spurious we would expect similar results for individuals who did not complete even primary schooling and hence were possibly not eligible for high wage jobs irrespective of English education in primary school. The results in column 2 and 5 of Tables 3 or 4 suggest that there was no significant effect of the policy on this group. The coefficients are either very small or positive. In general the results imply a significant difference in the wage outcomes for individuals who completed more than a primary level of schooling. There was no significant difference in wage outcomes, between younger and older cohorts in high and low public school intense districts, for individuals who never went to school or had not completed primary 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 since the return to education might have declined over time due to liberalization.

5.3 Differential District Trends

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. However, 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 labor market 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 coaching centers responds more to the policy transition in districts with fewer private schools. To see if indeed there is a differential trend across districts we conduct a falsification test. Table 6 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[36]. 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 (*West Bengal * Post * Intensity measure*). 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 preexisting 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 probability of attending public schools earned relatively lower wages. The estimates suggest that a decrease in the probability of learning English by 1% lowered weekly wages, in 2004, by approximately 1.6%. The average proportion of public schools in our sample of three states is 43%. Thus on average revoking English language instruction from public primary schools lowered wages by 69%. Moreover, for individuals with at least primary schooling, a decrease in the probability of learning English by 1% lowered weekly wages, by approximately 1.8%.

[36] 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

Table 7 presents the results after controlling for region-specific time trends. The results are larger than the two-way estimates but smaller in magnitude compared to the district level regressions.

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[37].

In fact a regression of the probability of working for a wage on English skills shows a positive relationship in our sample. 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 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 accrues 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,

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

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.[38] 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 investigate which mechanism is 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 the two-way fixed effects model and with the following specification:

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

where the dependent variable measures the log odds of working in occupation category *j* relative to occupation category *r*. 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.

Table 8 reports the estimation results using the specification in model 3. The main coefficient of interest for our analysis is that on the interaction term of public school exposure measure and post policy dummy which is reported in the table. These coefficients show the odds of working in one occupation relative to another as a function of the individual's district exposure to public schools in the post policy period.

If we construct an ordinal ranking of the occupations based on the skills they require and the average wages they pay, we can rank the occupations in the descending order as follows: PROF,

[38] 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.

MNGR, CLER, OPER, SERV and CRAFT. Our coefficient of interest, which is associated with the interaction term, is δ_2. A negative (positive) value of δ_2 implies that for individuals with lower degree of policy exposure have higher (lower) probability of working in a higher ranked occupation. Table 8 reports the multinomial coefficients of the interaction, δ_2 , estimated from model (3). They indicate the log odds of working in a specified occupation versus another from the two-way fixed- effects estimation, at the district level. Column 1 reports the estimation results from the full sample of West Bengal, without separate education categories. Columns 2 and 3 report the coefficients for below primary and above primary-educated group of individuals respectively.

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. When we consider the group of individuals who are educated above the primary level, we do find a high negatively significant coefficient at the 1% level of significance on our interaction term for the comparisons of all the occupation categories versus the category of Craft occupations. This suggests that for those educated at the primary level and above, holding other things constant, greater exposure to English, raises the odds of an individual working in high ranked occupations relative to craft and kindred occupations. 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 (which is higher ranked based on the ordinal ranking we defined) compared to craft and kindred occupation category. The log odds of working in a managerial occupation compared to craft and kindred category, row-2, column-2, shows a decline of 3.4% for a 1% increase in exposure to public schools in the post-policy period.

This lower (higher) likelihood of working in an occupation relative to another as a function of higher (lower) English policy exposure and hence differential access to English language learning shows that English language acquisition is an important determinant of occupational attainment of individuals. English skill can be considered as an important element of human capital[39] and thus it increases the probability of an individual attaining an occupation that has a need for highly productive workers with a greater amount of human capital such as professional workers and managerial workers.

[39] Chiswick and Miller (1995) explain that language has all the characteristics of human capital as it is embodied in the person, is productive in the labor market and is generated as a sacrifice of time and out of pocket resources.

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

Angrist, Joshua and Victor Lavy, "The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco", Journal of labor Economics, January 1997

Angrist, J., A. Chin and R. Godoy, "Is Spanish-Only Schooling Responsible for the Puerto Rican Language Gap?" , 2006, NBER Working Paper

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

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

Betts, Julian (1998), The Impact of Educational Standards on the Level and Distribution of Earnings", American Economic Review, March 1998, 266-275

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

Carliner Geoffrey, "The language ability of US immigrants: assimilatioon and cohort effects", 1995 NBER WP

Chin, Aimee (2005), Can redistributing teachers across schools raise educational attainment? Evidence from Operation Blackboard in India. Journal of Development Economics , 78, 384-405

Chiswick BR, Miller PW (1995) The endogeneity between language and earnings: international analyses. J Labor Econ 13:246–88

Chiswick BR, PW Miller, 2002 "Immigrant earnings: Language skills, linguistic concentrations and the business cycle", Journal of Population Economics", Springer

Chiswick BR, Taengnoi S, 2007 "Occupational Choice of High Skilled Immigrants in the United States", IZA DP No. 2969

Clingingsmith, D. (2007). "Bilingualism, Language Shift and Economic Development in India, 1931-1961."

Deaton, Angus (1997), The Analysis of Household Surveys, Johns Hopkins University Press

Grin Francois (2001), "English as Economic Value: Facts and Fallacies", World Englishes, Vol. 20, No 1, pg 67-78

Dustmann, Christian and Arthur van Soest, "Language and the Earnings of Immigrants,"
Industrial and Labor Relations Review 55:3 (2002), 473-492
Heckman J. James (2007), "The economics, technology, and neuroscience of human capability formation", PNAS, Vol 104, No. 33.

Hoyt Bleakley and Aimee Chin (2003), "Language skills and earnings: evidence from childhod immigrants", WP

Holla Alaka, (2006) "Do schools discriminate? School-choice, caste, and inequality in Karnataka, India ,

Judge, G.W., Willaim E.G, R.Carter, Tsoung C.L .(1980) The Theory and Pratice of Econometrics. New York: Wiley.

Kossoudji A. Sherrie, "English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men", Journal of Labor Economics, Vol. 6, No. 2, (Apr., 1988), pp. 205-228

Lang Kevin and Erez Siniver (2006), "The return to English in a non–english speaking country: Russian immigrants and native Israelis in Israel", Working paper

Levinsohn James, "Globalization and the returns to speaking English in South Africa, 2004, WP

McManus S. Walter, Labor market costs and language disparity: An interpretation of Hispanic Earnings differences, American Economic Review, 1985

Mitra Arup & Murayama Myumi "Rural to Urban Migration: A district level Analysis for India" IDE DISCUSSION PAPER No. 137

Mora T Marie, An overview of the economics of language in the US labor market, 2003, WP

Munshi Kaivan, Mark Rosenzweig (2005), "Traditinal institutions meet the modern world: caste, gender and schooling choice in a globalizing economy", WP

P. Duraisamy, "Changes to Education in India, 1983-94: By Gender, Age Cohort and Location, Center Discussion Paper 815, July 2000

Roy, Joydeep, Princeton University, "Abolition of Teaching of English from Primary Schools in West Bengal, India: An Evaluation", September 2003

Schmidt P , Strauss P , "The Prediction of Occupation Using Multiple Logit Models", International Economic Review, Vol. 16, No. 2 , June 1975, pp 471-486

Schmidt P , Strauss P , "Estimation of Models with Jointly Dependent Qualitative Variables: A Simultaneous Logit Approach" , Econometrica, Vol. 43, No. 4 , July 1975, pp745-755

Scrase J Timothy, "The Hegemony of English in India", 2004, Conference Paper at the 15th Biennial Conference of the Asian Studies Association of Australia

Shastry G Kartini, "Human Capital Response to Globalization: Education and Information Technology in India", Harvard University November 2007

Sridhar K. Kamal, "Language in Education: Minorities and Multilingualism in India", International Review of Education, Vol. 42, No. 4, The Education of Minorities. (1996), pp. 327-347

Topalova Petia, "Trade Liberalization, Poverty, and Inequality: Evidence from Indian Districts", NBER WP # 11614

Table 1A: Descriptive Statistics for West Bengal (Treatment State) based on NSS Data

Variable	Mean

Age (years)	30.68
Age at entry at school (years)	6.36
Job Experience (years)	8.56
Social Group	
Percentage Backward	27
Percentage Non-Backward	73
Percentage Females	41.39
Weekly Wages (deflated in 1982 Rs)	7
General Education	
Percentage Primary	12.76
Percentage Middle	19.95
Percentage Secondary	13.41
Percentage High Secondary	11.03
Percentage Graduate +	15.54
Percentage Others (Illiterates, Below Primary, Literate with no formal schooling)	27.31
Religion	
Percentage Hindus	87.64
Percentage Others	12.36
Occupational Distribution	
Percentage PROF	8.42
Percentage MNGR	8.24
Percentage CLER	24.93
Percentage CRAFT	11.57
Percentage OPER	43.73
Percentage SERV	3.11

Table 1B: Descriptive Statistics for Treatment and Control States Combined, NSS Data

Variable	**Mean**

Age (years)	30.08
Age at entry at school (years)	5.96
Job Experience (years)	7 .83
Social Group	
Percentage Backward	40.21
Percentage Non-Backward	59.79
Percentage Females	40.02
Weekly Wages (deflated in 1982 Rs)	62.33
General Education	
Percentage Primary	13.34
Percentage Middle	15.09
Percentage Secondary	17.39
Percentage High Secondary	10.75
Percentage Graduate +	12.06
Percentage Others (Illiterates, Below Primary, Literate with no formal schooling)	31.37
Religion	
Percentage Hindus	70.68
Percentage Others	29.32
Occupational Distribution	
Percentage PROF	8.82
Percentage MNGR	9.10
Percentage CLER	24.01
Percentage CRAFT	10.66
Percentage OPER	43.73
Percentage SERV	3.68

Table 1C: Average Probability of attending a public school

	Percentage Public School	**Percentage enrolled in public school**
West Bengal	0.3299	0.4666
Three States Combined	0.4283	0.4493

Table 2: Basic Regression of log wages on whether individual went to primary school after policy change

Dependent Variable : Log of real wage		
	(1)	(2)

Post	**-0.1519***** (0.0383)	**-0.138***** (0.0369)
Controls	Yes	Yes
District/Region Fixed Effects	No	Yes
Observations	3401	3401
R-squared	0.3079	0.3229

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

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*** (0.14)	-0.122 (0.314)	**-1.112*** (0.535)	-0.186 (0.14)	0.0470 (0.280)	**-1.340**** (0.625)
Post	Yes	Yes	Yes			
Birth Cohort Dummies				Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls						
Experience	0.0696*** (0.0063)	0.0159* (0.0088)	0.0241 (0.0246)	0.0978*** (0.0082)	0.0293 (0.0249)	0.0648*** (0.0184)
Experience	-0.0022*** (0.0001)	0.0004 (0.0005)	-0.0015 (0.0011)	-0.00404*** (0.00041)	0.0006 (0.0015)	-0.0050*** (0.0008)
SC-ST	-0.330*** (-.0321)	-0.0669 (0.0635)	-0.174** (0.0606)	-0.303*** (0.032)	-0.0287 (0.0709)	-0.0752 (0.0524)
Constant	4.570*** (0.026)	4.218*** (0.0388)	5.200*** (0.120)	3.734*** (0.13)	3.852*** (0.299)	4.072*** (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 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	Above Primary	Below Primary	All Individuals	Above Primary	Below Primary
Public School Enrollment * Post Policy Dummy	-0.187 (0.10)	-0.448 (0.25)	0.164 (0.087)	**-0.196**** (0.055)	**-0.568**** (0.16)	0.114 (0.086)
Controls						
Experience	0.0699*** (0.0039)	0.0140 (0.0100)	0.0197** (0.0037)	0.0990*** (0.007)	0.0089 (0.0060)	0.0317** (0.0056)
Experience square	-0.0022*** (0.0002)	-0.0002 (0.0004)	-0.0004 (0.0002)	-0.0041*** (0.0003)	-0.0022*** (0.0004)	-0.001** (0.0003)
SC-ST	-0.345*** (0.043)	-0.325** (0.091)	-0.175*** (0.024)	-0.317*** (0.035)	-0.246* (0.080)	-0.170*** (0.028)
Post	Yes	Yes	Yes			
Cohort				Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	4.645*** (0.013)	5.446*** (0.078)	4.313*** (0.0027)	3.755*** (0.020)	4.039*** (0.057)	4.208*** (0.029)
Observations	2766	1523	1243	2766	1523	1243
R-squared	0.23	0.20	0.11	0.30	0.35	0.17

Clustered standard errors 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 Dummy* West Bengal Dummy	**0.458*** (0.23)	**0.631***** (0.17)	**1.382**** (0.57)	**2.135**** (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).

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

Dependent Variable : Log of real wage		
	(1)	(2)
	All Individuals	Above Primary Education
Public School Intensity *Post Policy Dummy* West Bengal Dummy	**-1.671***** (0.079)	**-1.785***** (0.053)
District*Cohort	Yes	Yes
West Bengal * 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 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 Intensity *Post Policy Dummy* West Bengal Dummy	**-0.178** (0.078)	**-0.502***** (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 in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Two-way Fixed Effect Coefficient Estimates of Logit Model 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.