Revisiting the Language Experiment in India: Level Playing or Exacerbating Disparity?[1]

Bharti Nandwani[2]

Gitanjali Sen[3]

Abstract

Using a language policy and a consequent policy reversal as a source of exogenous variation, we study the impact of the removal of English language from primary grades of government schools on private investments in education in the Indian state of West Bengal. We show that the affected cohort has a four-percentage point higher likelihood of being enrolled in private school, increasing to a seven-percentage point for the richest. The economically better-off households among government school students spend more on private tuition. While the intention of English removal was to encourage higher participation of the disadvantaged, an unintended consequence of the policy has been the movement away from almost free public schools by the relatively well-off. The rich buy English education through higher (private) investment, while, the poor lose out on the skill premium, resulting in the further deepening of the divide.

Keywords: Language Policy, Bilingual Education, Private Investment, Private Tuition, English Language, India.

---

[1] We acknowledge the help of Aakansha Jain, whose assistance with the literature search has been extremely helpful. We would like to thank the anonymous referees on an earlier version, Aprajit Mahajan, for his generous comments, and the conference participants at the NEUDC 22, particularly Christopher Neilsen and Allan Hsiao, for their helpful comments.

[2] Indira Gandhi Institute of Development Research, Mumbai, India. Email: bharti@igidr.ac.in

[3] Corresponding author. Department of Economics, Shiv Nadar University, District-Gautam Buddh Nagar, State- Uttar Pradesh, Pin 201314 India, Email: gitanjali.sen@snu.edu.in / gitan_dutta@hotmail.com. Phone: +91 1207170100, Ext: 634.

JEL Codes: I20, I24, I28, I29

1. Introduction

"….children pick up languages extremely quickly between the ages of 2 and 8 and that multilingualism has great cognitive benefits to young students, children will be exposed to different languages early on (but with a particular emphasis on the mother tongue),…..There will be a major effort from both the Central and State governments to invest in large numbers of language teachers in all regional languages…., to satisfy the three-language formula in their respective States, and also to encourage the study of Indian languages across the country". (Excerpts from The National Education Policy, Government of India, 2020).

The economic returns to local language skills among the immigrant population (Kossoudji, 1988) or that of foreign language skills in the domestic labor market emphasizes the importance of bilingual education or inclusion of non-native languages in school curricula across the world (Angrist and Lavy, 1997, 2006; Barreto, 1998; Lang and Siniver, 2006; Azam et al., 2010). This is applicable particularly in countries that have been former European or American colonies where top government jobs or businesses require knowledge of the colonial language. Although bilingual education has been a norm in most of the countries with colonial history (Cafferty and Rivera-Martínez, 1981), increasing attempts to transition to the instruction in "local" language has garnered support due to its appeal in creating larger access. Instruction in "local" language is expected to remove social barriers for disadvantaged sections, who may otherwise find it difficult to cope with instructions in a non-native language (Angrist et al., 2008). It is also believed that children are able to learn more effectively if taught in their native languages at least in the initial few years. Therefore, countries like India, Malaysia, Pakistan (former English colonies), and Morocco (a former French colony) have continued the instructions in their respective "local" languages (same as native languages in these contexts) in public schools.

However, in India, even with the native language being used as a language of instruction in public schools in most of the states, the ideal timing of introducing the additional language of “importance” (that is, English) in the school curriculum has been a matter of debate. This is primarily due to the concerns that disadvantaged sections may find the additional non-native language a burden, which may discourage them from participating. Therefore, several states of India (including the eastern states of West Bengal and Tripura) have toyed with the removal of English from the curriculum of vernacular medium government-funded public schools[4].

On the contrary, the policy of teaching only the (non-English) native language in schools also raises serious concerns. Evidence from Psycholinguistics studies suggests that the ability to learn a second language is much higher among children if started early in life (Curran, 1961; Heckman 2007; Nikolov and Mihaljevic Djigunovic, 2006)[5]. Additionally, English is an important *lingua franca* in India, and most white collared jobs require a working knowledge of English. With the increasing labor market returns of English skills in a globalized world (Munshi and Rosenzweig, 2005; Chakraborty and Bakshi, 2012, Shastry, 2011), if teaching the English language early in life is abandoned, then its labor market premium may only be accessed by elites who can afford private investments in acquiring this language skill.

Against this backdrop, we empirically examine the impact of the abolition of English language from school curriculum of primary grades on private investment in acquiring English skill. We do that by using a language policy intervention followed by the policy reversal in the

[4] English is 44th on the list of languages with the most native speakers (Azam, et al 2013) and thus is considered as the non-native language in India.

[5] Although Jaekel et. al (2017), in their study of the Foreign Language policy in Germany do not find a supporting evidence for this.

Indian state of West Bengal as a source of exogenous variation in exposure to English language in schools. Even though we study the policy change in West Bengal, the findings of this study are relevant for other countries where there is policy debate around the inclusion of non-native language in school curriculum and the language policy centers around the native language.

West Bengal is an Eastern state of India where English has been taught as a second language in all vernacular (largely Bengali) medium government schools from the very first grade, till the year 1983. However, the Left Front government, which gained power in West Bengal in 1977, removed the teaching of English subject from primary classes (classes 1 to 5) from the year 1983, and English was introduced as a second language only in secondary school.

The main reason for the introduction of this policy was that English was considered as an additional burden for the underprivileged students. The removal of it from the primary level was expected to reduce the dropout rates in West Bengal (Bagchi, 1981). While the discussion around the significance of English, and preference for instructions only in native language had gained momentum, with the left Front government gaining power in 1977, the removal of English from the public school curriculum in primary classes was suddenly announced in 1983. This provides one source of variation in exposure to English language by cohorts. The second source comes from the gradual reversal of the policy two decades later. Increasing criticism of the English removal policy from the opposition parties, particularly, around the parliamentary elections in 1999, and rising demand for English education that has high labor market value by the middle class led the government to gradually reverse the policy. English was reintroduced in the 3rd grade from the year 1999, and from the 1st grade from the year 2004.

We expect the removal of English language from primary classes of public schools to create inequality in access to learning English language. The concerned and affected parents the economically better-off households who can afford to send their children to private schools (henceforth, the '*movers*') are expected to do so, since these schools were not covered under the policy and continued teaching English. We expect the well-off households to make this switch, because unlike public schools, private schools charge fees for providing education[6].

On the other hand, the households that are unable to afford private schools and continue enrolling in public schools (henceforth, the '*stayers*') may want to supplement their children's education by available means, including spending on private tuition or by purchasing additional materials, based on affordability. Private tuitions or supplementary education centres are pervasive in all Indian states, and more prevalent in West Bengal and have been shown to positively affect student performance in schools. Thus, the relatively well-off are expected to make higher private investments in education whereas the economically disadvantaged who cannot afford these private investments are expected to lose labor market premium of the English language skill. This is likely to exacerbate the existing gap between the rich and the poor, more so because the policy did not seem to have the beneficial effect of improving enrolment amongst the disadvantaged, the purported aim of its introduction. Roy (2004) does not seem to find evidence of an increased in enrolment or a reduction in dropouts due to this policy, particularly amongst the poorest households. Our data reveals that the current school enrolment rates among the 5-11 age cohort (relevant for primary grades) are 89 percent, 73 percent, 86 percent, and 95 percent in the years 1986-87, 1995-96, 2007-08, and 2014 respectively. The primary grade-relevant age cohorts of the first two rounds being affected and the last two rounds being unaffected by the policy, we would expect an increase in enrolment at least between the first two rounds of data (if the policy helped), but enrolment seems to decline in 1995-96. The fact that the primary graders in the last two rounds remained

[6] Such increased demand for private schools may even increase the cost (or price) of private schooling, making private school accessible to elites only.

unaffected, along with the increased enrolment in those two rounds too does not indicate that the policy may have helped in increasing enrolment[7].

Using the nationally representative National Sample Survey (NSS), a repeated cross-section data of households, we follow all three cohorts that are: 1) before the English-removal policy period (hence, exposure to bilingual education from the primary level), 2) after the English-removal policy period (no exposure to bilingual education at the primary level in public schools, and 3) during the period of the policy reversal (when bilingual education at the primary level was restored again). The exogenous variation in the time of roll out of the "English removal" policy from the primary level and the consequent policy reversal, along with the survey data on children's school attendance status, help us to construct the treatment variable of our interest. The variable *Affected* is constructed based on the information on whether the child is currently studying in the primary (grades I to V), middle (grades VI to VIII), or secondary level (grades IX, X) during the corresponding NSS round. The pooled cross section data helps us to ensure that we take care of the common time trend that may have affected the investment decisions on education over time.

Our estimation results suggest that, as expected, the economically well-off households among the cohorts *Affected* by the language policy have a higher likelihood of enrolling in private schools. Among the *stayers* in government schools, using a two-part model, we find that economically well-off households spend more on private tuition. These effects are robust to birth-year time-trends. The inclusion of district fixed effects in our regressions helps us to take care of the potentially confounding factors like differential access to public or private

[7] We are unable to produce a causal estimate of enrolment because a clean identification would require specific information on each student about the current grade of study and when the student dropped out. However, the current grade of study is asked to students that are currently enrolled only, and there is no other information on the time when the children dropped out. Since the policy affected the primary grade level students, without the knowledge of the specific grade when children drop out, estimating the causal impact on enrolment is challenging. Roy (2004) presented estimates of the effect of the policy on enrolment using Orissa as a comparison unit, so we avoid replicating the same exercise.

schools, supplementary resources, and school infrastructures across districts. The inclusion of grade-fixed effects ensures that factors like additional expenses incurred due to studying in higher grades do not confound our results. In addition to documenting higher spending on private tuitions by economically better-off *stayers*, we find evidence of lower spending on other types of educational components by the poor and middle-class. However, the affected cohort from the richest strata does not seem to spend differently on other components, as compared to their unaffected counterpart. We also find evidence of lower investments in books by all sections.

Our finding has important policy implications. While instruction in the mother tongue in early age has been found to aid in better learning outcomes, we show that ignoring non-native English language can lead to increased inequality in private investment in education. The inequality in private investment can create differential access to labor market opportunities resulting in further widening of the socio-economic divide between the rich and the poor. The narratives around the language policy in a multilingual country like India (with at least 22 scheduled languages and 122 regional languages) should therefore consider assigning equal importance to English along with its native languages. This would help to avoid the recent trends where the government took arbitrary decisions of converting vernacular medium public schools to English medium schools to serve popular demand (Joshi, 2014; Mukherjee, 2021; PTI, 2017; Sharma, 2020; TNN, 2021). The findings also suggest that it is important for the state to invest on bilingual education from an early age. The National Education Policy (Government of India, 2020) of the government of India emphasizes the importance of multilingualism in the cognitive education of children and encourages instruction in the local language in primary grades. However, it does not emphasize the importance of teaching English as a subject at the primary level.

While the pros and cons of bilingual education have been widely studied, there is limited work that focuses on the widening of the societal gap arising from ignoring globally recognized non-native languages. We contribute to this strand by estimating the impact of removing bilingual education (non-native English language) at the primary level on private investment in education.

While the existing literature produces evidence of the importance of the English language on labor market outcomes (Chakraborty and Bakshi, 2016; Shashtry, 2012) or occupational choice (Munsi and Rosenzweig, 2006) in India, our present work investigates the impact on private investment as an effect of the policy. Our research question is motivated by the above findings on the labor market premium of English language skills because the additional premium is expected to push up the private investment by the households that are able to pay. Our primary objective is to explore the potential of such a policy on deepening the socioeconomic divide by investigating differential private investments as an impact of the policy.

The work that is closest to our paper is Roy (2004), which studies the impact of the English removal policy in West Bengal on enrolment, dropouts and private investment in education. While the primary outcome of our paper is two different forms of private investments, that are enrolment in private schools and spending on private tuitions, we also use a considerably different methodology using the three phases of the policy change. Roy (2004) uses the introduction of the policy only, to capture variation in exposure to English language across cohorts and compares West Bengal with other (arbitrarily chosen) states to identify the impact of the policy. The educational policies in India being implemented at the state levels, almost all the states chosen to have been shown to have very different education outcomes from West Bengal.

Rather, we innovatively use the introduction *as well as* the reversal of the policy to estimate the causal impact. Since the policy was implemented throughout the state, finding a suitable counterfactual across space is challenging. Instead of comparing West Bengal with other neighboring states as done by Roy (2004), we focus on the comparison of cohorts, taking advantage of the policy reversal. This ensures that the comparison of treated and control groups does not come from different states with different educational policies. The policy reversal also helps us to avoid the comparison between the elder and younger cohort only, who could have different outcomes due to the difference in time periods. Rather, the way, we construct our unaffected groups (counterfactuals) come from different time period, and we are able to control for general time trends.

The rest of the paper is organized as follows. Section 2 discusses the background in detail. Section 3 explains the data, variables, and methodology. Section 4 elaborates on our main findings. Section 5 checks for heterogeneity in the main effects and Section 6 discusses the potential concerns and robustness checks. We conclude with policy recommendations in section 7.

2. Background

Beginning in 1757, British colonial rule in India lasted for close to two centuries which led to the institutionalization of English as an important *lingua franca*. Abandoning the initial policy of non-interference in the indigenous education system, this institutionalization of English began in 1835 when Governor-General Lord William Bentick, based on the recommendations of Thomas Babington Macaulay, made English the official language as well as the medium of instruction in all educational institutions. However, given the shortage of teachers who could teach in English, and opposition by a large section of society, this policy was partially reversed. Consequently, vernacular languages were restored as the medium of instruction in primary and secondary schools with the English language being a part of their curriculum. However, English was made the medium of instruction in higher education. The debate on the common language in multilingual India continued post-independence, and the constitution, recognizing that the country had already invested in the English language given its colonial past, named both Hindi and English as the official language. This means that states can legislate their own language for official communication within the state but the communication between the states has to be in English or Hindi.

The state of West Bengal, an eastern state of India, which houses the former British-India’s capital Calcutta (now Kolkata) continued with this cultural legacy taking advantage of the rich presence of English in all spheres of government activities. English was taught as a mandatory second language in all government-funded vernacular (Bengali) medium schools from the very first grade of the primary section. In addition, there has been a very small number of vernacular medium private schools that have English as a mandatory subject, and some private schools where the medium of instruction is English (also called English medium schools).

While the English language skill continued to generate higher labor market returns (Azam et. al., 2013; Chakarabarty and Bakshi, 2016), lower school enrolment in regions linguistically far away from English (Shastry, 2012) raised concerns about the inclusion of English in the school curriculum. English was believed to be an additional burden for disadvantaged students. Left-leaning policymakers and intellectuals thus advocated for instruction in *only* the local language at the primary level (Bagchi 1981). In the backdrop of this policy debate, the Left-Front government, (a coalition led by the Communist Party) that assumed power in the state of West Bengal in the year 1977, removed the subject of English language from primary grades of vernacular medium government-funded schools in the state from the year 1983[8]. Even though the initiative was expected to improve enrolment outcomes among the disadvantaged sections and thereby establish a level-playing field, it was also met with opposition by several academicians on the grounds that such abolition could result in mushrooming of private schools and coaching centers imparting English education which would only be accessed by the relatively well off (Acharya, 1982; Datta, 1981).

The English removal policy lasted for close to two decades but over time the opposition to this policy grew as the returns to English language in an increasingly globalized world increased. The increase in demand for studying English, particularly from the newly created middle class in West Bengal, created political pressure to reintroduce English in the school curriculum in government schools. The policy was finally reversed (Datta, 1981; Sen, 2015) when English was reintroduced from the 3rd grade in the year 1999, and then reintroduced from the 1st grade in the year 2004. The major policy shifts in the year 1983 and then in the year 2004 create an opportunity for a natural experiment, where we use variation in English exposure over time (also within each grade and in different survey years). This is explained in the next section in detail.

3. Data and Methods

3.1 Data and Variable Construction

We use four rounds of the NSS data, a nationally representative household survey of social consumption expenditures, conducted by the National Sample Survey Organisation (NSSO) of the government of India. The surveys that we use are specially designed to capture

[8] This was one of the many policy changes (including land reforms, devolution of local decision-making to village councils) brought in by the Left-Front government that particularly appealed to the poor rural voters. However, this particular language policy change was never discussed either in election manifesto or was never informed to people prior to the announcement. So, this was more sudden policy change.

expenditures on education (also known as education rounds). The four rounds were conducted in the years 1986-87 (42nd round), 1995-96 (52nd round), 2007-08 (64th round), and 2014 (71st round). NSS surveys are cross-section in nature and do not follow the same households over the years.

The education rounds have detailed information on education particulars of currently studying members of the household including the current grade of study, type of institution, and expenditures on education among others. Additionally, the data has details on the consumption expenditures incurred by the household as well as demographic and education-related information of all household members. We use the information on the type of school attended to construct our first outcome variable, which takes value 1 if the child attends private school, and 0 otherwise. Additionally, we use the information on expenditures on private tuition incurred by children studying in grades I to X in public schools to construct our second outcome variable, which is annual expenditures incurred on private tuition as a percentage of annual per capita consumption expenditures. In order to check if students substitute private tuition expenses by any other educational expenses, we also look at *other expenditures on education* (which include cost of books, examination fees, costs of uniforms or any other expenses) as a percentage of annual per capita consumption expenditure.

In addition, we construct the following variables that could potentially affect educational expenses and systematically differ for children exposed to the intervention.

At the *individual* level: Sex of the child- a dummy variable assuming a value one if the child is a female; and reported age of the child in years.

At the *household* level: household size; distance to the nearest school - a categorical variable which assumes an indicator 1 if the school is less than a kilometre from the house, 2 if it is between one to two kilometres, 3 if it is between two to five kilometres, and 4 if it is more than five kilometres away; and four household level indicators capturing quartile distribution of monthly per capita consumption expenditures of the household (MPCE) in a survey year.

Additionally, we use indicators for major caste affiliations of the household head to capture the difference between the forward castes and the socially disadvantaged household that could influence private investment in education. The major caste groups are, Schedule Tribe (ST)-a socio-economically marginalised indigenous group of India, Schedule Caste (SC) -recognised as disadvantaged based on the history of discrimination and economic marginalisation, and all others, who are relatively better off. We also control for the location of residence (urban or rural), gender of household head, and education level of the household head. The latter is a categorical variable capturing the education level of the head with indicators ranging between 0 to 6, representing illiterate, below primary, primary, middle, secondary, higher secondary education, and graduate or higher degree of the head respectively.

Table 1 here

Table 1 reveals that only 47% of children in the affected cohort take private tuition as compared to 73% of the unaffected cohort. Consequently, we observe that average private tuition expenses (as a percentage of total education expenses) for the affected cohort is lower (26%) as compared to the unaffected cohort (44%). The affected cohort spends about 6.6% of their household's annual per capita consumption expenditures (PCE) on private coaching, which is close to half of that (about 12.4%) for the unaffected cohort. Affected and unaffected cohorts seem to spend as high as 16.5% and 23.3% of their annual PCE respectively, on all components of education, including private tuition. Thus, the unconditional average spending on education for the affected group seems to be less as compared to the unaffected cohort.

Figure 1 reveals the extent of private tuition market in Indian states and of West Bengal. The figure indicates that the proportion of students opting for private tuition in West Bengal is disproportionately higher than the rest of the states in a pooled sample across all survey rounds. A similar trend is seen over the years when we look at this distribution across Indian states by each survey round, as presented in appendix figures A1-A4.

Figure 1 here

3.2 Empirical Framework

Since the policy of English abolition is applicable to primary grade children only, we use the grade-survey year combination in the four rounds of NSS data to construct the *Affected* and *Unaffected* cohortI.

Table 2 here

We do that in such a way that the *Unaffected* cohort consists of children who have crossed the primary level before the English abolition policy (that is, before the year 1983), or studying at the primary level after the policy-reversal in the year 1999. Table 2 summarises the mechanism of construction of the *Affected* and *Unaffected* cohort, where for students in each grade, we have *Affected* and *Unaffected* cohorts due to our innovative method of using the repeated cross-section data. As the table shows, children studying at the primary and middle level in 42nd Round, children studying at the primary, middle and secondary levels in the 52nd round, and those studying at the secondary level in the first half of 64th round (i.e. children surveyed in 2007) constitute the *affected* group. Since our primary outcome variable which captures private investment in education varies by grade of study, our method of construction of *Affected* cohort helps us to compare the outcome within each grade level.

Our identification strategy is very close to Angrist and Lavy (1997) which studies the impact of a change in language of instruction in Morocco on test scores and labor market returns. Just like the case of West Bengal, the policy in Morocco was rolled out in the whole country at once, which did not leave any spatial variation in treatment exposure. Hence, the paper uses the combination of years of schooling and cohort of birth to generate variation in exposure. Using a similar strategy, we use the survey year and the corresponding level (grade) of study of the children in that year to construct the affected cohort. We do not use birth year as a proxy for years of schooling because the survey provides the information on whether the child was studying during the survey period and at what level. This is a more precise measure than using birth year or age, because in India a significant share of students does not study at age-relevant grades, and reporting of age has a high measurement error. The age-grade distribution for the state of West Bengal, reported in Appendix table A1, shows significant variability in age within grades.

We estimate the impact of the language abolition policy on the likelihood of enrolling in private school for the *Affected* cohort by estimating the following regression model. The sample is restricted to currently studying children in vernacular medium schools[9] in grades I to X:

$$Y_{ichdt} = \alpha + \beta\, Affected_{ict} + I_i'\gamma + H_h'\delta + \theta_d + Birth\ year_i + \mu_c + \varepsilon_{ichdt} \quad \ldots\ldots\ldots (1)$$

The dependent variable Y_{ichdt} assumes a value 1 if a child *i* in grade *c* in household *h*, district *d* surveyed in year *t* is enrolled in private schools; 0 otherwise. Among the covariates, 'I_i' and 'H_h' include the individual and household level control variables (mentioned in the data section), respectively, that are likely to influence the decision on private investment and might differ among the affected and unaffected cohorts. It is possible that the *affected* cohort spends differently on private tuitions as compared to *unaffected*, because of the difference in grade of study and not due to the exposure to the policy. Our regression specifications thus have grade

[9] The percentage of students studying in English medium public schools is negligible. Only 1.6% of students in our sample report studying in English medium public schools.

fixed effects (μ_c) to control for grade-specific unobservable like levels of difficulty for which students may need additional support. Note that all the rounds except the 71st round of NSS have information on the current grade of study. For this round, we construct the current grade by subtracting the age at entry in school from the current age (and adding one to it) with the assumption that there are no repeaters. We acknowledge that this may introduce noise in the construction of grade variable for the 71st round. Therefore, we check the robustness of our results to addition of fixed effects for levels of education rather than grades, and our findings do not change (table A6).

We also include district fixed effects (θ_d) to ensure that district specific time invariant characteristics do not affect our estimates as certain districts may have better quality schools or infrastructures than others. Additionally, the inclusion of birth year trend ($Birth\ year$) helps us to take care of the general trend in private investment over the years. We construct the birth year by subtracting the age of the child from the year of survey. We do not cluster standard errors in our primary specification. However, the appendix table A5 presents estimates of the impact after clustering standard errors at the grade level, and our findings remain unaltered[10].

Figure 2 here

For those who continue their education in a government school (*stayers*), the immediate effects of compensating investment by parents, if any, could be reflected in their expenditure on private tuition of their children. This seems plausible since the supply of private schools at least in the short run might not have kept pace with the increased demand created by the removal of English from public schools. This has also been pointed out by existing work evaluating the West Bengal English abolition policy (Roy 2004; Chakarabarty and Bakshi, 2016). Private tuition thus might be an additional alternative medium for private investment,

[10] We could not bootstrap standard errors due to the small sample size.

especially because there is some evidence that private tuition aids in improving student learning outcomes (Dongre and Tewary, 2015).

Hence, we also investigate the impact of the policy on the percentage of per capita annual education expenditures spent on private tuition and other education expenses. The last component includes school fees, examination fees, and expenses for books, notebooks, uniforms and transport costs. Additionally, the supply of private schools, at least in the short run might not increase in response to the increased demand for English education, as has also been pointed by existing work evaluating the West Bengal English abolition policy (Roy 2004; Chakarabarty and Bakshi, 2016), and therefore private tuitions might be an additional medium for private investment.

A Two-Part Model is better suited to study the impact of English abolition policy on expenditures on private tuition. This is because spending on private tuition involves a conscious choice of sending children to private tuition, as can be observed by a large frequency of zeroes in the data (figure 2). The zero expenditures on private tuition arise from the parental decision on participation (opting for private tuition), rather than zero expenses. Once parents choose private tuition for their children, then only we observe positive values of expenses[11]. Thus our outcome variable is a discrete-continuous variable with two features: y_i=0 when parents do not enrol in private tuition and $y_i > 0$ when parents spend a positive amount on tuition. Since the decision of spending on private tuition happens in two stages, an OLS model is unable to account for this. Rather, a Two-Part model, which explicitly models the two decisions is likely to be a better choice (Madden, 2006).

Like a double hurdle model, two-part model fits a binary choice model for the selection part and an appropriate regression model for the conditional part. As compared to double hurdle

[11] It is not possible to have zero expenses in our data if children opt for private tuitions (because it is not free of cost).

model, which is better suited for discrete outcome variables, a two-Part model does not allow for the outcome variable to take zero value once the hurdle is crossed. It is also important to note that a two-part model does not require any assumption about correlation between error terms of the participation equation and expenditure equation to get consistent estimates. Additionally, we believe that Heckman selection model would not be appropriate for our outcome variable as the zeroes reported for the private tuition expenses are true zeroes and not censored values of positive outcomes[12].

We use a Probit model to explain the participation in private tuition and an OLS model to explain the positive expenditure on private tuition. The overall predicted expected value is the product of expectations from the first and second parts of the model.

Finally, we estimate an OLS model to investigate the impact on other components of education expenditures for the *stayers* using specification (1). Since this sample consists of currently studying children, and all of them need to spend positive amount on other components that include school fees, notebooks, transport, or other supplies, the OLS model is a better fit here.

4. Discussion of Results

4.1 Moving to Private Schools (Movers)?

Table 3 here

Panel A of Table 3 reports the results. Our preferred model specification is presented in the last column. The columns before that help to check the robustness through varying covariates and fixed effects. All the specifications indicate that the affected cohort seems to have a higher likelihood of enrolling in private schools as compared to the unaffected cohort. The magnitude of the coefficient indicates that there is around 3 percentage point higher likelihood of enrolling

[12] We use *twopm* package of Stata for empirical estimation of private coaching expenses (Belotti et al.).

in private schools for the *Affected* cohort. Given that the average likelihood of enrolling in private schools is 24 percent, this represents a 12.5 percentage increase. However, this effect could be driven by overall trend of increasing enrolment in private schools, but our estimates are robust to addition of birth-year trend in our final specification. Additionally, since the unaffected cohort consists of not only those that entered primary grades before 1983, but also after 2004, it seems unlikely for the increasing trend of enrolment in private schools to have confounded the effect of the policy. We also check the robustness of our results to district specific time trend to address the concern that growth of private schools could be faster in some districts as compared to the other and that could be associated with exposure to the policy (e.g. availability of government schools may be less and hence exposure to policy could be less) (see Appendix Table A2).

Panel B of Table 3 examines if the economic conditions of the household matter for the impact of the policy on the likelihood of enrolling in private schools. We do that by dividing the annual per capita consumption expenditure of the household into four quartiles. Consumption expenditure is used as a proxy for the economic conditions of the household as the NSS does not collect data on household income. As expected, the panel shows that the reported increase in the likelihood of enrolling in private schools is driven by relatively well-off households (those in the third and fourth quartiles) and there is no impact on the affected cohort belonging to poor households. This result is robust to district-specific time trends (Appendix Table A2). Our results are consistent with the existing literature that has documented that preference for English education is an important reason why households switch to private schools from public schools (Kumar and Choudhary 2021; Kingdon 2020; Lahoti and Mukhopadhyay, 2019; James and Woodhead 2014).

4.2 Effects on the Stayers in Government Schools: On Private Tuition

The results for private tuition expenses are presented in Table 4. As documented for enrolment in private schools, economically well-off households could be more likely to opt for private tuition, and spend more on private on that (Azam, 2016). We, therefore, estimate the impact on *Affected* separately for four consumption expenditure quartiles.

Table 4 here

The first column presents estimates for the complete sample, and the next four columns present the results for quartiles 1 to 4 of per capita annual consumption expenditure. Our results suggest that there is no significant difference in spending on private coaching between the affected and unaffected cohorts in the full sample, but the affected cohort in the poorest quartile seems to spend less on private tuition. However, in the other three quartiles, the affected households seem to spend more on private tuition. This suggests that the poor in the affected group reduce their spending on private tuition by 3 percentage points, and the middle class and rich increase their spending by 1 to 2 percentage points.

As explained earlier, since an OLS model does not take into account the decisions of enrolling and spending on private tuition separately, we present estimates from the two part model in Table 5.

Table 5 here

Panel B reports marginal impacts obtained from estimating the binary selection equation. Results suggest no significant impact on the likelihood of taking private tuition for the affected cohort in the full model specification (column 1). In line with the OLS results (Table 4), this specification suggests that, on average, the affected group does not spend differently from the unaffected group on private tuition. However, the second column indicates that the likelihood of opting for private tuition by the poorest quartile is about 7 percentage points less among the affected cohort. As compared to the average, this amounts to a 12 percent decline in the likelihood of attending private tuition. However, for the middle-income households (second and third quartile of expenditures distribution), the affected cohort's likelihood of attending private tuition seems to be higher by 5 to 6 percentage points, a 10 percent impact as compared to the average. Richest households in the sample have a 3 percent higher likelihood of opting for private tuition (however, the p-value is 0.20). While the reduction in the likelihood of attending private coaching for the poorest household may seem a bit puzzling, it could indicate that poor households, who were earlier opting for private tuition to cope up with English subjects taught in school, now no longer have to do so because of the abolition of the teaching of English. Before the abolition of English education, necessity seemed to be the driving force, and after the abolition, affordability became a driving force in their decisions to invest in private tuition.

Panel A of table 5 reports the combined marginal impacts obtained from estimating both the binary selection equation as well as the conditional expenditure equation. Column 1 shows that there is no impact of the policy on private tuition expenses on the affected cohort in the full sample. However, column 2 suggests that private coaching expenses are about 3.5 percent point lower for the affected cohort belonging to the poorest quartile, driven by lower likelihood of joining private tuition as well as lower spending on tuition for those using private coaching services. Among the richer households, the combined marginal effects seem to be weakly higher by 0.9 to 1.5 percent for the affected households. Overall, the share of spending on private tuition by poor households seems to be lower by 36 percent whereas it is higher by 10 percent for the relatively well-off as compared to the average. These estimates are robust to district time trends (Appendix table A5), the use of education-level fixed effects instead of grade-fixed effects (Appendix table A6), and the clustering of standard errors at the grade level (Appendix table A7). The full specifications with coefficients on all the control variables are reported in Appendix tables A3 and A4.

This finding, at first, might indicate that poorest *Affected* households have become better off after the policy as their likelihood of attending as well as spending on private tuition declines. However, the documented increase in private tuition expenditure and movement to private schools by relatively well-off households suggests that children value English education which was freely provided in government schools before the policy. The abolition of English language forces relatively better-off households to either switch to private schools or supplement government school education with private tuition. The poorest households get deprived of the opportunity to learn English at the primary level which has been documented to provide higher labor market returns. This could mean a loss of lifetime income even though their current expenses see a fall. In the next part, we investigate the impact of this policy on other components of education expenditures.

It is also important to note that the usually well-off households enroll their children in private schools. Therefore, we observe poorer students in government schools (Table 3). Hence, the negative impact on spending that we see among the poor and affected households attending public school is likely to be a lower bound, and the true impact on spending may not be as lower as estimated. It also implies that the positive effect on spending for rich households among the government school stayers may be an underestimate of the true impact. Affected in government schools may have actually spent higher than estimated.

4.3 Effects on the Stayers in Government Schools: Other Educational Expenses

Table 6 here

Panel A of table 6 presents OLS estimates of other education expenses as a percentage of annual PCE. Panel B and C reports the disaggregated estimates, where the former panel reports estimates of expenses on books only, and the latter reports estimates for expenses on any other fees, payments, examination fees or uniforms. The spending on transports is also a component of other expenses as defined in panel A, but we do not take that into account in the disaggregated expenses in panel B or C because of a large number of zeroes reported for transport-related expenses.

Overall, the *Affected* cohort seems to spend about 1.08 percentage point less than the unaffected on other components of education expenditures in the full sample (column 1 in panel A). The poorest three quartiles of households seem to spend about 3.6, 1.9 and 2.12 percentage points less on other components of education excluding private tuition (columns 2 to 4). This amounts to about 20 to 40 percent lower expenses for the poorest three quartiles as compared to the sample average. The affected cohort among the richest households does not seem to spend differently on other education expenses. This finding suggests that middle-class households who increase their spending on private tuition, cut down on other educational expenses due to limited financial resources, whereas, rich households do not seem to make such adjustments.

As we disaggregate the spending further and estimate the spending on books in panel B, the affected cohort from all economic classes seems to spend less on books. This is expected because English, which was taught as an additional subject earlier has been removed from the primary level in public schools. Even though books are supposed to be distributed free of cost at the primary level in public schools, it may indicate that parents may have purchased supporting books for English, and post-abolition of English, the spending on books may have dropped for all.

Additional insights emerge in panel C, when we estimate the impact on spending on other fees, payments, and uniforms (this category cannot be disaggregated further because of the nature of data collection). The affected cohort from the richest households spends about 0.90 percentage points higher than the unaffected cohort (column 5). In the full sample (column 1), there is about 0.55 percentage points higher spending by the affected cohort. We thus find a difference in spending by the affected cohort, and the spending patterns align with the economic conditions of households. This indicates that the economically better-off parents of the affected children supplement their children’s education by additional expenses, which the poor parents are unable to afford.

5. Extensions: Heterogeneity

5.1 Heterogeneous Effects

We examine if there is any heterogeneity in the impact of the policy on the affected cohort by the gender of the child. Existing work has established evidence of male bias in private education investment in developing countries (Azam and Kingdon, 2013). Since private coaching is an additional investment to supplement children’s education, the bias is expected to be even more. Table 7 reveals that the likelihood of attending private coaching is about 6 percentage points less for the female children as compared to the male children among the poorest households in the affected cohort (first column of Panel B). Expectedly, the combined marginal impacts reported in panel A suggests that there is 2.7 percentage points lower spending on private coaching for females amongst the poorest households. However, we do not find this to be the case for the richer households (presented in columns 2-4).

Table 7 here

In columns 5 to 8, we report the heterogeneity of the impact by education of the father of the child. Results suggest that the impact of the policy on the likelihood of attending private tuition does not depend on whether the father of the child has completed high school education or not for poor and middle-income households. However, there is weak evidence that the father holding a high school degree reduces the likelihood of attending private tuition and spending on private tuition for the richest *Affected* households. This finding suggests that these parents have the necessary skills to teach their children English at home. Also note that if parents’ education signifies parents' ability to invest on a child, then this effect is to an extent captured by household expenditures.

Table 8 here

Table 8 reports the heterogeneity in the impact of the policy by caste affiliation of the households. Rows 2 and 3 report the interaction between *Affected* and a dummy indicating if the household head belongs to Scheduled Caste (SC) or Scheduled Tribe (ST), respectively – the two most marginalized and disadvantaged communities of India. First columns in Panels A and B suggest that the policy reduces the lower likelihood of attending private tuition as well as spending on private tuition for the poorest ST households. We, however, do not find any differential impact on SCs. This result is expected given that STs are the most economically marginalized and geographically isolated groups even when compared with SCs. Given their geographical isolation and prominent involvement in agriculture, it is possible that they do not consider English as important for their occupation as compared to other caste groups.

Table 9 here

Table 9 examines if the reported increase in private tuition expenses for the affected cohort is driven by investment on students in primary grades when the policy shock happens, or in middle and secondary grades when they are exposed to English subject for the first time. Impact on primary level students would provide supporting evidence of our argument that well-off households make up for the potential loss in English skill premium by investing on private coaching for their children. Impact on students studying above primary grade, on the other hand, would point towards additional investment by rich households because students face difficulty in coping up with a new language in senior grades.

The first four columns in both the panels are from the primary school sample and the last four columns are from the secondary school sample. In line with the mechanism highlighted earlier, Panels A and B indicate that none of the coefficients are significant when the sample is restricted to secondary school going children. However, we find evidence of negative coefficients for the poor quartiles and positive coefficients for the relatively well-off households for the primary school sub-sample. This table thus suggests that the documented higher expenses on private coaching made by rich households and lower expenses by the poor households in the affected group are driven by primary school students.

6. Potential Concerns and Test for Robustness

Even though the timing of the introduction and reversal of the English removal policy was potentially exogenous, there could be events that coincided with the years in which the policy was applicable and could therefore confound our estimates. One such factor could be an increasing preference for private investment in education. This could confound the observed impact of the policy, particularly, if the preference for private school was higher for the affected cohort – those who were in primary classes after 1983 and before 2004. While this may seem unlikely because cohorts attending primary school after 2004 are unaffected by the policy, and they too should have faced a higher preference for private schools in that case. Moreover, even if this were to be the case, we would observe an increase in enrolment in all private schools, including English medium private schools, as well as increased investment in private coaching for students in English medium private schools.

Table 10 here

We perform this test and the results reported in Table 10 suggest that for the affected cohort there has been neither any increase in enrolment in private English medium schools nor in private tuition and other education expenses for those who attend private English medium schools. While vernacular medium private schools could be an alternative school option for students from the affected cohort (as private schools could still continue to have English as a subject in the primary section), English medium schools could hardly be considered as an alternative. Adjusting to the new medium of English from a vernacular one in higher classes involves a risky decision, and English medium schools are much more expensive as compared to vernacular medium government schools. These results thus show that a higher preference for private investment in education amongst the affected cohort is highly unlikely to be driving the observed result.

We identify the impact of the English removal policy by including class fixed effects in all our specifications and thus using the variation in the status of getting affected within the same class across different survey years. While this ensures that class-specific time-invariant factors cannot confound our results, it does not address the possibility of time-varying class-specific variables affecting the education choices of the affected cohort. For example, there could be an expansion of private schools or coaching classes over time particularly catering to certain grades which could confound our results. We address this concern by adding class-specific survey year trends that take care of all the class-specific factors that change over survey years. The results reported in Tables 11 and 12 confirm that our estimates remain robust.

Table 11 and 12 here

We also provide further evidence to credibly establish that the cohort exposed to the English removal policy in West Bengal makes higher *private* investment in education (particularly the relatively well-off), by presenting estimates from a difference in difference specification. We use the neighboring state Odisha as the control group where no such language policy was adopted. Odisha has been chosen as the comparison state because it shares geographical as well as cultural proximity with West Bengal. Additionally, amongst all the neighbouring states of West Bengal, Odisha is the closest in terms of the percentage of students opting for private tuition. However, we acknowledge that Odisha could have been undertaking many other education policy changes during our sample period (1986 – 2014) and some of these changes could potentially affect the affected cohort more and may also have implications for private investment in education. Even though we have controlled for policy changes by adding state-specific time trends in our specification, this does not completely address the issue. Therefore, this test is only presented as a robustness test for our main findings. Table 13 confirms that the affected cohort in West Bengal has a differentially higher likelihood of enrolling in private vernacular medium schools. Table 14 shows that the likelihood of attending private coaching is lower for the poor households amongst the affected cohort. But there is a weak evidence of higher spending by the relatively well-off amongst the affected cohort (p-value 0.14). These findings strengthen our main result of an increase in private investment in education following the removal of English subject from government schools.

Table 13 and 14 here

One potential concern would be about the way in which we construct the *Affected* variable. Cohort in 4th (5th) grade in 1983 (1984), did not have English only in these 2 years (1983 and 1984), as the policy of teaching English at the secondary level started in 1983. So, they started English again in 1985 in 6th grade. These students missed out English only for two years and not in all the primary years. Hence, it could be argued that they should not be counted as "*affected*" ideally. However, the whole curriculum till the end of secondary school was converted to the new (reduced) standard, the textbooks from 6th grade too had to be changed to the preliminary level of learning English. Hence this cohort would still be affected with respect to the skill attainment because they studied under the new curriculum at the secondary level. Thus, without loss of generality one can safely assume that even with one or two years of exposure to the English language during primary school years, this cohort during the transition could not pick up the English skill like their older cohort. Nevertheless, we confirm that our results are robust to excluding children studying in grade 7 in the survey year 1986-87 from our sample[13].

[13] Results available on request.

7. Conclusion

Bilingual education or teaching in a non-native language from early on in life has been a contentious issue throughout the world. While the labor market literature has been able to establish positive premiums, there are other serious concerns of discouraging disadvantaged sections by promoting early education in a non-native language. This has been a policy debate in countries with colonial history like India.

Having been ruled by the British for about two centuries, bilingual education involving the teaching of English as a subject became a norm in most states of India after independence. Education is a joint subject of the federal and the state governments in India, and states have implemented their own language polices. West Bengal, an eastern state of India, which houses the former capital of British India, Kolkata, changed its language policy from the year 1983, with the removal of English subject (hence bilingual education) at the primary level of the government schools. However, this policy was reversed after close to two decades. This gives us the opportunity of a natural experiment to study the impact of the removal of English education policy on private investment in education.

Since private schools did not come under the policy domain, our results show that the affected parents from the economically better-off households send their children to private schools leading to higher private investments in education. Additionally, our two-part model suggests that affected cohorts, particularly those belonging to rich households, studying in government schools have a higher likelihood of opting for private tuition. We also find that private tuition expenses for the affected cohort among the richer households are higher and that of the poorer households are lower due to the policy. A similar disaggregated effect is found on other components of education expenditures of government school students. Most of the affected cohort, on average, spend less on other forms of educational expenses, except the richest households who do not seem to spend differently.

The above findings indicate an increased private investment in learning English through enrolling in private schools or supplementing by private tuition or additional expenses by the households depending on their ability to pay. The altruist parents seem to have attempted to bridge the potential gap in children's learning generated by the English removal policy. This creates an additional channel widening the already existing socio-economic gap between the rich and poor in the future. This finding should be of interest to policymakers studying not only the effectiveness of bilingual education policy but also the importance of English language ability in a country that has already invested in English education. While the National Education Policy of India (Government of India, 2020) delves into encouraging multilingualism and rightly extends support on providing flexibility to the regional (state) governments on encouraging regional languages, it does not specifically focus on the importance of English language which has been found to be strongly associated with higher labor market returns. The latter seems to be equally important, in a globalized world with increasing preference for learning English (as is evidenced by increasing private investment in English language) for establishing an equitable society.

References

Acharya, P. 1982. Abolition of English at Primary Level in West Bengal, *Economic and Political Weekly*, 17(4): 124-128, January 23.

Angrist, J. D., and Victor, L. 1997. The effect of a change in language of instruction on the returns to schooling in Morocco. *Journal of Labor Economics* 15:S48–S76 (January, Part 2).

Angrist, J. D., Chin, A., and Godoy, R. 2008. Is Spanish-only schooling responsible for the Puerto Rican language gap? *Journal of Development Economics*, 85(1-2):105-128.

ASER Center. 2020. Annual Status of Education (ASER) Report for Rural West Bengal for survey 2018-2020. Published by Pratham.

Azam, M., 2016. Private Tutoring: Evidence from India, *Review of Development Economics*, 20(4): 739-761.

Azam, M., and G. G. Kingdon. 2013. Are girls the fairer sex in India? Revisiting intra-household allocation of education expenditure. *World Development* 42: 143-164.

Azam, M., Chin, A. and Prakash, N. 2013. The returns to English-language skills in India, *Economic Development and Cultural Change*, 61(2):335–67.

Bagchi, J. 1981. Storm over English in West Bengal: A Comment. *Economic and Political Weekly*, 16(22): 993–994.

Belotti, F., Deb, P., Manning, W. G., and Norton, E. C. 2015. twopm: two-part models. *Stata Journal*, 15(1):3-20. ISSN 1536-867X.

Barreto, A. A., 1998. Language, Elites and the State: Nationalism in Puerto Rico and Quebec. Praeger, Westport, Connecticut.

Cafferty, P. SJ., Rivera-Martínez, C., 1981. The Politics of Language: The Dilemma of Bilingual Education for Puerto Ricans. Westview Press, Boulder, Colorado.

Chakraborty, T. and Bakshi, S. K. 2016. English language premium: Evidence from a policy experiment in India, *Economics of Education Review*, 50:1–16.

Chatterjee, S. 2018. Do Private Tutors Enhance English Language Ability? Regression Discontinuity Evidence from a Policy Experiment in India. *Bulletin of Economic Research*, 70:139-149. https://doi.org/10.1111/boer.12142

Curran, C. A. 1961. Counseling skills adapted to the learning of foreign languages. *Bulletin of the Menninger Clinic*, 25:78-93.

Datta, B. 1981. Storm over English in West Bengal. *Economic and Political Weekly*, 16(14): 619–622. http://www.jstor.org/stable/4369675

Dongre, A., and Tewary, V. 2015. Impact of private tutoring on learning levels. *Economic and Political Weekly*, 50(41):72-80.

Government of India. 2020. The National Education Policy, Ministry of Human Resource Development, accessed at: https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf on 7th February, 2022.

Guha Thakurta, S.N. 1998. Rigmarole over English in West Bengal, *Economic and Political Weekly*, 33(22):1300-1303.

Business Of Private Tutorials In India Now A Multi-Billion Dollar Industry, accessed online at: http://www.businessworld.in/article/Business-Of-Private-Tutorials-In-India-Now-A-Multi-Billion-Dollar-Industry/28-07-2016-100972/ (accessed on 15 July 2020)

Jaekel, N., M. Schurig, M. Florian, and M. Ritter. 2017. From Early Starters to Late Finishers? A Longitudinal Study of Early Foreign Language Learning in School. *Language Learning*, 67(3): 631–664.

Joshi, S. 2014. U.P. govt. to open English-medium schools, November 29, The Hindu, accessed online https://www.thehindu.com/news/national/other-states/up-govt-to-open-englishmedium-schools/article6646342.ece , on May 15, 2022.

Kingdon, G. G. 2020. The Private Schooling Phenomenon in India: A Review, *The Journal of Development Studies*, 56(10):1795-1817, DOI: 10.1080/00220388.2020.1715943

Kossoudji, S. A. 1988. English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men, *Journal of Labor Economics*, 6(2):205-228.

Kumar, D., and P. K. Choudhury. 2021. Determinants of Private School Choice in India: All about the Family Backgrounds?. *Journal of School Choice* 15.4: 576-602.

Lahoti, R., and R. Mukhopadhyay. 2019. School choice in rural India: Perceptions and realities in four states. *Economic and Political Weekly*, 54.49: 51-57.

Madden D. 2008. Sample Selection Versus Two-part Models Revisited: The Case of Female Smoking and Drinking. *Journal of Health Economics*, 27(2):300-7. doi: 10.1016/j.jhealeco.2007.07.001. Epub 2007 Nov 29. PMID: 18180064.

Mukherjee. M., 2021. West Bengal to have English medium school at every block: CM, 10th May. https://www.educationworld.in/west-bengal-to-have-english-medium-school-at-every-block-cm/ , accessed on May 15, 2022.

Munshi, K., and Rosenzweig, M. 2005. Traditional Institutions Meet the Modern World: Caste, Gender and Schooling Choice in a Globalizing Economy. *American Economic Review*, 96 (4): 1225-1252.

Nikolov, M. and J. M. Djigunovic. 2006. Recent Research on Age, Second Language Acquisition, and Early Foreign Language Learning. *Annual Review of Applied Linguistics*, 26: 234-260.

Press Trust of India (PTI), 2017. West Bengal plans to introduce English medium teaching in state-run schools. The Indian Express, September 2, accessed online https://indianexpress.com/article/india/west-bengal-plans-to-introduce-english-medium-teaching-in-state-run-schools-4825441/ on May 15, 2022.

Sen, S. 2015. Abolishing English in schools: implications for higher education in West Bengal. Inter-Asia Cultural Studies 16.2: 269-281.

Sharma, H. 2020. Share of kids in English medium steadily up, dips in poorer states, The Indian Express, July 20, accessed online https://indianexpress.com/article/education/english-medium-schools-national-statistical-office-native-languages-6514045/ on May 15, 2022.

Shastry, G. K. 2012. Human Capital Response to Globalization: Education and Information Technology in India. The Journal of Human Resources 47(2): 287–330. http://www.jstor.org/stable/23214392.

Times News Network (TNN), 2021. English-medium schools to be set up in every block: CM Mamata Banerjee, The Times of India, May 11, accessed online https://timesofindia.indiatimes.com/city/kolkata/eng-medium-schools-to-be-set-up-in-every-block-mamata/articleshow/82536696.cms, on May 15, 2022.

James, Z., and M. Woodhead. 2014. Choosing and changing schools in India’s private and government sectors: Young Lives evidence from Andhra Pradesh. *Oxford Review of Education* 40.1: 73-90.

Table 1: Descriptive Statistics of Students of Public School Sample

Outcomes and Covariates	***Affected***	***Unaffected***
Private School Enrolment (binary: 1=Yes, 0=No)	0.34	0.13
Private Coaching (binary: 1=Yes, 0=No)	0.47	0.73
(Private Coaching / Education) expenses for *i*	26.0	44.2
Private Coaching of child *i* / Annual PCE	6.6	12.4
Education Expenses of child *i* / Annual PCE	16.5	23.3
Other Education Expenses / Annual PCE	9.8	10.9
Individual Characteristics (Covariates):		
Sex (female=1, male=0)	43.0	49.0
Age in Years	10.9	10.8
Currently Studying in Grade:		
1	15.7	8.8
2	16.5	10.9
3	13.1	10.7
4	12.0	10.5
5	11.0	9.6
6	10.3	10.3
7	8.7	8.8
8	3.6	12.0
9	5.2	9.3
10	4.0	9.2
Household Level Covariates		
Household Size	6.88	5.51
Distance to nearest school (in Km.)	1.36	1.23
MPCE Quartile 1	0.34	0.30
MPCE Quartile 2	0.27	0.29
MPCE Quartile 3	0.24	0.24
MPCE Quartile 4	0.15	0.17
Caste Affiliation of Household:		
Scheduled Tribes (ST)	5.0	6.0
Scheduled Castes (SC)	27.0	29.0
Others (including OBC)	68.0	65.0
Sex of Head (Female=1, Male=0)	4.6	7.8
Education of Head (Categories):		
Not literate	23.7	32.7
Below primary	18.8	16.7
Primary	24.9	21.7
Middle	16.5	15.1
Secondary	7.1	6.7
Higher secondary	4.3	3.1
Graduate	4.1	3.1
Post graduate and above	0.53	0.6
Year: (1986-87: Round 42)	42.5	8.4
Year: (1995-96: Round 52)	53.6	0
Year: (2007-08: Round 64)	3.9	51.7
Year: (2014: Round 71)	0	39.9
Observations	6868	7726

Table 2: Construction of Affected Group by Grade level and by Year of Survey

Level / Grade	Unaffected	Affected
Level: Primary		
Grade: 1 to 4	All from 2014 data	All from 1995-96 data
	All from 2007-08 data	All from 1986-87 data
Grade 5	All from 2014 data	All from 1995-96 data
	All from 2007-08 data partially unaffected	All from 1986-87 data
Level: Middle		
Grade 6	All from 2014 data	All from 1995-96 data
	All from 2007-08 data partially unaffected	All from 1986-87
Grade 7	All from 2014 data	All from 1995-96 data
	All from 2007-08 data partially unaffected	All from 1986-87
Grade 8	All from 2014	All from 1995-96
	All from 1986-87	
	All from 2007-08 partially unaffected	
Level: Secondary		
Grade 9	All from 2014	All from 1995-96
	All from 1986-87	2007 affected
	2008 unaffected	
Grade 10	All from 2014	2007 affected
	all from 1986-87	
	2008 unaffected	

Note: Source is NSS Social Consumption Expenditure on Education 42(1986-87), 52(1995-96), 64(2007-08) and 71(2014) rounds. Sample - students currently studying between 1st and 10th grades in vernacular medium schools in West Bengal.

Table 3: Movers- OLS Estimation of Schools Choice on Affected

Panel A: Full Sample Test for Robustness to Covariates (Private + Public School)

	(1)	(2)	(3)
Dependent Variable- Binary Choice of School Currently Enrolled in: Private =1, Public=0			
Affected	0.202***	0.057***	0.036***
	(0.000)	(0.000)	(0.000)
Observations	19,269	19,184	19,184
R-squared	0.055	0.171	0.258
Birth year trend	No	No	Yes
District FE	No	No	Yes
Grade FE	No	Yes	Yes
Other Covariates	No	Yes	Yes

Panel B: Final Specification using Sub-samples of PCE Quartiles (Private + Public School)

	(1)	(2)	(3)	(4)
Dependent Variable- Binary Choice of School Currently Enrolled in: Private =1, Public=0				
Affected	-0.006	0.016	0.077***	0.065***
	(0.705)	(0.370)	(0.000)	(0.001)
Observations	5,507	5,079	4,761	3,837
R-squared	0.279	0.240	0.245	0.230
Birth year trend	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes
Other Covariates	Yes	Yes	Yes	Yes
Sample Quartile	1	2	3	4

Note: P-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Quartile 1 is the poorest and following this order Quartile 4 is the richest households according to the PCE distribution. All specifications include Rural-Urban dummy (Urban=1, Rural=0).

Table 4: OLS Estimation of Private Coaching Expenses as a Percentage of Annual PCE (Public School Sample)

	(1)	(2)	(3)	(4)	(5)
Affected	0.368	-3.097**	1.004*	1.273**	2.009*
	(0.426)	(0.016)	(0.067)	(0.045)	(0.056)
Observations	14,524	4,618	4,077	3,508	2,321
R-squared	0.170	0.111	0.346	0.285	0.206
Birth year trend	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Other Covariates	Yes	Yes	Yes	Yes	Yes
Sample Quartile	Full	1	2	3	4

Note: P-values in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications include Rural-Urban dummy (Urban=1, Rural=0).

Table 5: Two-Part Model of Annual Private Coaching Expenses as a Percentage of Annual PCE (Public School Sample)

Panel A: Conditional Expenditure Equation (Continuous-OLS)					
	(1)	(2)	(3)	(4)	(5)
Affected	-0.10	-3.55***	0.34	0.89	1.48
	(0.82)	(0.00)	(0.52)	(0.14)	(0.15)
Observations	14,524	4,615	4,074	3,508	2,305
Birth year trend	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Other Covariates	Yes	Yes	Yes	Yes	Yes
Sample Quartile	Full	1	2	3	4

Panel B: Selection Equation (Binary –Probit)					
	(1)	(2)	(3)	(4)	(5)
Affected	0.01	-0.07**	0.05*	0.06**	0.03
	(0.68)	(0.01)	(0.06)	(0.01)	(0.20)
Observations	14,534	4,625	4,074	3,508	2,305
Birth year trend	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Other Covariates	Yes	Yes	Yes	Yes	Yes
Sample Quartile	Full	1	2	3	4

Note: P-values in parentheses *** p<0.01, ** p<0.05, * p<0.1. All specifications include Rural-Urban dummy (Urban=1, Rural=0).

Table 6: OLS Estimates of Other Education Expenses as a Percentage of Annual PCE (Public School Sample)

	(1)	(2)	(3)	(4)	(5)
Panel A: Dependent Variable is Other Education Expenses (Excludes Private Tuition)					
Affected	-1.08**	-3.56**	-1.90***	-2.12***	0.25
	(0.03)	(0.03)	(0.00)	(0.00)	(0.59)
Observations	14,512	4,613	4,072	3,508	2,319
R-squared	0.13	0.12	0.48	0.27	0.27
Panel B: Dependent Variable is Part of Other Education Expenses Spent on Books					
Affected	-1.63***	-4.26***	-1.61***	-2.29***	-0.85***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	12,034	3,566	3,314	3,026	2,128
R-squared	0.08	0.08	0.45	0.13	0.33
Panel C: Dependent Variable is Part of Other Education Expenses on Fees, Uniform etc,					
Affected	0.55**	0.41	0.08	0.17	0.90***
	(0.05)	(0.65)	(0.76)	(0.53)	(0.00)
Observations	13,374	4,200	3,746	3,264	2,164
R-squared	0.08	0.08	0.29	0.28	0.21
Sample Quartile	Full Sample	1	2	3	4

Note: P-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Estimates in all columns are from full model specifications, including the individual and household level covariates, urban dummy, birth year trend, district fixed effects, and grade fixed effects. Total Education Expenses = Expenses on private tuition + Other Education Expenses. Other Education Expenses = Expenses on Books + Expenses on other payments or fees, examination fees, uniform + Expenses on Transport. We do not include transport in panel B or C because of a large number of zeros there. However, transport cost is included in panel A.

Table 7: Heterogeneity Based on Head's Education- Two Part Model of Private Coaching Expenses as a Percentage of Annual PCE (Public School Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Conditional Expenditures Equation (Continuous – OLS)							
Affected	-2.68**	0.59	0.97	1.22	-3.46***	0.44	1.01	2.45**
	(0.05)	(0.31)	(0.16)	(0.32)	(0.01)	(0.41)	(0.11)	(0.04)
Affected X Female	-2.08	-0.59	-0.16	0.52				
	(0.13)	(0.32)	(0.83)	(0.71)				
Affected X High-School					-1.82	-0.97	-0.69	-2.25
					(0.57)	(0.35)	(0.48)	(0.12)
Observations	4,615	4,074	3,508	2,305	4,615	4,074	3,508	2,305
Sample Quartile	1	2	3	4	1	2	3	4
	Panel B: Selection Equation (Binary –Probit)							
Affected	-0.04	0.06**	0.07***	0.04	-0.07***	0.05*	0.05**	0.05*
	(0.15)	(0.04)	(0.01)	(0.17)	(0.01)	(0.06)	(0.05)	(0.06)
Affected X female	-0.06**	-0.02	-0.03	-0.02				
	(0.02)	(0.44)	(0.34)	(0.60)				
Affected X High-School					0.06	-0.00	0.05	-0.05
					(0.38)	(1.00)	(0.15)	(0.14)
Observations	4,625	4,074	3,508	2,305	4,625	4,074	3,508	2,305
Sample Quartile	1	2	3	4	1	2	3	4

Note: P-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The first four columns check for heterogeneous impact based on child's gender, and the last four columns check for heterogeneous impact based on Head's education (whether above high school or not).Estimates in all columns are from full model specifications, including the individual and household level covariates, Urban dummy, birth year trend, district fixed effects, and grade fixed effects.

Table 8: Heterogeneity Based on Caste Affiliations- Two Part Model of Private Coaching Expenses as a Percentage of Annual PCE (Public School Sample)

	(1)	(2)	(3)	(4)
Panel A: Conditional Expenditures Equation (Continuous – OLS)				
Affected	-4.04***	0.68	0.80	1.35
	(0.00)	(0.23)	(0.21)	(0.20)
Affected X SC	2.33	-0.98	0.41	0.84
	(0.11)	(0.14)	(0.63)	(0.66)
Affected X ST	-5.56*	-1.44	-1.08	2.15
	(0.10)	(0.33)	(0.67)	(0.79)
Observations	4,615	4,074	3,508	2,305
Sample Quartile	1	2	3	4
Panel B: Selection Equation (Binary –Probit)				
Affected	-0.07**	0.05**	0.05*	0.03
	(0.02)	(0.04)	(0.07)	(0.29)
Affected X SC	0.02	-0.03	0.04	0.04
	(0.39)	(0.25)	(0.18)	(0.37)
Affected X ST	-0.09*	0.05	0.02	0.02
	(0.08)	(0.40)	(0.82)	(0.85)
Observations	4,625	4,074	3,508	2,305
Sample Quartile	1	2	3	4

Note: P-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
The Omitted caste category is all others. Estimates in all columns are from full model specifications, including the individual and household level covariates, urban dummy, birth year trend, district fixed effects, and grade fixed effects.

Table 9: Expenditures on Private Coaching – Separate Sub-samples for Primary and Secondary Levels

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Conditional Expenditures Equation (Continuous – OLS)							
Affected	-3.91***	0.45	1.68**	0.59	-2.58	-0.02	0.67	1.48
	(0.0)	(0.42)	(0.05)	(0.7)	(0.45)	(0.99)	(0.51)	(0.32)
Observations	3159	2470	1901	991	1456	1604	1607	1301
Sample Quartile	1	2	3	4	1	2	3	4
Primary School Sample	Yes	Yes	Yes	Yes	No	No	No	No
Secondary School Sample	No	No	No	No	Yes	Yes	Yes	Yes
	Panel B: Selection Equation (Binary –Probit)							
Affected	-0.13***	0.10*	0.23***	0.05	-0.04	0	0.02	0.03
	(0.0)	(0.06)	(0.0)	(0.53)	(0.24)	(0.86)	(0.51)	(0.18)
Observations	3163	2470	1901	991	1462	1604	1607	1301
Sample Quartile	1	2	3	4	1	2	3	4
Primary School Sample	Yes	Yes	Yes	Yes	No	No	No	No
Secondary School Sample	No	No	No	No	Yes	Yes	Yes	Yes

Note: P-values in parentheses, *** p<0.01, ** p<0.05, * p<0.1
Estimates in all columns are from full model specifications, including the individual and household level covariates, urban dummy, birth year trend, district fixed effects, and grade fixed effects.

Table 10: Robustness check: private English medium schools

	(1) Private	(2) Coaching expenses	(3) Other expenses
Affected	-0.006	1.251	2.925
	(0.926)	(0.740)	(0.677)
Observations	973	824	823
R-squared	0.149	0.191	0.148
Birth year trend	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Class FE	Yes	Yes	Yes
Other covariates	Yes	Yes	Yes

Note: P-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Private is a dummy variable taking a value 1 if the child is enrolled in private school. All specifications include Rural-Urban dummy (Urban=1, Rural=0).

Table 11: Robustness to grade time trends

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable- Binary Choice of School Currently Enrolled in: Private =1, Public=0				
Affected	0.058***	-0.031	0.048*	0.115***	0.077***
	(0.000)	(0.241)	(0.068)	(0.000)	(0.007)
Observations	19,184	5,507	5,079	4,761	3,837
R-squared	0.285	0.306	0.267	0.277	0.262
Birth year trend	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Other Covariates	Yes	Yes	Yes	Yes	Yes
Grade time trend	Yes	Yes	Yes	Yes	Yes
Sample Quartile	All	1	2	3	4

Note: P-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Quartile 1 are Poorest and following this order Quartile 4 are the richest households according to the PCE distribution. All specifications include Rural-Urban dummy (Urban=1, Rural=0).

Table 12: Two-Part Model: robustness to grade time trends

Panel A: Conditional Expenditure Equation (Continuous-OLS)

	(1)	(2)	(3)	(4)	(5)
Affected	0.69	-1.88	2.58***	1.39*	0.88
	(0.26)	(0.29)	(0.00)	(0.10)	(0.52)
Observations	14,524	4,615	4,074	3,508	2,305
Birth year trend	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Other Covariates	Yes	Yes	Yes	Yes	Yes
Grade time trend	Yes	Yes	Yes	Yes	Yes
Sample Quartile	Full	1	2	3	4

Panel B: Selection Equation (Binary –Probit)

	(1)	(2)	(3)	(4)	(5)
Affected	0.06***	0.01	0.09**	0.08**	0.07*
	(0.00)	(0.82)	(0.02)	(0.04)	(0.10)
Observations	14,534	4,625	4,074	3,508	2,305
Birth year trend	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Other Covariates	Yes	Yes	Yes	Yes	Yes
Grade time trend	Yes	Yes	Yes	Yes	Yes
Sample Quartile	Full	1	2	3	4

Note: P-values in parentheses *** p<0.01, ** p<0.05, * p<0.1. All specifications include Rural-Urban dummy (Urban=1, Rural=0).

Table 13: Difference-in-difference specification

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable- Binary Choice of School Currently Enrolled in: Private =1, Public=0				
Affected	-0.130***	-0.154***	-0.182***	-0.114***	-0.114***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Affected X WB	0.096***	0.104***	0.161***	0.100***	0.064*
	(0.000)	(0.000)	(0.000)	(0.002)	(0.076)
WB	0.194***	-0.173**	-0.189**	0.393***	1.110***
	(0.000)	(0.012)	(0.025)	(0.000)	(0.000)
Observations	27,520	8,479	7,085	6,485	5,471
R-squared	0.226	0.257	0.219	0.216	0.212
Birth year trend	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Other Covariates	Yes	Yes	Yes	Yes	Yes
State round trend	Yes	Yes	Yes	Yes	Yes
Sample Quartile	All	1	2	3	4

Note: P-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Quartile 1 are Poorest and following this order Quartile 4 are the richest households according to the PCE distribution. All specifications include Rural-Urban dummy (Urban=1, Rural=0).

Table 14: Two-Part Model: difference-in-difference specification

Panel A: Conditional Expenditure Equation (Continuous-OLS)

	(1)	(2)	(3)	(4)	(5)
Affected	-0.03	0.23	-0.06	-0.77	1.08
	(0.96)	(0.90)	(0.95)	(0.44)	(0.42)
Affected X WB	0.31	-1.81	0.83	1.60	-0.11
	(0.66)	(0.35)	(0.39)	(0.14)	(0.94)
Observations	21,410	7,276	5,753	4,950	3,388
Birth year trend	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Other Covariates	Yes	Yes	Yes	Yes	Yes
Grade time trend	Yes	Yes	Yes	Yes	Yes
Sample Quartile	Full	1	2	3	4

Panel B: Selection Equation (Binary –Probit)

	(1)	(2)	(3)	(4)	(5)
Affected	0.03	0.04	0.03	0.05	0.06*
	(0.11)	(0.28)	(0.40)	(0.18)	(0.05)
Affected X WB	-0.04*	-0.08*	-0.05	-0.06	-0.03
	(0.07)	(0.07)	(0.27)	(0.17)	(0.44)
Observations	21,420	7,286	5,753	4,950	3,388
Birth year trend	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Other Covariates	Yes	Yes	Yes	Yes	Yes
Grade time trend	Yes	Yes	Yes	Yes	Yes
Sample Quartile	Full	1	2	3	4

Note: P-values in parentheses *** p<0.01, ** p<0.05, * p<0.1. *WB* is a dummy indicating that the state is West Bengal. All specifications include Rural-Urban dummy (Urban=1, Rural=0).

Figure 1: Proportion of Students Opting for Private Tuitions across States of India (Pooled Sample across all Survey Rounds)

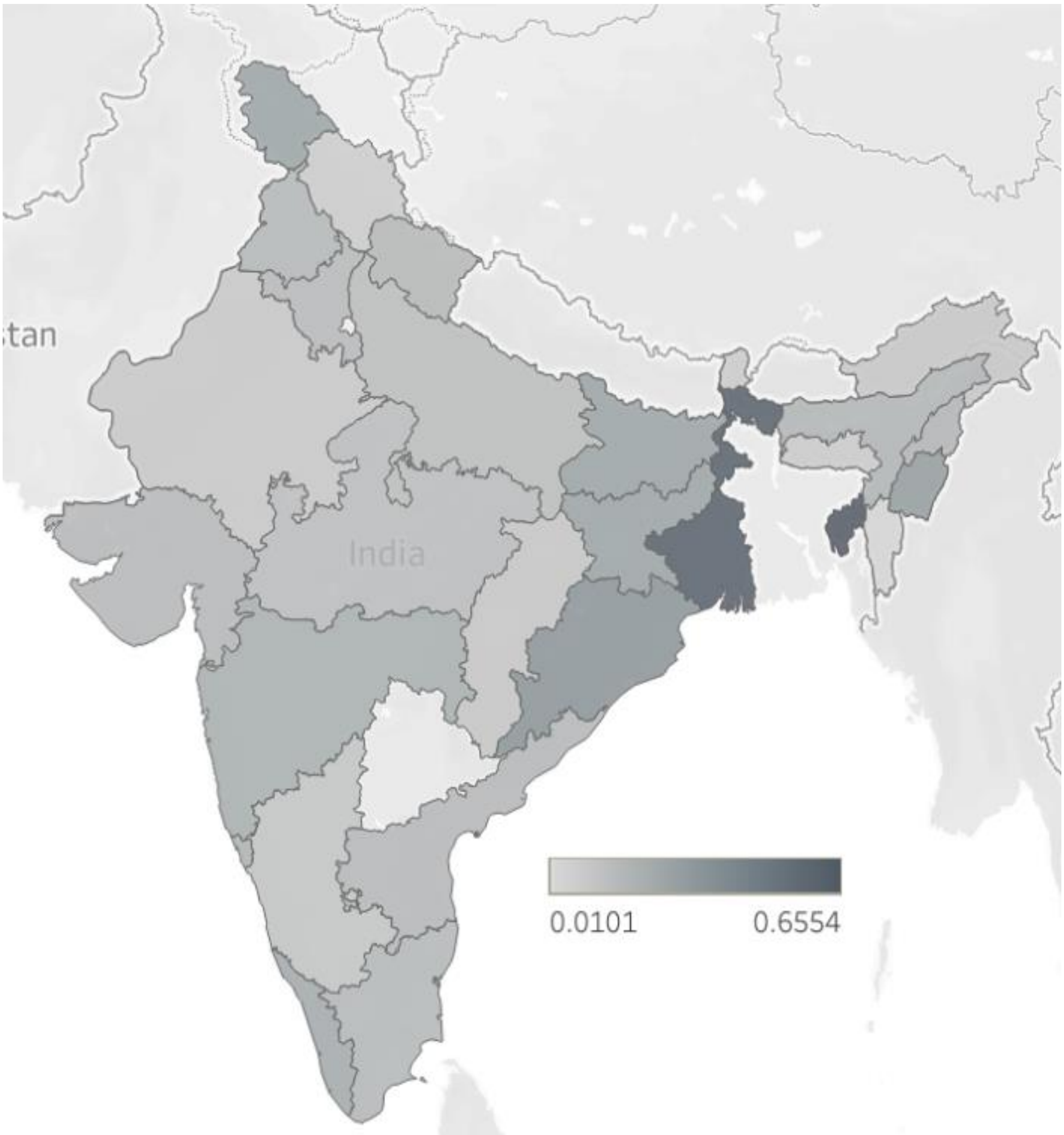

Figure 2: Histograms Plotting Tuition Expenses of Currently Enrolled Children