Economics of Education Review 50 (2016) 1–16

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

English language premium: Evidence from a policy experiment in India☆

Tanika Chakraborty[a,b,*], Shilpi Kapur Bakshi[c]

[a] Indian Institute of Technology Kanpur, Kanpur, India
[b] IZA, Bonn, Germany
[c] The Energy and Resources Institute (TERI), India

ARTICLE INFO

Article history:
Received 2 August 2014
Revised 21 October 2015
Accepted 21 October 2015
Available online 30 October 2015

JEL classification:
H4
I2
J0
O1

Keywords:
English premium
Language
Triple difference
Education policy
Wage
Occupation

ABSTRACT

In this paper, we estimate the English premium in a globalizing economy, by exploiting an exogenous language policy intervention in India that abolished teaching of English in public primary schools. Our results indicate that a 10% lower probability of learning English in primary schools leads to a decline in weekly wages by 8%. On an average, this implies 26% lower wages for cohorts exposed to the policy change. We find supporting evidence that occupational choice played an important role in determining this wage-gap.

1. Introduction

There is a longstanding interest in estimating the economic returns to the human capital embodied in language skills. An extensive literature emphasizes the importance of language skills in the context of the economic assimilation of immigrants (Carliner, 2000; Chiswick & Miller, 2003; Dustmann & Fabbri, 2003). Relatively less research has explored the importance of foreign language skills within the domestic labor markets of economies.[1] Ever since their independence, many of the former European colonies faced the dilemma about the choice of language to be encouraged in educational

☆ We thank Sukkoo Kim, Sebastian Galiani, Charles Moul, Bruce Petersen, and Robert Pollak for their invaluable advice and support, Barry Chiswick and two anonymous referees for their 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.

* Corresponding author at: FB 626, Indian Institute of Technology Kanpur, Kanpur, UP 208016, India. Tel.: +91 9005834827.

E-mail addresses: tanika@iitk.ac.in, tanikac@gmail.com (T. Chakraborty), shilpikapur@gmail.com (S.K. Bakshi).

http://dx.doi.org/10.1016/j.econedurev.2015.10.004
0272-7757/

[1] Few exceptions are Angrist and Lavy (1997), Angrist, Chin, and Godoy (2008), Levinsohn (2007), and Azam et al. (2013) who study the returns to foreign language skills in the domestic labor market. Gao and Smyth (2011) and Godoy et al. (2007) study the returns to the knowledge of the majority language in the domestic labor market.

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. For instance, after 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). 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 favoring the teaching of English in schools is the high perceived returns to English in the labor market. Teaching only native language in schools 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. For instance, Shastry (2012) finds that regions with lower costs of acquiring English skills attracted more information technology jobs in India post liberalization. Munshi and Rosenzweig (2006) show that, in the 1990s, English premium increased by about 10% for men and 27% for women in Bombay. Higher returns to English skill is likely to foster the growth of avenues for private English training at a premium. Individuals who can afford private schooling and coaching would acquire the necessary skills to obtain jobs requiring English skills. However, poor households may not be able to respond to these changes to take advantage of the global opportunities. This in turn would exacerbate the existing inequality. India's liberalization experience thus provides an excellent opportunity to revisit the debate on the optimal language policy in primary schools. In this paper, we investigate the rewards to English language skills in the labor market in India, post its economic liberalization. The paper closest to our work is Azam, Chin, and Prakash (2013) who find significant returns to English skills, which increases with the level of education, in the Indian labor market. They use detailed information from the India Human Development Survey 2004–2005 to control for individual characteristics. However, the lack of longitudinal data or the absence of an exogenous variation in English skills makes it difficult to completely do away with concerns of unobserved heterogeneity.

We exploit a language policy intervention in the state of West Bengal in India that generates plausibly exogenous variability in English skills. In West Bengal, the medium of instruction in majority of the government run schools is Bengali and in a few cases it is Hindi. English is used as a medium of instruction only in a small fraction of private schools. However, English was taught as a subject in all government run primary schools starting from first grade. Beginning in 1983 teaching of English was revoked from primary grades (grades 1 through 4) in all government run schools and introduced as a part of the curriculum starting only from secondary school.[3] Cohorts who were already enrolled in school before 1983 were exempted from the policy change and continued to learn English, as a subject, in primary grades. Cohorts who started school after 1983 did not study English in primary grades. Moreover, private schools were out of the purview of this policy. Hence an individual's exposure to this policy change is determined both by the year of birth and by the probability of attending a public school as opposed to a private school.[4] We construct district level measures of the fraction of public school and the fraction of children enrolled in public school as proxies for the probability of attending public school. We start by combining district and cohort variations generated by this exogenous language policy intervention in a two-way fixed effects model to estimate the English skill premium in West Bengal. However, an inherent problem with this two-way-fixed-effects strategy is the possibility that districts with different intensities of public school exposure might also have other differences that vary over time. Specifically, districts with a higher intensity of public schools might have observed a different path of economic development or growth in labor markets compared to districts with a lower intensity of public schools. Indeed, Muralidharan and Kremer (2008) show that regions with higher per capita income are less likely to have private schools in India. To correct for these confounding district trends we use as controls districts from states which did not experience any change in language policy during that period. This allows us to eliminate factors that varied between high and low public-school intense districts for each cohort. The underlying identifying assumption is that the difference over time between high and low public school intensity districts in the treatment state is the same as the difference between similar districts in the control states, in the absence of the policy intervention.

Our estimates suggest that a 10% decrease in the probability of learning English in primary school led to a decline in weekly wages by 8%. On the average, this implies 26% lower wages for cohorts not exposed to the policy change. Had we not accounted for the time varying differences between the low and high intensity districts, our inference about the English premium would have been much lower. Close examination 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.

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

[2] For example, Frenchis commonly used in the labor market of many African countries and English in the case of many former British colonies in Asia.

[3] Around the same period similar changes were taking place in other parts of India. Education being a state level policy, governments in different states were experimenting with Language policies according to their ideologies. A complete documentation of these changes does not exist. Hence, in this paper, we restrict our attention to states for which we could find an official document on language policy in schools.

[4] 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).

Moroccan language policy change was a country-wide phenomenon, they use variation in individuals' years of schooling and cohort of birth. A disadvantage of using variation 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 bias the results. 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. We then account for potentially different time effects between high and low exposure districts using control states that did not experience the policy change.

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, with 22 official languages, and the large-scale economic liberalization efforts undertaken in the recent decades. 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 from their infancy, they would easily gain considerable advantage over others in the race of life" (Guha, 2011). Since India's 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 across the country. 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 the white collar jobs. The debate over promoting indigenous languages versus English in schools was further fueled in recent decades by the expansion of high-skilled jobs in the export oriented sectors. If English skills are indeed at a premium, then excluding English from public schools will reduce economic opportunities for the poor. From a public policy standpoint this would imply a rethinking of previous policies which might have lost their initial relevance in the age of globalization.[5]

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 for our estimation of English language premium. Section 4 describes the data used in the analysis. The results of the empirical estimation on wage outcomes are then discussed in Section 5. Section 6 explores the effect of the policy on occupational choice. Section 7 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 respective state administrations. As a result, in many cases, education policies in India have been influenced by regional political ideologies. One of the major policies the state governments have experimented with is the position of English language in the primary school curriculum. 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. English as medium of instruction is only practiced by a few private schools in the country. The second variant, teaching English as a subject, is more commonly observed in both private and government schools. However the grade at which English is introduced as a subject differs 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.[6]

In independent India, government run schools in West Bengal taught English as a subject in primary school, while Bengali, and in a very few cases Hindi, remained the medium of instruction for all other subjects. The debate on teaching of English in primary school curriculum was initiated in 1977 when the communist party of India, heading a coalition called the Left Front, came to power in West Bengal. For instance, a prominent Bengali intellectual, Jashodhara Bagchi, debated against teaching of English in schools in an article in the Economic and Political Weekly in 1981 (Roy, 2004). The left leaning intellectuals and policymakers held the view that English is an elitist language from the colonial era which discouraged school participation of children from disadvantaged backgrounds. They argued that abolition of English from primary school would increase enrollment and probability of school completion and hence improve the average educational standard of the population and reduce inequality.[7,8] However, while the Left Front government was in favor of reducing the prominence of English in West Bengal, the abolition of English from primary school was brought about unexpectedly in 1983 without any prior notice. Effective from 1983, teaching of English was abolished in primary grades

[5] 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.

[6] In India, primary school education typically covers grades 1–5.

[7] 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. If true, this would imply that our estimates provide a lower bound for the returns to English.

[8] Roy (2004) provides a thorough discussion of the political backdrop in which this policy was undertaken. He finds that the policy failed to achieve its desired objectives in terms of greater enrolment in school in urban regions.

of government schools.[9] Since the ban was implemented by the state government which funded and directly managed the government schools across West Bengal, all such schools, were strictly required to adhere to the new policy with immediate effect. On the other hand, private unaided schools and government aided private schools remained outside the purview of the policy since they were privately managed and hence not mandated to follow the guidelines of the government.[10]

With the new policy, English was abolished as a subject in primary school (grades one through five) and was introduced as a subject at the beginning of secondary school, grade six onwards, in 1983. Children entering government primary schools after 1983 did not learn English in primary school. Since the entry age at primary school is around 6 years, this meant that children who were born after 1977 and attended a government school did not learn English in primary grades. Children born before 1977 and thus entered primary school before 1983 were not affected by the change and continued to learn English. This implies that in the pre-policy period, a child would learn English for the entire duration of 10 years of schooling starting from age 6. In the post-policy period, children attending government schools in West Bengal learnt English only in secondary school, effectively for 5 years, starting from age 11. Our contention is that learning English in primary schools would have improved the ability to read and write English. This would be the case due to (i) a standalone effect of 5 additional years of learning English, (ii) a cumulative effect arising out of an early start age and continuing to learn English for a longer period of time and (iii) the 'critical period' hypothesis which suggests language skills are best acquired before age 12. Note that the nature of the policy allows us to compare children with potentially low English skills to children with potentially higher English skills. This is as opposed to English being not taught at all to one group of children. Hence our identification strategy would provide more conservative estimates of the English skill premium in the labor market.

One outcome of the English abolition policy could have been an outflow of students from public to private schools that continued teaching English. However, this would require an increase in capacity of the existing private schools, or opening up of new private schools, to accommodate more students. 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, who brought about the English abolition policy, would allow an unfettered expansion of private schools as it would undermine the very policy of the state government (Roy, 2004). This is unlikely to happen in the short run since private schools need clearances from the state government for any such structural changes. Indeed, Kingdon (2005) notes that private schools started expanding in India only post 1991. She shows that private school enrolment in the primary grades of West Bengal was as low as 5% in 1995–1996. Since our identification strategy rests on individuals who started school latest by 1987, these numbers are even less of a concern for us. Finally, note that if some of the more able children actually moved to private schools in response to the policy change then our intent to treat estimates would provide more conservative measures of the effect of the English abolition policy on wages. This is because, our identification strategy exploits treatment variation at the aggregated district level. If more children moved to private schools in treatment districts then they are more likely to be equipped with English skills compared to control districts. Then children in treatment districts are more likely to earn higher wages reducing our estimate of the treatment effect.

Another potential threat to our identification strategy could arise from the said objective of the policy. Particularly, cohorts entering school after English abolition could be of lower ability as they are not required to pass English as a subject anymore in primary schools. If this is true, then our estimates would be confounded as the treatment effect not only captures impact of the policy change but also lower ability of the students. Existing evidence on the effect of the policy on educational attainment is mixed. Roy (2004) finds no improvement in educational attainment as reflected in enrollment rates, age at entry to school or dropout rates as a result of the policy change for urban regions. However, he finds significant improvement in age at entry to school and reduction in dropout rates in the rural areas, even for students from poorer backgrounds. Hence we restrict our study to urban areas. We use a more direct measure of ability – grade repetition information available in the India Human Development Survey (IHDS, 2005) data – to test whether ability composition of the students changed due to the policy, using a difference-in-differences strategy. We do not find any significant impact of the policy on the ability composition of children as proxied by grade repetition.[11] Overall, compositional changes in the ability of children due to the policy change seem less of a concern for our urban sample of study.

3. Identification strategy and empirical specifications

We use the exogenous language policy shift in the state of West Bengal in India to identify the returns to English skills in the backdrop of the country's large scale economic liberalization program. Exposure to English learning opportunity varied by cohort and district of birth. Since the policy

[9] The policy was scaled back in West Bengal in 1999 when English was reintroduced from grade 3 and was then completely revoked in 2004–2005 when it began to be taught from grade 1 itself.

[10] 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 estimates would only be conservative.

[11] We also repeated our analysis in Tables 2 and 5 using the IHDS data after controlling for ability of individuals as proxied by an indicator on grade repetition. The results are similar to what we found in our main specification using the NSS data. Results are available from the authors upon request. We thank an anonymous referee for this suggestion.

was introduced in 1983, only children who went to school post 1983 were exposed to the abolition of English. 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 exposed to the policy. Our estimates would reflect English skill premium under the assumption that learning English in primary school significantly affects individual's English skills. In Table A.1 we provide some suggestive evidence on the effect of learning English in primary school on English skills of individuals using the IHDS data. IHDS collected self-reported data on individual's English ability in 2004–2005. 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 of West Bengal.[12] Since the policy was revoked in West Bengal starting from 2004, we consider only children who joined the first grade before 2004 (see footnote 12). Column 1 shows that a child is 18 percentage points more likely to speak English if he/she attends a private school as opposed to a public school. Column 2 disaggregates the school types further to see if children in private aided schools have similar English skills as those in public schools. This would be the case if the aided schools also observed the English abolition policy. The findings imply that 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 in primary grades as opposed to the effect of English being used as a medium of instruction, we exclude, in column 3, the schools with English as the medium of instruction. This was a very small fraction of private unaided 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. 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, those who would have been exposed (or not exposed) to the full effect of learning English in primary grades. These results provide some suggestive evidence on the first stage – not learning English in primary grades is associated with lower English skills of individuals.

Absence of data on individual specific English ability, around the time of the policy change, prevents us from identifying the first stage causally. Hence, in our main analysis, we restrict our attention to 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.

[12] Note however that these findings are only suggestive of the first stage. We are not able to employ the identification strategy as in our main analysis since there are not enough observations in IHDS with information on English speaking ability for cohorts that went to school around the policy years.

Our empirical analysis proceeds in two steps. First, we compare individuals across districts (or regions) and cohorts with varying degrees of policy exposure across districts (regions) in West Bengal. Next we highlight the importance of controlling for time varying effects across districts with different degrees of policy exposure. We introduce the control states of Haryana and Punjab which experienced no similar policy change during the reference period and account for these 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 the labor market outcomes of individuals 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 can be considered as a probability measure of individual i having studied in a public school in district d (or region r) in 1983. We construct two measures of public school exposure. The first measure is the probability of attending a public school using region level public school enrollment figures from the Survey on Participation in Education conducted by the National Sample Survey Organization (NSSO) between July 1986 and June 1987. This is constructed as follows,

$$IP_r^E = \frac{G_r^E}{N_r^E}$$

where G_r^E is the number of students enrolled in public schools in region r. N_r^E is the corresponding total number of students enrolled in both public and private schools. IP_r^E is the *Public School Enrollment Measure* and calculated as the proportion of students enrolled in public schools and hence affected by the policy change. One difficulty with this estimate is that the NSS is representative only at the region level, an administrative boundary larger than a district, and thus generates very little variation in IP_r^E (there are only four regions in West Bengal).

Hence we alternatively use data from the Fifth All India Education Survey (AIES)-1986 which contains information on the number of public and private schools at the district level. Using this data, we construct a second measure of public school exposure, and call this the Public School Intensity Measure,

$$IP_d^S = \frac{G_d^S}{N_d^S}$$

where G_d^S is the number of public schools in district. N_d^S is the corresponding total number of public and private schools. IP_d^S is the *Public School Intensity Measure* and is calculated as the proportion of public schools in a district reflecting the potential probability of a student attending a public school. Table 1a reports the average probability of attending a public school based on these two measures for our treatment (West Bengal) and control (Haryana and Punjab) states. For all three states combined, the average probability of being exposed to the Language Policy change, according to the Public School

Table 1a
Average probability of attending a public school.

State	(1) Proportion of public schools in the state[a]	(2) Proportion of students enrolled in public schools[b]
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 of English ban) refers to government run schools and private school (not affected by the policy of English ban) includes both government-aided privately managed schools and unaided private schools.
[a] Estimates based on Fifth All India Educational Survey (AIES)-1986.
[b] Estimates based on Survey on Participation in Education: NSS 42nd Round: July 1986–June 1987.

Intensity measure, is 55%. According to the Enrollment measure, it is about 45%.

It is important to note here that both our measures of exposure to public schools are based on data collected around the same time period (1986 for AIES data and 1986–1987 for NSS data). For AIES, 1986 was the earliest year after the policy change for which the 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 from alternative avenues of English training. For the public school enrollment measure, we use the data from *NSS* (1986–1987) 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.[13] On the one hand, higher the probability of attending a public school, lower is the probability of learning English. On the other, younger cohorts are deprived of English training in the primary school. 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. This is captured by the following equation:

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

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 entered primary school after the policy change. It takes a value 1 if an individual enters primary school in or after 1983 and 0 otherwise. Thus ($IP_d^S \times Post$) 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 years of job experience, years of job experience-squared, education, religion, caste 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 might have affected the pre- and post-policy cohorts differently. Controlling for cohort fixed effects reduces the likelihood that the effects of the policy change are confounded with other changes that occurred over time. D_d are time invariant district dummies capturing district specific characteristics which might affect individuals in the high and low public school-intense districts differently.

This two-way-fixed-effect model compares wage outcomes between cohorts entering school before and after the policy change and between districts with a high and low probability of learning English. Since wages are measured at the individual level, we cluster the standard errors at the district level. α_2 is our coefficient of interest and 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 NSS rounds from which we get the wage data do 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 he/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 three states combined that we use in our sample. In addition, Topalova (2007) notes that less 4% of the population in urban areas moved for reasons of economic consideration (or employment). Thus we assume that district of current residence (or of employment) of an individual is approximately the same as the schooling district.

3.3. Threat to DID estimates: differential 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 bias our estimate of α_2. In other words, the estimates of α_2 might be confounded in the presence of pre-existing year specific effects that vary directly with the percentage of public schools in a district.

[13] This strategy is similar to Card and Krueger (1992). More recently it has also been used by Duflo (2004) to study the impact of school expansion on education and wages.

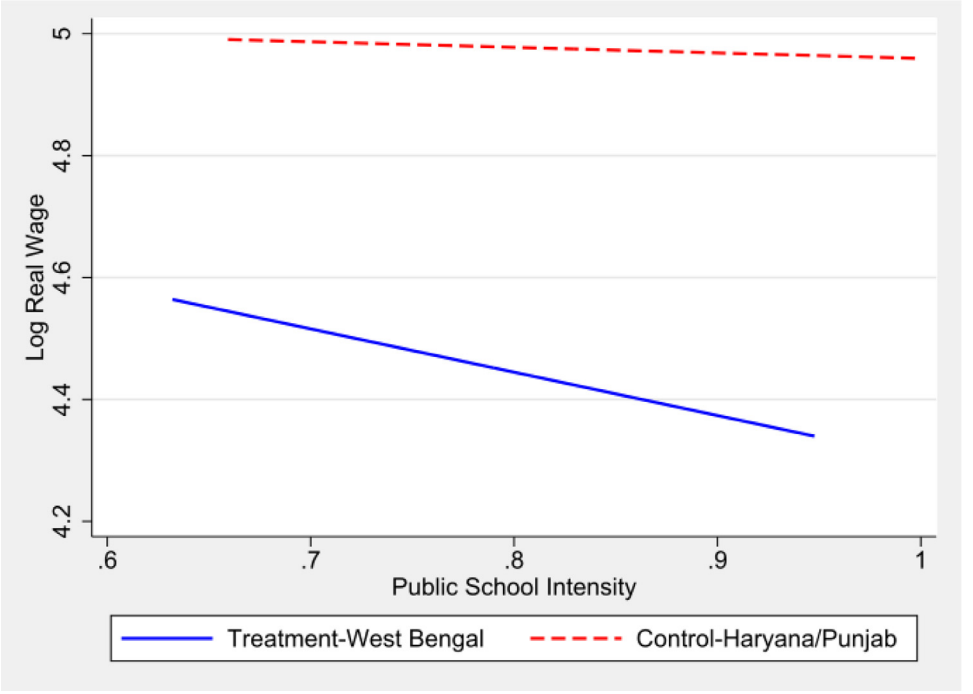

Fig. 1. Policy exposure: intensity of public schools.
Note: Estimates based on NSS rounds 1999–2000 and 2004–2005 and AIES 1986. Public School Intensity is measured for each district as the percentage of government schools in total number of schools.

3.4. Accounting for differential time trends

As mentioned earlier, education policies are governed by state authorities and the policy under review was 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 as they continued to teach English from the first grade in public schools at the time when West Bengal experienced the change in its language policy. While other states that did not experience any change in Education policy around 1983 might also qualify as control states, our choice of Punjab and Haryana are driven by two constraints. Firstly, we are restricted by the availability of AIES data, around the time of the policy change, only for these three states. Secondly, ever since independence, several states experimented with education and language policies. However, a complete documentation of these changes does not exist. Hence, in this paper, we restrict our attention to the states for which we could find an official document on language policy in schools.

Figs. 1 and 2 show the idea behind our identification strategy. Fig. 1 plots the predicted weekly wages against public school intensity separately for the treatment (West Bengal) and the control states (Haryana and Punjab). Our basic contention is that higher prevalence of public schools should affect wages only in West Bengal since English was abolished only in public schools in West Bengal and not in Haryana and Punjab. This is exactly what we find. Wages are lower in places where public school intensity in higher only in the case of West Bengal.

Fig. 2 plots the average weekly wages by cohort separately for the treatment (West Bengal) and the control states (Haryana and Punjab). The general trend shows that wages are going down over time, reflecting, in part, the negative correlation between job experience and cohort of birth. However, our interest lies in the relative difference in wages between the treatment and control states. The graph shows that for older cohorts, born before 1975, who were not exposed to the policy (cohorts below 20 in the graph) the trend in wages in the treatment state closely follows the wage trend in control states. On the other hand, for younger cohorts, born after 1977, who were affected by the policy (cohorts above 20 in the graph), the wages fall relatively more in West Bengal compared to Haryana and Punjab.

We now turn to our regression specification, similar in spirit to a triple difference strategy.

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

In this regression β_2 gives the causal estimate of the effect of language policy in West Bengal on wage outcomes. 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. The interactions $IP_d^S \times D_c$ between public school intensity and cohort dummies denote district-time trends that account for any time varying differences 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. Thus we include the time

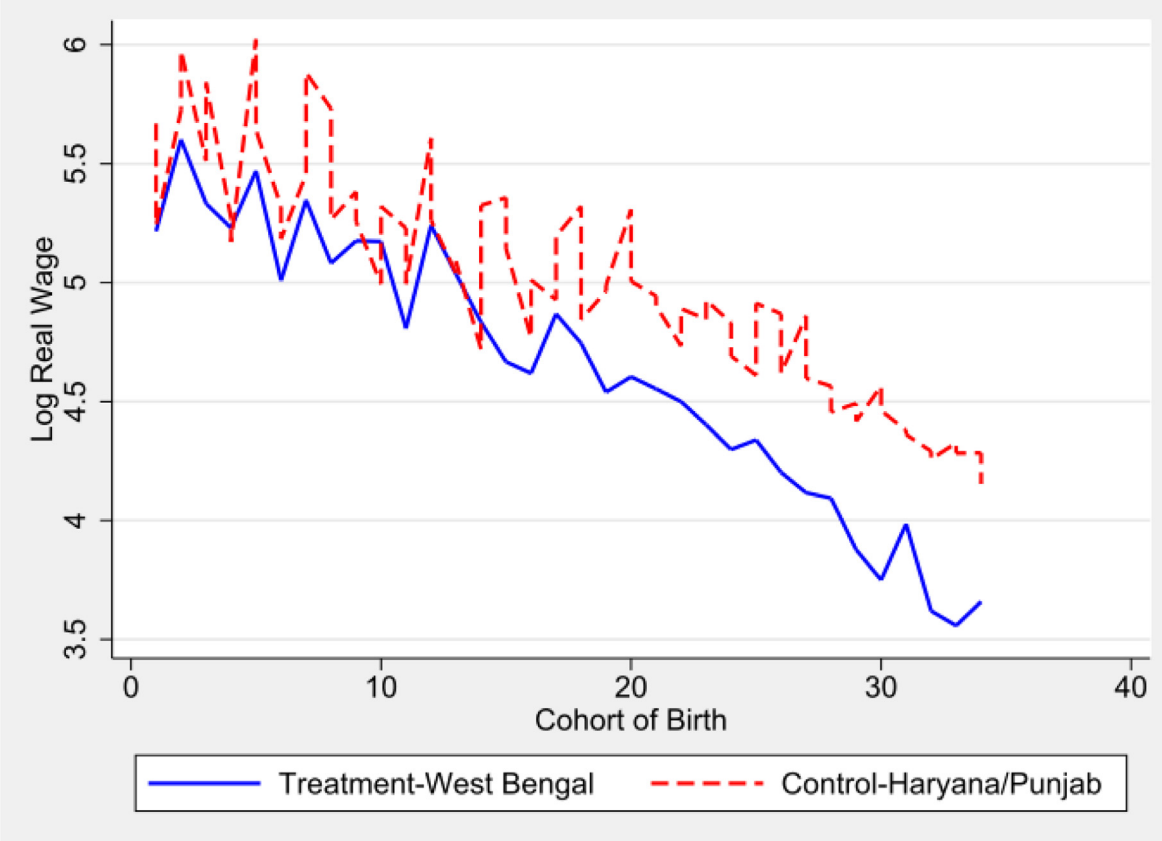

Fig. 2. Policy exposure: cohort of birth.
Note: Estimates based on NSS rounds 1999–2000 and 2004–2005. 1–20 reflect the older cohorts with year of birth 1954–1975. Cohorts 21–32 are the ones born between 1976 and 1987.

varying state effects, WB $\times$ D_c, that differences out all state specific time varying factors. This identification strategy allows districts with different public school intensity to have different cohort specific effects. However, it rests on the assumption that the difference in trends between the high and low public school intensity districts is identical across West Bengal, Haryana and Punjab. This assumption is likely to hold once the state specific factors (WB) have been controlled for.

4. Data

The data used in this paper comes from two sources: The All India Educational Survey (AIES) and the National Sample Survey (NSS) conducted by the Government of India. The AIES, conducted every 5–7 years, is a census of schools in India and provides district level information on the number of public and private schools across districts in India. The information is collected and disseminated separately for each state. The district level Public School Intensity Measure is constructed using the Fifth AIES-1986 round. For the region level Public School Enrollment Measure we use the Survey on Participation in Education, NSS 42nd Round, July 1986–June 1987 (see Section 3.1 for more details).

The individual level data to study the impact on wages and occupational choices are taken from the Employment and Unemployment Survey conducted by the NSSO between July 1999 and June 2000 and Employment and Unemployment Survey conducted between July 2004 and June 2005. The Employment and Unemployment rounds are 5-yearly surveys. The survey includes information on household characteristics like household size, social group and religion. It also includes detailed demographic information including age, sex, location, educational level, school attendance, occupational status and industry of occupation for those employed. For this paper, we pool the data from the two rounds (NSS 55th and 61st Round) since these are the only two rounds that allow us to observe labor market outcomes for cohorts entering primary school before and after the policy change.

We restrict our sample to the salaried individuals in the working age group 17–45 at the time of the NSS 55th Round (1999) and NSS 61st Round (2004–2005). Individuals who were below 17 years at the time of the survey would not be in the formal labor market that requires any knowledge of English and hence we exclude them from our study sample. 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 policy change should be felt only by those individuals

who were born in and after 1977 and hence aged 6 years or below in 1983. Individuals who are born after 1977 would be 17–22 years in 1999 (NSS 55th Round) and would be 17–27 years in 2004–2005 (NSS 61st Round). These individuals who potentially joined school in the post policy period form the treatment group in our analysis. We compare our treatment group to individuals in the age group 23–40 in NSS 55th Round (1999–2000) and those in the 28–45 age group in the NSS 61st Round (2004–2005). 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 the wage data from the two rounds. 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 six broad categories following the approach used by 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

Descriptive statistics on the social and demographic characteristics of the individuals in our sample are reported in Table 1b. For the treatment state of West Bengal, the average age in our sample is about 31 years with an average age at entry to school a little over 6 years. For the control states of Haryana and Punjab, the average age in the sample is about 30 years and the average age at entry to school is little less than 6 years. Average years of job experience is 8.5 years in West Bengal, and about 8 years in the states of Punjab and Haryana.[14] The percentage of individuals who did not compete primary education form about 26% of the sample in West Bengal and about 24% in the states of Haryana and Punjab. The distributions of the sample over the different categories of education and occupation have minor differences across the states. The average weekly wages in 1982 Indian Rupees was 71 in West Bengal compared to 91 in Haryana and 88 in Punjab.[15] Given the observed differences between these states it is possible that there exist differences on unobserved dimensions. To account for these differences our triple difference specification includes state and state-cohort specific effects.

[14] Years of job experience is calculated as *job experience=minimum {age-15, age-age at highest education}*.

[15] The current exchange rate between Rupee and Dollar is approximately 60 INR to 1 USD.

5. Results

5.1. Average impact using English learning probability

Results from the estimation of model (1) are reported in Tables 2 and 3. Table 2 uses the district level Public School Intensity Measure while Table 3 uses the region level Public School Enrollment Measure. Since older individuals would have been in the labor market for a longer time and hence earn higher income than the younger cohorts by virtue of their years of job experience, each column controls for years of job experience and a quadratic in the same. We also include dummies for different social and religious groups, in all our regressions. Since job experience is calculated from education we cannot separately control for education. Instead, we control for an indicator whether a person completed primary schooling. We cluster the standard errors for any within district correlations. Column 1 shows the results from our model in Eq. (1). While the direction of the estimates suggest that individuals who were exposed to the English language abolition policy get lower wages compared to individuals who continued to learn English in primary school, the coefficients are not significant.

Since the NSS data from which we construct the Public School Enrollment Measure is representative only at the region level, the estimation using this measure is conducted at the region level. Region being 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 exposure to public schools, we find similar results as in the case of our district regressions. The estimates reported in Column 1 of Table 3 suggest a negative impact of exposure to the language policy on wages. The coefficient is now marginally significant.[16] Overall, both at the district and the region level with different measures of the exposure to the English language policy, the estimates suggest modest or no significant effect of the English abolition policy on wages.

5.2. Heterogeneity of impact

One problem with the two-way fixed effects analysis is that younger cohorts in districts with lower public school concentration (or higher private school concentration) could be earning a higher return to human capital due to higher labor market growth in these districts. This would imply that 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 that require some knowledge of English. Specifically, the results should not hold for those individuals who would potentially be unaffected by the language policy

[16] Since there are very few clusters at the region level, inference based on clustered standard errors can be misleading. Hence, we conduct two sensitivity checks. First, we bootstrap the standard errors and second, we run the fixed effect model without clustering the standard errors. In both cases, the coefficient in column 1 becomes insignificant. However, the coefficient in column 3 for the 'above primary' subsamples remains significant at 95% confidence level.

Table 1b
State level descriptive statistics.

	(1) West Bengal	(2) Haryana	(3) Punjab
Age in years	30.69 (8.30)	29.68 (8.28)	29.87 (8.38)
Age at entry at school in years	6.36 (3.09)	5.37 (3.63)	5.94 (3.89)
Job experience in years	8.56 (9.12)	8.38 (8.88)	8.18 (9.07)
Weekly wages (deflated in 1982 INR)	71.04 (148.69)	90.60 (165.38)	87.94 (165.15)
Education			
Proportion primary	0.1231 (0.0376)	0.1379 (0.0512)	0.1171 (0.0224)
Proportion middle	0.2164 (0.0564)	0.1243 (0.0446)	0.1072 (0.0184)
Proportion secondary	0.1361 (0.0213)	0.1956 (0.0431)	0.2075 (0.0323)
Proportion high secondary	0.1106 (0.0355)	0.1381 (0.0638)	0.1457 (0.0291)
Proportion graduate and above	0.1508 (0.0486)	0.1719 (0.0647)	0.1727 (0.0582)
Proportion others (illiterates, below primary, literate with no formal schooling)	0.2628 (0.0657)	0.0.2322 (0.0880)	0.2496 (0.0776)
Religion			
Proportion Hindus	0.883 (0.114)	0.954 (0.250)	0.450 (0.497)
Social group			
Proportion disadvantaged	0.316 (0.129)	0.420 (0.150)	0.465 (0.089)
Proportion females	0.413 (0.047)	0.414 (0.054)	0.405 (0.037)
Occupational distribution			
Proportion PROF	0.1016 (0.0598)	0.1005 (0.0522)	0.0980 (0.0330)
Proportion MNGR	0.0632 (0.0580)	0.0627 (0.0332)	0.1199 (0.0385)
Proportion CLER	0.2607 (0.0478)	0.02598 (0.0913)	0.2325 (0.0498)
Proportion CRAFT	0.1134 (0.0483)	0.1014 (0.0568)	0.0956 (0.0349)
Proportion OPER	0.4195 (0.0859)	0.4392 (0.1073)	0.03938 (0.0753)
Proportion SERV	0.0414 (0.0342)	0.0361 (0.0330)	0.0601 (0.0617)

The statistics are based on Employment and Unemployment Survey: NSS 55th Round: July 1999–June 2000 1999 and Employment and Unemployment: NSS 61st. Round: July 2004–June 2005; Standard deviation reported in parenthesis.

Table 2
Two-way fixed effect with public school intensity measure (West Bengal): district level.

Dependent variable: log of real wage			
	All individuals (1)	Below primary education (2)	Above primary education (3)
Public School	−0.095	0.0604	−1.358**
Intensity × Post Policy Dummy	(0.184)	(0.274)	(0.638)
Controls			
Experience	0.049***	0.0291	0.0650***
	(0.007	(0.0257)	(0.0186)
Experience Square	−0.0019***	0.0005	−0.0050***
	(0.0003)	(0.0016)	(0.0008)
Above Primary Education	0.7124*** (0.0414)		
Caste Category	−0.179***	0.0022	−0.0827
	(0.034)	(0.0543)	(0.0546)
Religion Hindu	−0.0619 (0.045)	−0.1445 (0913)	0.1278 (0.1120)
Cohort Fixed Effects	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes
Constant	3.728***	3.975***	3.992***
	(0.1580)	(0.354)	(0.242)
Observations	2766	1243	1523
R-squared	0.388	0.2361	0.3724

Clustered standard errors at district level in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.
Results based on NSS rounds 1999–2000 and 2004–2005 and AIES 1986 for the state of West Bengal. Public School Intensity is measured for each district. Post Policy Dummy indicates whether an individual started primary school after 1983. Column 1 includes all individuals with wage information. Column 2 restricts to the subsample of individuals who did not complete primary school. Column 3 restricts to the subsample of individuals who completed primary school.

but would still be affected by any other district wide changes. Columns 2–3 in Table 2 show the estimates separately for those with a maximum of primary schooling and those with more than primary schooling respectively at the district level. The results in Column 3 indicate a very strong negative effect of the policy on individuals who completed primary schooling, and hence more likely to join jobs that require English skills. In this case, a 10% decrease in the probability of learning English (or a 10% reduction in exposure to English) in primary school leads to approximately a 7.4% reduction in wages. Table 3 shows the analogous results at the region level. The estimates are smaller than at the district level implying a 2.6% reduction in wages for a 10% reduction in exposure to primary school. The results in column 2 of

Table 3
Two-way fixed effects with public school enrollment measure (West Bengal): region level.

Dependent variable: log of real wage			
	All individuals (1)	Below primary education (2)	Above primary education (3)
Public School Enrollment ∗ Post Policy Dummy	−0.157* (0.055)	0.102 (0.094)	−0.567** (0.16)
Controls			
Experience	0.0514*** (0.0055)	0.0327** (0.0057)	0.0089 (0.0061)
Experience Square	−0.0020*** (0.0003)	−0.0011** (0.0003)	−0.0021*** (0.0004)
Above Primary Education	0.7089*** (0.0555)	–	–
Caste Category	−0.1846*** (0.0304)	−0.1302** (0.0286)	−0.245* (0.086)
Religion Hindu	0.0706* (0.0238)	−0.1360*** (0.0139)	−0.0107 (0.0652)
Cohort	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes
Constant	3.8045*** (0.1093)	3.8363*** (0.1356)	4.281*** (0.1852)
Observations	2766	1243	1523
R-squared	0.37	0.17	0.32

Clustered standard errors at region level in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Results based on NSS rounds 1986, 1999–2000 and 2004–2005 for the state of West Bengal. Public School Intensity is measured for each administrative region. Post Policy Dummy indicates whether an individual started primary school after 1983. Column 1 includes all individuals with wage information. Column 2 restricts to the subsample of individuals who did not complete primary school. Column 3 restricts to the subsample of individuals who completed primary school.

Tables 2 or 3 suggest that individuals who did not complete primary schooling were not affected by the language policy change. 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 exposed to the English abolition policy. These findings are similar to Angrist and Lavy (1997). They find no wage premium of French skills in Morocco for education less than primary school level but significant premium for individuals with secondary schooling. Note that these results are also compatible with an alternate line of interpretation. It is possible that English writing and speaking skills are obtained as a cumulative effect of learning English for a longer period of time. Hence only children who learnt English in primary as well as secondary school are the ones benefiting. Hence while these results are interesting and provide suggestive evidence of the negative impact of the English abolition policy, they are not definitive evidence.

5.3. Falsification: pre-existing trends

While estimates from the two-way-fixed-effects model suggests that revoking English from primary school only moderately reduced wage outcomes of individuals exposed to the policy, it 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 funds for education. Hence districts with higher public school concentration might have experienced a higher economic and social development. In the absence of the language policy this would imply higher wages for individuals in districts with more public schools which will underestimate the policy effects. If true, then our inference of a modest policy effect based on the estimates in column 1 of Tables 2 and 3 would be wrong. To see if indeed there is a differential trend across the treatment and control districts we conduct a falsification test. Table 4 reports the results of the control experiment using cohorts, none of whom were affected by the policy change. Individuals born between 1950 and 1974 entered school prior to the start of the language policy and hence were never treated by the abolition of English from public primary schools. We consider 1965 as the pseudo treatment year and compare individuals born during 1950–1964 (control) to individuals born during 1965–1974 (placebo) in an analysis similar to Eq. (1). The results in columns 1 and 2 suggest spurious positive treatment effects. The positive significant coefficient 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 pre-existing time trends that might be downward biasing the two-way fixed effect estimates.[17]

5.4. Correcting for district-cohort trends

To correct for these confounding effects we compare our two-way fixed effects estimates in West Bengal to similar estimates that we derive from the control states of Punjab and Haryana which did not experience any change in their education policies. The estimates of model 2 are reported in Table 5 (district level) and 6 (region level). As before all

[17] Throughout the paper, by downward bias we mean that the magnitude of the coefficient is biased towards zero – a more conservative estimate.

Table 4
Falsification: pseudo treatment year.

Dependent variable: log of real wage		
	(1) All individuals	(2) Individuals above primary education
Public School Intensity × Post	0.488*	0.611***
	(0.26)	(0.17)
Controls[a]	Yes	Yes
Cohort Fixed Effects	Yes	Yes
District Fixed Effects	Yes	Yes
Observations	2670	1606
R-squared	0.47	0.32

Note: Clustered standard errors at district level in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Results from control experiments using cohorts who were never affected by the language policy change – born between 1950 and 1974. Results based on NSS rounds 1999–2000 and 2004–2005 and AIES round 1986 for the state of West Bengal. Public School Intensity is measured for each district. Post Policy Dummy indicates whether an individual started primary school after the pseudo policy year 1965. Column 1 includes all individuals with wage information. Column 2 restricts to the subsample of individuals who completed primary school.

[a] Includes control for individual's years of job experience, square of years of job experience, caste category, religion, and indicator for primary school completion.

Table 5
District specific trends: district level (Punjab, Haryana and West Bengal).

Dependent variable: log of real wage		
	(1) All individuals	(2) Above primary education
Public School Intensity × Post× West Bengal	−1.462***	−1.887***
	(0.104)	(0.0819)
West Bengal× Cohort	Yes	Yes
Public School Intensity × Cohort	Yes	Yes
District Fixed Effects	Yes	Yes
Cohort Dummies	Yes	Yes
Controls[a]	Yes	Yes
Observations	5000	2023
R-squared	0.485	0.567

Clustered standard errors at district level in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Results based on NSS rounds 1999–2000 and 2004–2005 and AIES 1986. West Bengal forms the treatment state and Punjab and Haryana form the control states. Public School Intensity is measured for each district. Post Policy Dummy indicates whether an individual started primary school after 1983. Column 1 includes all individuals with wage information. Column 2 restricts to the subsample of individuals who completed primary school.

[a] Includes control for individual's years of job experience, square of years of job experience, caste category, religion, and indicator for primary school completion.

regressions include controls for years of job experience, a quadratic in years of job experience, and the social group of the individuals. The coefficient of interest in these specifications is that on the triple interaction term ($IP_d^S \times Post \times$ WB).

The results indicate a significant negative impact of the Language Policy on future returns in the labor market. 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 negative coefficient estimate of 1.462 in column 1 of Table 5 suggests that a decrease in the probability of learning English by 10% lowered weekly wages, in 2004, by approximately 8% 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 an average, abolishing English from public primary schools lowered wages by approximately 26% for cohorts exposed to the policy change.[18] For individuals with at least primary education, the language premium is estimated to be 33%.

Table 6 presents the results with enrollment measures after controlling for region-specific time trends. The results are similar. A decrease in public school enrolment by 10% lowered weekly wages, in 2004, by approximately 8% for cohorts born in West Bengal in the post policy period.

The above results point to the importance of accounting for pre-existing differences in trends between the potential treatment and control districts. Falsification tests using cohorts born much before the policy change suggest that districts with high share of public schools actually experienced higher wage growth than districts with low share of public schools in the absence of the policy. Without correcting for

[18] 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.

Table 6
Region specific trends: region level (Punjab, Haryana and West Bengal).

Dependent variable: log of real wage		
	(1) All individuals	(2) Above primary education
Public School Enrollment × Post Policy × West Bengal	−0.290***	−0.5112***
	(0.073)	(0.090)
Public School Enrollment × Cohort	Yes	Yes
Region Fixed Effects	Yes	Yes
Cohort Fixed Effects	Yes	Yes
Controls[a]	Yes	Yes
Observations	5000	2832
R-squared	0.385	0.346

Clustered standard errors at region level in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Results based on NSS rounds 1986, 1999–2000 and 2004–2005. West Bengal forms the treatment state and Punjab and Haryana form the control states. Public School Intensity is measured for each administrative region. Post Policy Dummy indicates whether an individual started primary school after 1983. Column 1 includes all individuals with wage information. Column 2 restricts to the subsample of individuals who completed primary school.
[a] Includes control for individual's years of job experience, square of years of job experience, caste category, religion, and indicator for primary school completion.

Table 7
Falsification: pseudo treatment states.

Dependent variable: log of real wage			
	Madhya Pradesh	Orissa	Rajasthan
Public School	−0.297	2.355*	0.453
Intensity × Post Policy Dummy	(0.592)	(0.667)	(0.743)
Controls[a]	Yes	Yes	Yes
Cohort Fixed Effects	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes
Constant	3.961***	1.437	3.623***
	(0.451)	(0.603)	(0.423)
Observations	2,322	962	1,581
R-squared	0.400	0.451	0.357

Note: Clustered standard errors at district level in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Results based on NSS rounds 1986, 1999–2000 and 2004–2005. In column 1, 2 and 3 Madhya Pradesh, Orissa and Rajasthan form the pseudo-treatment states, respectively. In all columns, Punjab and Haryana form the control states. Public School Intensity is measured for each administrative region. Post Policy Dummy indicates whether an individual started primary school after 1983. Each column consists of all individuals with wage information.
[a] Includes control for individual's years of job experience, square of years of job experience, caste category, religion, and indicator for primary school completion.

these pre-existing time-varying differences, the estimates of English premium are significantly downward biased.

5.5. Falsification: pseudo-treatment states

Next we conduct some more sensitivity tests using pseudo-treatment states which did not experience a similar policy change around the same time as West Bengal.[19] We claim that the English abolition policy lowered the wages in West Bengal (treatment state) compared to Punjab and Haryana (control states). Then in the absence of an English abolition policy, the pseudo-treatment states should not have a similar treatment effect compared to the control states. Note, however, that we cannot rule out the presence of other changes in these pseudo-treatment states around the same time.

We use West Bengal's neighboring state Orissa and the states of Madhya Pradesh and Rajasthan as the pseudo treatment states since these states did not have a similar change in English teaching policy around 1983.[20] Note that we are only able to repeat this analysis at the region level since information on the public school intensity measure at the district level is not available for these States. The difference-in-differences estimates, similar to Table 3, are reported in Table 7. Overall, we find that in the pseudo treatment states, the wages were either insignificantly different or higher from the control states of Haryana and Punjab. This lends support to the claim that the lower wages in West Bengal, compared

[19] We thank an anonymous referee for this suggestion.

[20] We could not include Bihar, the other neighboring state of West Bengal, as it had a similar policy change approximately around the same time, though we could not find any documentation about the exact timing of the change in policy in Bihar.

to Punjab and Haryana, is driven by the English abolition policy in the former.[21]

5.6. 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 learning 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 biasing our estimates downwards.

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. 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 Eq. (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. Underlying mechanism: occupational attainment

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 economic liberalization in the country are both more likely to hire English skilled workers and also are the sectors that offer relatively higher wages. Thus wage premium to English skills could be the result of an inequality in the choice of occupations available to English-skilled and unskilled workers. In addition, lack of English skills may create search costs which may then change the order of occupational preferences or access to certain occupations. 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.

Table 8
Two-way fixed effect estimates of occupational choice.

Dependent variable: log odds of working in a specified occupation		
	(1) All individuals	(2) Above primary education
(PROF/CRAFT)	−0.7912	−4.253***
	(1.063)	(0.818)
(MNGR/CRAFT)	0.057	−3.457***
	(1.037)	(1.054)
(CLER/CRAFT)	−0.071	−3.899***
	(0.871)	(0.558)
(OPER/CRAFT)	1.176	−3.527***
	(0.961)	(0.693)
(SERV/CRAFT)	0.055	−4.276***
	(0.845)	(0.657)
(PROF/SERV)	−0.846	0.0233
	(0.773)	(0.97)
(MNGR/SERV)	0.002	0.8194
	(0.72)	(1.03)
(CLER/SERV)	−0.125	0.3775
	(0.302)	(0.551)
(OPER/SERV)	1.121	0.7488
	(0.600)	(0.663)
(PROF/OPER)	−1.967**	−0.7256
	(0.943)	(0.846)
(MNGR/OPER)	−1.119*	0.0705
	(0.596)	(0.940)
(CLER/OPER)	−1.247*	−0.3714
	(0.639)	(0.371)
(PROF/CLER)	−0.720	−0.354
	(0.790)	(0.760)
(MNGR/CLER)	0.128	0.4419
	(0.709	(1.099)
(PROF/MNGR)	−0.848	−0.7961
	(0.969)	(1.253)
Cohort Fixed Effects	Yes	Yes
District Fixed Effects	Yes	Yes
Controls[a]	Yes	Yes
Observations	3872	2186
Pseudo R^2	0.0878	0.0989

Note: Clustered standard errors at district level in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.
Results based on NSS rounds 1999–2000 and 2004–2005 and AIES 1986. Each coefficient reflects the log-odds ratio of working in a specified occupation relative to another and is associated with the Interaction of Public School Intensity and Post Policy Indicator, from a multinomial logit model.
[a] Includes control for individual's years of job experience, square of years of job experience, caste category, religion, and indicator for primary school completion.

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 employed in one of the occupational groups relative to another group.

The empirical specification involves a variant of the model in Eq. (1):

$$Log\left(\frac{P_j}{P_r}\right)_{icd} = \delta_1 + \delta_2 IP_d^S \times Post + D_c + D_d + \delta_3 X_i + e_{icd} \quad (3)$$

[21] As a sensitivity check, we also repeated our main regressions, with West Bengal as the treatment state and a combination of Orissa, Madhya Pradesh and Rajasthan as control states, at the region level. These results are similar in spirit to our main findings although the magnitude of the effect varies across specifications. We are unable to conduct a district level analysis as district level data are not available for these states.

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 primary 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, most of the coefficients are negative with a few 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 higher 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 4.25% in the log odds of working in a professional occupation compared to craft and kindred occupation category.

This greater (lesser) likelihood of working in a higher ranked occupation as a function of greater (lesser) exposure to English learning 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, at least in part, driven by the lack of occupational mobility for individuals with little or no English skills but otherwise similar educational attainment.

7. Conclusion

English is an increasingly valued skill in the labor market in this era of globalization. In this paper we estimated the returns to English skills in the Indian economy by using an exogenous change in English learning opportunity in primary school. The results suggest that individuals who are more likely to have training in English earn significantly higher relative wages and have better occupational outcomes. 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 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. 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 (2004) investigates the same policy and does not find many beneficial effects on school enrollment, age at entry to school and dropout rates in the urban region and the together with the results of this paper, it suggests that such policies might be regressive and actually increase inequality. Additionally, it might have important implications for gender gaps in labor force participation. Women 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 women proportionately more than men. This might reduce the male–female gap in labor force participation or wages. As a part of future research, it would be interesting to measure whether labor market outcomes were affected disproportionately for women due to the English abolition policy in West Bengal.

Appendix A

Table A.1
School type and English skills.

Dependent variable: English speaking ability					
	(1) All schools	(2) All schools	(3) Non-English medium schools	(4) Non-English medium schools	(5) 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 2004–2005. 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 show the difference in English skills arising from learning English only as an additional subject in primary school – they exclude the English medium schools.

References

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

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, 15*(1), 548–576.

Azam, M., Chin, A., & Prakash, N. (2013). The returns to English-language skills in India. *Economic Development and Cultural Change, 61*(2), 335–367.

Card, D., & Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of Political Economy, 100*, l–40 February.

Carliner, G. (2000). The language ability of US immigrants: Assimilation and cohort effects. *International Migration Review, 34*(1), 158–182.

Chiswick, B. R., & Miller, P. W. (2003). The complementarity of language and other human capital: Immigrant earnings in Canada. *Economics of Education review, 22*(5), 469–480.

Desai, S., Vanneman, R., & National Council of Applied Economic Research, New Delhi. India Human Development Survey (IHDS) (2005). *ICPSR22626-v8*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor] 2010-06-29 http://doi.org/10.3886/ICPSR22626.v8.

Duflo, E. (2004). The medium run effects of educational expansion: Evidence from a large school construction program in Indonesia. *Journal of Development Economics, 74*, 163–197.

Dustmann, F., & Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the UK. *Economic Journal, 113*(489), 695–717.

Gao, W., & Smyth, R. (2011). Economic returns to speaking 'standard Mandarin' among migrants in China's urban labour market. *Economics of Education Review, 30*(2), 342–352.

Godoy, R., Reyes-García, V., Seyfried, C., Huanca, T., Leonard, W. R., McDade, T., et al. (2007). Language skills and earnings: Evidence from a pre-industrial economy in the Bolivian Amazon. *Economics of Education Review, 26*(3), 349–360.

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

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement, 5*(4), 475–492.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*(1), 153–162.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences of the United States of America, 104*(33), 13250–13255.

Kingdon, G. (2005). Private and public schooling: The Indian experience. Available online at: http://www.ksg.harvard.edu/pepg/PDF/events/MPSPE/PEPG-05-15geeta.pdf (accessed 01.05.15).

Kingdon, G. (2008). School-sector effects on student achievement in India. In R. Chakrabarti, & P. Peterson (Eds.), *School choice international: Exploring public-private partnerships* (pp. 111–140). Cambridge, Massachusetts: 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, 6*(2), 205–228.

Levinsohn, J. (2007). Globalization and the returns to speaking English in South Africa. In A. Harrison (Ed.), Globalization and *poverty*. Chicago, USA: University of Chicago Press.

Munshi, K., & Rosenzweig, M. (2006). Traditional institutions meet the modern world: Caste, gender and schooling choice in a globalizing economy. *American Economic Review, 96*(4), 1225–1252.

Muralidharan, K., & Kremer, M. (2008). Public and private schools in rural India. In R. Chakrabarti, & P. Peterson (Eds.), *School choice international: Exploring public-private partnerships*. Cambridge, Massachusetts: MIT Press.

NASSCOM (2004). *Strategic review* (pp. 185–194). New Delhi: National Association of Software and Service Companies.

Roy, J. (2004). *Can changing academic standards affect educational outcomes? Evidence from a policy experiment in India* Unpublished manuscript, Georgetown University, Washington D.C., USA.

Shastry, G. K. (2012). Human capital response to globalization: Education and information technology in India. *Journal of Human Resources, 47*, 287–330.

Topalova, P. (2007). Trade liberalization, poverty, and inequality: Evidence from Indian districts. In Harrison (Ed.), *Globalization and poverty*. University of Chicago Press.