Intergenerational Effects of Compulsory English Education: Evidence from India

Tanika Chakraborty* Somdeep Chatterjee† Shubham Ojha‡

Abstract

This paper examines the intergenerational effects of parental exposure to compulsory English education on household educational investments and children's learning outcomes in India. Our identification strategy relies on the generalized difference-in-differences approach that leverages the differential timing of state-level mandates to make English compulsory in the primary academic curriculum, leading to differences in parents' years of exposure between 1979 and 2004. We find that parental exposure to compulsory English significantly increases educational investments: the probability of children attending private schools rises, and households spend more on out-of-school tutoring. However, these investments do not translate into better learning outcomes. Instead, children of treated parents experience declines in reading and math scores. We provide evidence that this paradoxical effect is explained by a substitution mechanism: compulsory English reduced parents' own likelihood of completing schooling, weakening their capacity to provide educational support. As a result, parents substitute household inputs with private schooling and tuition of heterogeneous quality. Our findings highlight how well-intentioned language policies can generate unintended intergenerational trade-offs between aspirations, investments, and actual learning.

Keywords - English exposure, intergenerational effects, educational investment, language policy
JEL Classification - I21; I25; I28; J62

*Indian Institute of Management Calcutta

†Indian Institute of Management Calcutta

‡Indian Institute of Management Calcutta

1 Introduction

Over the past few decades, there have been marked improvements in schooling availability in most parts of the world Barro and Lee (2013). Even though this has meant that more children attend school, learning deficits persist extensively in developing countries like India, where increased access to schooling has not consistently translated into enhanced learning (Glewwe and Muralidharan, 2016). For instance, the Annual Status of Education Report (ASER) consistently highlights that a significant fraction of children, despite attending school, are laggard with basic reading and arithmetic skills (ASER, 2017). As an example, although the net enrolment rate in India is close to 100 percent, only 43 (33) percent of sixth (seventh) graders in India could read a vernacular text at the second-grade level, and around one-quarter of fifth-grade students were able to solve a math (division) problem.

While a bulk of the interventions addressing these learning deficits have focused on school-level inputs (Kremer et al., 2013), in this paper, we study the alternate inputs channel, viz, household inputs. This is also a growing literature focused on information provision (Jensen, 2010, Loyalka et al., 2013), conditional cash transfers (Baird et al., 2011, Maluccio et al., 2009), scholarship programmes (Blimpo, 2014), etc. Specifically, our paper looks at a potential intergenerational effect, i.e., the impact of parents' (fathers) exposure to English language education in shaping their children's educational outcomes.

Essentially, we complement emerging evidence that suggests parental aspirations and educational investments play a critical role in shaping learning outcomes (Banerjee and Duflo, 2011, Duflo, 2012, Jensen, 2010), including growing evidence on the role of intergenerational mobility and children's educational performance (Chakraborty et al., 2019, Serneels and Dercon, 2021). The role of aspiration in driving economic behavior is well known. For instance, **?** shows how economic growth is affected by income aspirations. At the individual level, aspirations are known to affect decision making, such as investment in the future (Banerjee and Duflo, 2011, Duflo, 2012).

In India, the knowledge of English is a significant predictor of people's aspirations. For example, a survey conducted in Karnataka showed that most parents considered English skills to be essential for employment opportunities and economic development. In accordance with that belief, the survey findings also show that the parents feel that 'English should be made compulsory at the primary level' (Deccan Herald, 2010). Similar observations and arguments on how English shapes aspirations for social mobility in India is discussed in a 2021 EPW article (EPW Engage, 2021). This widespread aspiration is not merely symbolic but has important implications for decisions on investment in human capital, including preferences for private schooling and additional tutoring and related complementary household-level inputs.

Against this backdrop, we want to study the role that parental exposure to English in school played in their decision to invest in their children's education and, consequently, the educational outcomes of children. The underlying hypothesis is that exposure to English, as a subject in school, affects aspirations, which in turn manifests as higher investment in children's education. Our identification strategy relies on leveraging variation in state-level policy adoption regarding compulsory English education in primary schools across India.

Given the limited rigorous empirical evidence directly linking parental exposure to English during their schooling years and subsequent inter-generational educational outcomes, our study contributes to this literature by exploiting a quasi-experimental setting. As mentioned above, we take advantage of the staggered implementation of compulsory English language education across Indian states from 1979-80 to 2011-12. The variation across states and cohorts allows us to causally estimate how early parental exposure to English education shapes aspirations, educational investments, and, ultimately, children's learning outcomes.

We use nationally representative repeated cross-sectional data from ASER to measure learning outcomes, whereas data on policy variation is constructed by scraping various records from administrative sources and announcements. We find that parental exposure

to compulsory English education in primary school leads to an increase in educational investment. To measure educational investment, we consider four outcomes, grouped into two broad categories: (a) school education and (b) shadow education. Exposure to one additional year of compulsory English increases the probability of children attending private school by 0.6 percentage points, which is a 1.6 percent increase relative to the control mean. We also observe effects on other measures of educational investment: exposure raises the probability of attending a private pre-school by 0.6 percentage points. The probability of attending private tuition increases, but the effect is statistically insignificant. On the intensive margin, we find a 0.019 standard deviation increase in tuition expenditure. These results are robust to using the father's exposure as the main treatment, which accounts for the patrilocal nature of Indian society and migration patterns. However, this surge in investment does not translate into improved learning outcomes; instead, children of parents exposed to compulsory English experience a decline in standardized reading and math scores by 0.013 and 0.008 standard deviations, respectively.

A key mechanism underlying this outcome is a potential reduction in parents' own educational attainment due to compulsory English education, leading to a substitution of educational inputs within the household. As evidence for this, we find that parents exposed to compulsory English were 1.5 percentage points less likely to complete primary school and 1.7 percentage points less likely to complete high school relative to the control group. This decline in parental human capital, an essential input in the education production function, potentially weakened their capacity to directly support their children's learning. Therefore, due to limited own human capital, these parents may have compensated by substituting their direct involvement with external, school/market-based inputs, such as enrolling children in private schools and providing private tutoring. However, as noted in the education literature (Muralidharan and Kremer, 2006, Singh, 2015), the quality of private schools and low-cost tuition centers in rural India can be highly heterogeneous. This heterogeneity, along with parents' potentially limited capacity to effectively evaluate

the quality of these external inputs or complement them due to their own lower educational attainment, could explain why increased financial investments do not translate into better learning outcomes, and may even lead to adverse effects.

This paper makes three main contributions. First, it provides causal evidence of the intergenerational effects of education policy, showing how parental exposure to compulsory English in primary school shaped household aspirations, educational investments, and children's learning outcomes, a relatively less explored dimension in prior studies that largely focus on contemporaneous effects on students' outcomes (Glewwe and Muralidharan, 2016, Chakraborty and Bakshi, 2016, Banerjee et al., 2010). Second, we marginally contribute to the literature on the education production function by identifying an input substitution mechanism: compulsory English reduced parents' own attainment, weakening their human capital and capacity to provide direct educational support, which in turn led them to substitute toward market-based investments such as private schooling and tutoring. This extends work that emphasizes the role of parental aspirations and inputs (Andrabi et al., 2012a, Banerjee and Duflo, 2011, Serneels and Dercon, 2021) by showing that when parental human capital is diminished, substitution into heterogeneous external inputs can systematically weaken learning outcomes. Third, the paper contributes to the literature on the unintended consequences of well-intentioned reforms, demonstrating how a policy designed to raise aspirations and human capital inadvertently reduced parental educational attainment and transmitted adverse effects across generations, resonating with evidence on the unintended impacts of education and governance reforms (Banerjee et al., 2010, Blimpo, 2014).

The rest of the paper is organized as follows. Section 2 provides the institutional context and policy background on the introduction of compulsory English across Indian states. Section 3, describes the data sources and outlines the construction of our key variables in sub-section 3.1 and sets out the empirical strategy in sub-section 3.2. And in Section 5, we discuss robustness checks, including alternative exposure measures and state-level sensitiv-

ity. Section 6 examines spillovers to children's learning outcomes. Section 7 explores the potential mechanism by analyzing the effect of compulsory English on parents' own educational attainment. Finally, Section 8 concludes by highlighting the broader implications of our findings for the design of language and education policies.

2 Context and Policy Background

The role of the English language in Indian education is rooted in its colonial legacy and subsequent state-level policy choices. Beginning in 1757, British colonial rule institutionalized English as a lingua franca across administrative, legal, and educational domains. In 1835, Governor-General Lord William Bentinck, following Macaulay's recommendations, declared English both the official language and the medium of instruction in educational institutions. However, widespread opposition and a shortage of qualified teachers led to a partial reversal of this policy: vernacular languages remained the medium of instruction in primary and secondary schools, while English was retained as a subject and as the medium of instruction in higher education.

The language debate persisted after independence. The Constitution of India recognized both Hindi and English as official languages, with states retaining autonomy to legislate on their internal language policies. Since education is placed on the concurrent list, broad policies are framed at the central level, but the design and implementation of curricula fall largely within the domain of state governments. This division has produced significant variation across states in the timing and manner of introducing English in school education.

State governments have experimented with two main approaches: (i) using English as the medium of instruction, which is largely confined to private schools, and (ii) teaching English as a compulsory subject alongside regional languages, a practice more widely adopted in both public and private schools. The second approach has been shaped not only by regional political ideologies but also by practical supply-side constraints, particularly the shortage

of teachers trained to teach English. Consequently, the grade at which English was first introduced as a compulsory subject varied substantially across states.

We exploit this variation in the introduction of the English language as a compulsory subject in different states across India. We look at the introduction of compulsory years of schooling for all states from 1979 to 2004. This leads to variation in potential years of exposure for children enrolled during this period. For example, while some states introduced English from grade 1, others delayed its introduction to grade 5 or beyond. These differences, rooted in both political ideology and institutional capacity, imply that two children of the same age but from different states could have had markedly different exposure to English in their primary schooling. We leverage this heterogeneity to construct a measure of parental exposure: Panel A of Table 1 reports exposure for fathers' cohorts, and Panel B for mothers' cohorts.[1]

3 Data and Empirical Strategy

3.1 Data

3.1.1 ASER Data

Our data come from the Annual Status of Education Report (ASER), conducted by Pratham, a non-profit organization in India. ASER provides nationally representative repeated cross-sectional data, capturing the educational profiles of children aged 5 to 16 in rural India. The survey is conducted annually and serves as a critical resource for assessing schooling and learning outcomes across the country. For our analysis, we leverage eight rounds of ASER data spanning the years 2009 to 2018 (excluding 2015) and covering 21 states. The data include information on educational investment variables such as school choice, tuition

[1] We restrict attention to potential years of exposure in primary schools.

choice, tuition amount, and pre-school choice. Additionally, the dataset includes test scores for both reading and math, offering insights into children's fundamental learning abilities. ASER's unique household-based testing approach ensures that all children, regardless of their enrollment status, are assessed—allowing us to include currently enrolled students, dropouts, and those who have never attended school.

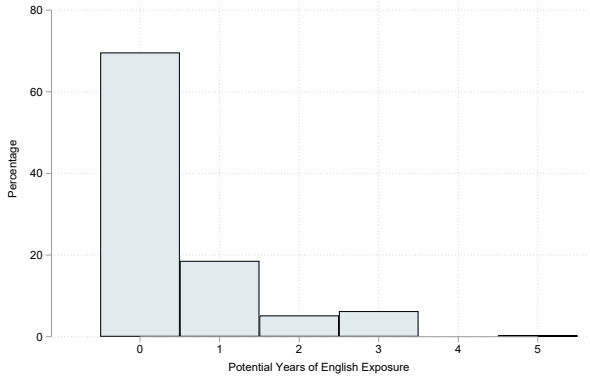
The ASER survey is conducted in October each year and covers a random sample of 20–30 households per village in 20 villages across each of India's approximately 580 rural districts. The reading test evaluates literacy proficiency at four levels: letter recognition, word recognition, reading a short paragraph (grade 1 level), and reading a short story (grade 2 level). The math test comprises four levels: single-digit number recognition, double-digit number recognition, two-digit subtraction with carry-over, and three-digit by one-digit division. Each child is assigned an integer score ranging from 0 to 4 based on the highest level they can comfortably complete.

These data allow us to exploit variation across time, states, and birth cohorts to assess the impact of parental exposure to English on children's educational investments and outcomes. We categorize our main outcomes of interest into two broad groups: educational investment and learning outcomes. Educational investment captures first-order changes in parental investment or behavior towards their children's education, while learning outcomes measure students' academic performance. We measure educational investment using four proxies. First, pre-school choice is a binary variable that takes a value of 1 if the child attends a private kindergarten and 0 if they attend an Anganwadi (a public pre-school). Second, school choice is a binary variable that takes a value of 1 if the child attends a private school and 0 otherwise. Third, tuition choice is also a binary variable indicating whether the child receives private tutoring. Finally, the tuition amount captures the intensive margin of private tutoring and represents a standardized measure of the amount spent on tuition.

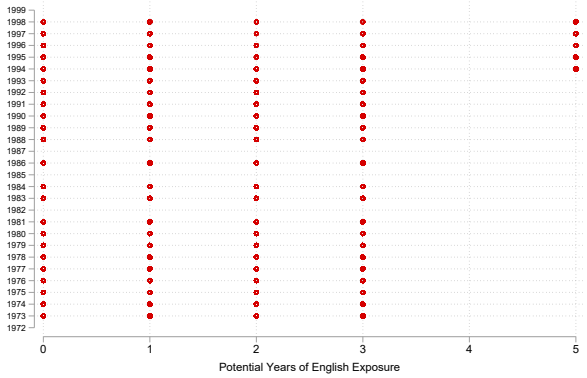
Figure 1: Descriptive on the Variation

Panel 1: Father's Exposure

Panel A: On Average

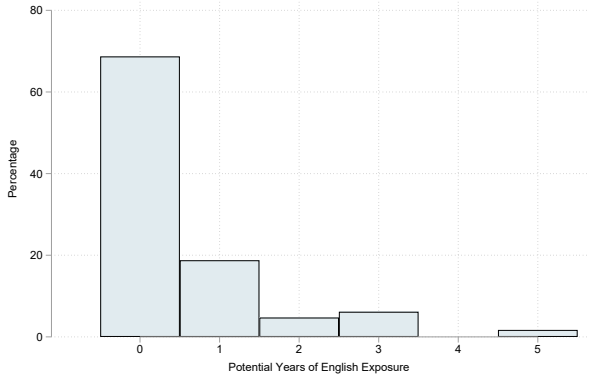


Panel B: Cohort-Wise

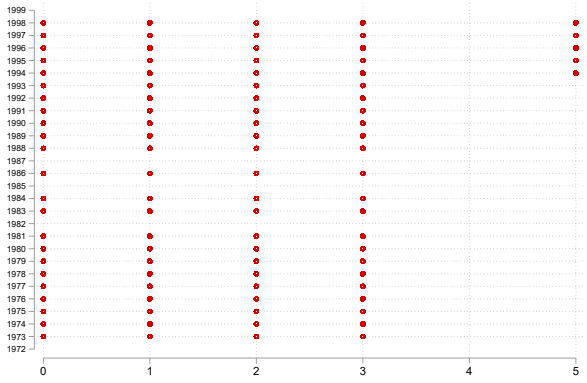


Panel 2: Mother's Exposure

Panel C: On Average

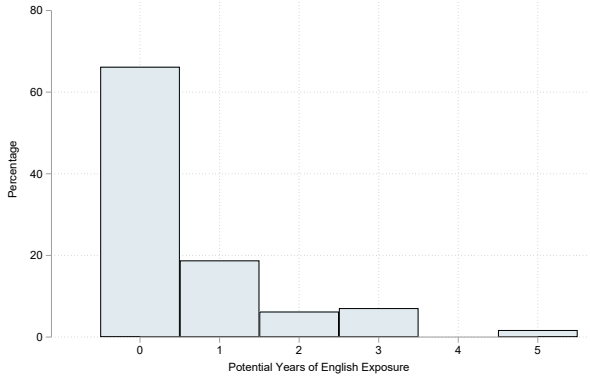


Panel D: Cohort-Wise

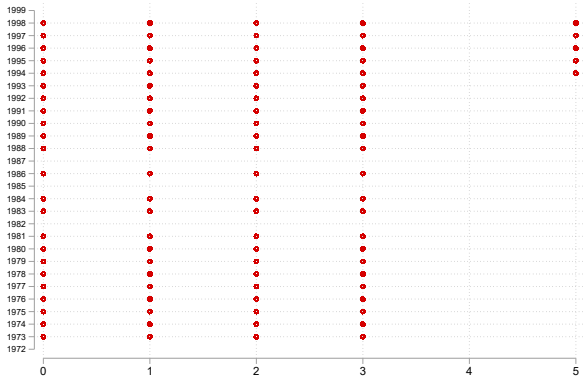


Panel 3: Maximum of Parental Exposure

Panel E: On Average



Panel F: Cohort-Wise



3.1.2 Administrative Data

We incorporate administrative records that track the years in which English was taught as a compulsory subject across states. This dataset spans academic years from 1979–80 to 2011–12, providing a long-term view of English language education policy over three decades. During this period, there has been a general trend toward expanding English instruction in primary education, though substantial regional differences persist, with some states adopting compulsory English much later than others.

We exploit this variation in education policy across states and years to define our treatment. Since ASER data are at the individual level, they allow us to link parents' birth cohorts with state-level policies. Specifically, using each parent's birth year and state of residence, we map individuals to their potential exposure to compulsory English education.[2] For example, an individual born in 1971 would have entered primary school at age six in 1977. Their assigned exposure corresponds to the policy in force in their state that year. Thus, an individual born in Tamil Nadu in 1977 would have had one year of compulsory English exposure, while a peer born in Rajasthan in the same year would have had none.[3]

Compulsory English policies varied widely across states. In some Southern states, English was mandated from grades 1–12, whereas in others it was introduced only from grade 5 onward. One natural way to measure exposure is to count the total number of years for which English was compulsory for each cohort. However, this measure risks endogeneity: individuals who persisted through grades 10 or 12 may differ systematically from those who dropped out earlier. To avoid this concern, we focus exclusively on potential exposure during primary schooling (grades 1–5). Figure **??** shows the distribution of parental exposure

[2]The compulsory years of English schooling refer to public schools. While some private schools may have offered English instruction across all grades, during the 1980–2000 period, the relevant parental cohorts in our data—private schools were not widespread, and only a small share provided English as the medium of instruction, even then (Chakraborty and Bakshi, 2016).

[3]In India, preschooling was not common during this period; children typically entered school directly in grade 1 at age six. We therefore assign policy exposure based on the birth year plus six, which we refer to as the school-entry age. The relevant policy is the one in place in that state at the school-entry year.

to English in primary school. The largest share of individuals had zero years of exposure, and notably, no cohort experienced four years of exposure, reflecting that English was not taught continuously from grade 2 onward in any state during this period.

Finally, whether parental exposure is additive across both parents is theoretically ambiguous. We therefore construct our treatment as the maximum of the father's or mother's exposure to English in primary school.[4] The distribution of this maximum parental exposure is presented in the last row of Figure **??**, where we highlight variation across potential years of exposure and across cohorts.

3.2 Identification Strategy

The main objective of this paper is to estimate the causal effect of potential parental exposure to the English language on children's educational investment. We exploit the fact that different states introduced compulsory English education at different points in time. This implies that a parent's exposure to English in primary school is jointly determined by her year of birth and the English education policy in her state during that year.

To see how birth year and the timing of implementation generate variation in potential program exposure, consider the following example. Take Gujarat, where English was introduced in 1989. A parent born in 1983 was six years old in 1989 and thus was potentially exposed to English education in primary school. By contrast, a parent born in 1978 was already eleven years old in 1989 and therefore not exposed. Children born after 1989 in Gujarat were also exposed to English education.

It would, of course, be problematic if exposure only varied by birth cohort, since younger cohorts are generally more likely to have been exposed to English than older ones. Staggered implementation, however, allows us to account for cohort-specific factors because it

[4] Alternative constructions of the treatment variable are discussed in Section 5.

generates variation in program exposure within cohorts. For instance, consider two children: Rohit from Andhra Pradesh and Seema from Haryana, both born in 1994. Andhra Pradesh had compulsory English education in primary school from 1979, while Haryana had not introduced it at all. Thus, although Rohit and Seema belong to the same cohort, their potential exposure differs because of state policy. This example illustrates how implementation generates variation in years of exposure. Figure **??** presents a scatter plot of potential exposure against birth year, demonstrating that there is considerable variation in our data along both dimensions. In the second column of the figure, we present cohort variation separately for fathers, mothers, and our main exposure variable, defined as the cohort that received the maximum exposure among the parents. We observe two patterns. First, there is significant variation within cohorts. Second, when comparing older cohorts with younger cohorts, the variation in English exposure in primary schools is relatively smaller for the latter.

Our empirical unit is a state × birth-cohort cell. Treatment, which is defined as potential exposure to English in primary school, and control groups are fixed for specific cohorts based on state policy year and school entry age, and do not change over time for each unit. Consequently, our design differs from the canonical staggered-timing panel, where the same unit (e.g., a state) switches from untreated to treated at different dates—a setting in which two-way fixed effects can mix valid and "forbidden" comparisons and generate negative-weight under treatment heterogeneity (Roth et al., 2023). Instead, our comparisons are across fixed treated and fixed control state × cohort cells. Our identifying assumption is a cohort-by-state version of parallel trends: absent the policy, differences in children's educational investments across cohorts would have evolved similarly across states (conditional on the fixed effects we include). We measure treatment intensity as the years of potential exposure.

We estimate the following baseline model, which exploits variation in the years of English

language exposure in primary school, which is

$$y_{itcs} = \alpha + \beta \cdot \text{Exposure}_{itcs} + \phi \cdot \text{Controls}_{itcs} + \delta_t + \delta_c + \delta_s + \gamma_{st} + \varepsilon_i \quad (1)$$

Here, y_{itcs} denotes the educational investment in child i, surveyed in year t, where the parent (father or mother) belongs to birth cohort c and resided in state s during their primary schooling. The variable Exposure_{itcs} measures the potential years of exposure to compulsory English (ranging from 0 to 5 years), and the coefficient β captures the ITT effect of this exposure.

As mentioned earlier, we take the maximum exposure in primary education across both parents. For example, if the father belongs to the 1988 cohort and the mother to the 1994 cohort, and in the father's schooling state, compulsory English lasted for 1 year in 1998, while in the mother's schooling state, it lasted for 3 years in 1993, then we consider the mother's exposure, or vice versa.

Controls include demographic factors such as gender, household size, and maternal schooling. In addition, we include fixed effects for survey year (δ_t), birth cohort (δ_c), and state (δ_s). These parameters capture differences in outcomes associated with survey timing, parental cohort, and state of schooling, respectively.

This empirical specification allows us to control for systematic shocks to outcomes that are correlated with, but not directly attributable to, program exposure. Survey timing is important because educational investment may vary naturally over time, independent of policy exposure. Cohort effects are also relevant since older parental cohorts typically had less exposure to English education compared to younger cohorts, a pattern that is clearly visible in Figure 1, where intensity of exposure is highest for the younger groups. Finally, state differences matter because the timing of compulsory English adoption varied widely across states. Early implementers included both relatively advanced states such as Andhra Pradesh and Karnataka and lagging states such as Uttarakhand. Since early implementers

may also differ systematically in educational outcomes from late implementers, controlling for state fixed effects accounts for these differences.

4 Results

4.1 Main Results

In this section, we look at the impact of parental exposure to the English language in primary school on the educational investment by parents in their children's education. In Table 1, we provide results using equation 1 from the empirical strategy section. We use four outcome variables to measure investment in education of a child: (i) whether a child is enrolled in a public or private school; (ii) whether a child attends a private pre-school (kindergarten) versus an anganwadi; (iii) whether a child attends tuition or not; and (iv) conditional on tuition attendance, how much the household invests (the intensive margin).

Table 1: Effect of Parents' Exposure to English on Educational Investment in Children

	Educational Investment							
	Schooling				Tuitions			
	Early		Public-Private		Extensive		Intensive	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure	0.006***	0.006***	0.006**	0.006**	0.001	0.002	0.014***	0.019***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.004)
Controls	×	✓	×	✓	×	✓	×	✓
Survey Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Father Birth Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Mother Birth Year FE	✓	✓	✓	✓	✓	✓	✓	✓
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	231102	229074	1406253	1392989	1408827	1396276	116065	115603
Outcome Mean (Control)	o.407	0.407	0.374	0.374	0.148	0.148	0.046	0.046

Notes: Standard errors clustered at the state × birth-cohort level are reported in parentheses. Odd-numbered columns present estimates without controls, while even-numbered columns include controls. The dependent variables are grouped into two categories: *Schooling* and *Tuitions*. Schooling outcomes include (i) early school choice, measured by attendance in anganwadi or kindergarten, and (ii) the choice between public and private schools. Tuition outcomes include (i) the extensive margin, indicating whether the child receives any tuition, and (ii) the intensive margin, measured by tuition expenditure. All specifications employ a difference-in-differences design. The sample is drawn from ASER survey rounds 2009–2022 (excluding 2015 and 2017).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In odd-numbered columns, we provide results without controls, and in the even-numbered columns, we add the full set of controls described in the empirical strategy section. We

find that exposure to English education leads to an increase in educational investment in children. Specifically, increasing parental exposure by one year raises the probability of a child attending a private school by 0.6 percentage points (about a 1.6 percent increase relative to the control mean) and the probability of attending a private pre-school by 0.6 percentage points (about a 1.5 percent increase relative to the control mean).

For investment beyond schooling, we look at tuition attendance and the amount spent on tuition. We find a small but positive effect on the probability of attending tuition (0.2 percentage points, or about 1.4 percent relative to the control mean). On the intensive margin, conditional on tuition attendance, parental exposure is associated with a 0.019 standard deviation increase in the amount spent on tuition.

4.2 Dynamic Specification

In this section, we examine the dynamic effect of exposure to the English language in primary school. We estimate the following equation with exposure measured through a vector of four dummy variables denoting 1, 2, 3, and 5 years of exposure, with 0 years serving as the excluded category.

$$y_{itcs} = \alpha + \beta' \cdot \mathbf{Exposure}_{itcs} + \phi \cdot \text{Controls}_{itcs} + \delta_t + \delta_c + \delta_s + \gamma_{st} + \varepsilon_i \tag{2}$$

Here, y_{itcs} denotes the educational investment in the child. The coefficient estimates for the vector of exposure dummies capture the change in educational investment associated with each level of exposure relative to no exposure. The remaining variables are defined as in equation 1.

The results from the dynamic specification suggest that the overall positive effects documented in Table 2 are largely driven by parents who received the maximum exposure to English education in primary school, namely five years. At lower levels of exposure (one or two years), the effects are small or even negative for some outcomes, while three

years of exposure shows stronger effects on tuition choice. This pattern highlights that the relationship between exposure and educational investment is non-linear: a minimum threshold of exposure appears necessary for parents' own schooling experience to translate into meaningful investments in their children's education. Notably, the coefficients for five years of exposure in Table 2 are larger than the average marginal effect reported in Table 1. In relative terms, the impact of five years of exposure is about five to eight times larger than the average effect, underscoring how the linear specification in Table 1 masks the underlying heterogeneity across exposure intensities.

Table 2: Dynamic Effect of Parents' Exposure to English on Educational Investment in Children

	Educational Investment			
	Schooling		Tuitions	
	Early	Public-Private	Extensive	Intensive
	(1)	(2)	(3)	(4)
Exposure Intensity				
One	-0.022***	0.018	-0.020	-0.006
	(0.008)	(0.021)	(0.013)	(0.026)
Two	0.012	-0.029***	0.008	0.034*
	(0.011)	(0.010)	(0.012)	(0.020)
Three	-0.002	0.002	0.002	0.061***
	(0.010)	(0.009)	(0.010)	(0.014)
Five	0.030***	0.050***	0.009	0.014
	(0.009)	(0.013)	(0.009)	(0.049)
Controls	✓	✓	✓	✓
Survey Year FE	✓	✓	✓	✓
Father Birth Year FE	✓	✓	✓	✓
Mother Birth Year FE	✓	✓	✓	✓
State FE	✓	✓	✓	✓
Observations	229074	1392989	1396276	115603
Outcome Mean (Control)	0.4079	0.3746	0.1486	0.0464

Notes: *Notes*: Standard errors clustered at the state × birth-cohort level are reported in parentheses. Outcomes are grouped into *Schooling* (early entry, public vs. private choice) and *Tuitions* (extensive and intensive margins). The structure follows the even-numbered columns of Table 1. The coefficients correspond to OLS estimates of β' in equation 2, where the excluded category is zero years of exposure. One year of exposure corresponds to English introduced in grade 5, two years in grade 4, and so on (no state-mandated English from grade 4, hence four years of exposure is not observed). The sample is drawn from ASER survey rounds 2009–2022 (excluding 2015 and 2017).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

For the intensive margin of tuition, however, we do not find a statistically significant effect

at five years of exposure. This result should be interpreted with caution, as the sample size for this outcome is relatively small. Fewer than 2 percent of households with tuition expenditure data fall into the category of parents exposed for five years. Thus, the lack of significance may be driven more by limited statistical power than by the absence of an effect.

4.3 Heterogeneity Based on Child's Gender

In this section, we examine whether the impact on educational investment differs based on the gender of the child. This is an important dimension to consider because the impact may be driven by boys, as in developing countries, especially India, where son preference is prevalent. This often leads to lower investment in girls. On the other hand, since the treatment is exposure to English education, it could also help change social norms, as the language increases one's exposure to other cultures. Therefore, we investigate whether the impact is driven solely by boys or not. To do this, we conduct a sub-sample analysis: we run the main specification given in equation 1 separately for boys and girls across all measures of educational investment. Table 3 presents the heterogeneity analysis by the gender of the child

Table 3: Heterogeneity Based on Gender: Effect of Parents' Exposure to English on Educational Investment in Children

	Educational Investment							
	Schooling				Tuitions			
	Early		Public-Private		Extensive		Intensive	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expsorue	0.006***	0.005***	0.007***	0.005*	0.002	0.002	0.017***	0.022***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.006)	(0.006)
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Survey Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Father Birth Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Mother Birth Year FE	✓	✓	✓	✓	✓	✓	✓	✓
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	121128	107946	733326	659663	734392	661884	66242	49360
Outcome Mean (Control)	0.436	0.376	0.409	0.335	0.165	0.129	0.067	0.0144

Notes: Standard errors clustered at the state × birth-cohort level are reported in parentheses. Odd-numbered columns present estimates for the boy child, while even-numbered columns present estimates for the girls. The dependent variables are grouped into two categories: *Schooling* and *Tuitions*. Schooling outcomes include (i) early school choice, measured by attendance in anganwadi or kindergarten, and (ii) the choice between public and private schools. Tuition outcomes include (i) the extensive margin, indicating whether the child receives any tuition, and (ii) the intensive margin, measured by tuition expenditure. All specifications employ a difference-in-differences design as given in equation 1. The sample is drawn from ASER survey rounds 2009–2022 (excluding 2015 and 2017).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We find that parents' exposure to English significantly increases the educational investment for both boys and girls. Specifically, the probability of attending a private school rises for both genders; however, the effect is relatively stronger for boys (0.7 p.p) compared to girls (0.5 p.p). For tuition choice at the intensive margin, that is, the amount spent on tuition, the effects are positive and significant for both boys and girls. Interestingly, the magnitude of the effect is larger for girls (2.2 percentage points) than for boys (1.7 percentage points). These results suggest that while son preference is a well-documented phenomenon in India, the benefits of parental exposure to English do not appear to be restricted to boys alone.

5 Robustness

5.1 Alternate Measure of Exposure

In this section, we reproduce our main results using an alternative measure of the exposure variable. Specifically, we restrict the exposure to the father only—that is, the potential number of years a father was exposed to English education in primary school. We adopt this measure because marriages in India are generally patrilocal, and since our focus is on the intergenerational impact of English exposure on children's educational investment, the mother's exposure may be measured with error if she moved to a different location after marriage, which is often the case. To address this potential mismeasurement, we use the father's exposure as the treatment variable in this section, while controlling for the mother's exposure. Apart from this modification, we follow the same estimation strategy, clustering the standard errors at the state-cohort level, where the cohorts are defined by the father's birth year.[5]

Our findings remain consistent with the main results. The magnitude and direction of the coefficients across outcomes are broadly similar, suggesting that the economic significance of the estimates is unchanged. The only exception is school choice, where the coefficient becomes statistically insignificant. Overall, these results reinforce the robustness of our main findings from Table 1.

[5]Although marriages in India are predominantly patrilocal, with around 66% of women over the age of 22 migrating for marriage, the migration distance is not always large. In both rural and urban areas, about two-thirds of women moved more than an hour away from their parental homes Farahzadi (2024). Moreover, most migration occurs within the state: 87.5% of all migrants remain within the same state, with 92.6% of females and 65.6% of males migrating intra-state, while only 7.2% of females migrate across states (see here). These figures are based on the Periodic Labour Force Survey (PLFS) 2022, whereas the cohorts we study (1973–1998) likely experienced even higher rates of within-state migration, given the dominance of agriculture as the primary occupation and the constraints on mobility and information at that time.

Table 4: Effect of Fathers' Exposure to English on Children's Educational Investment

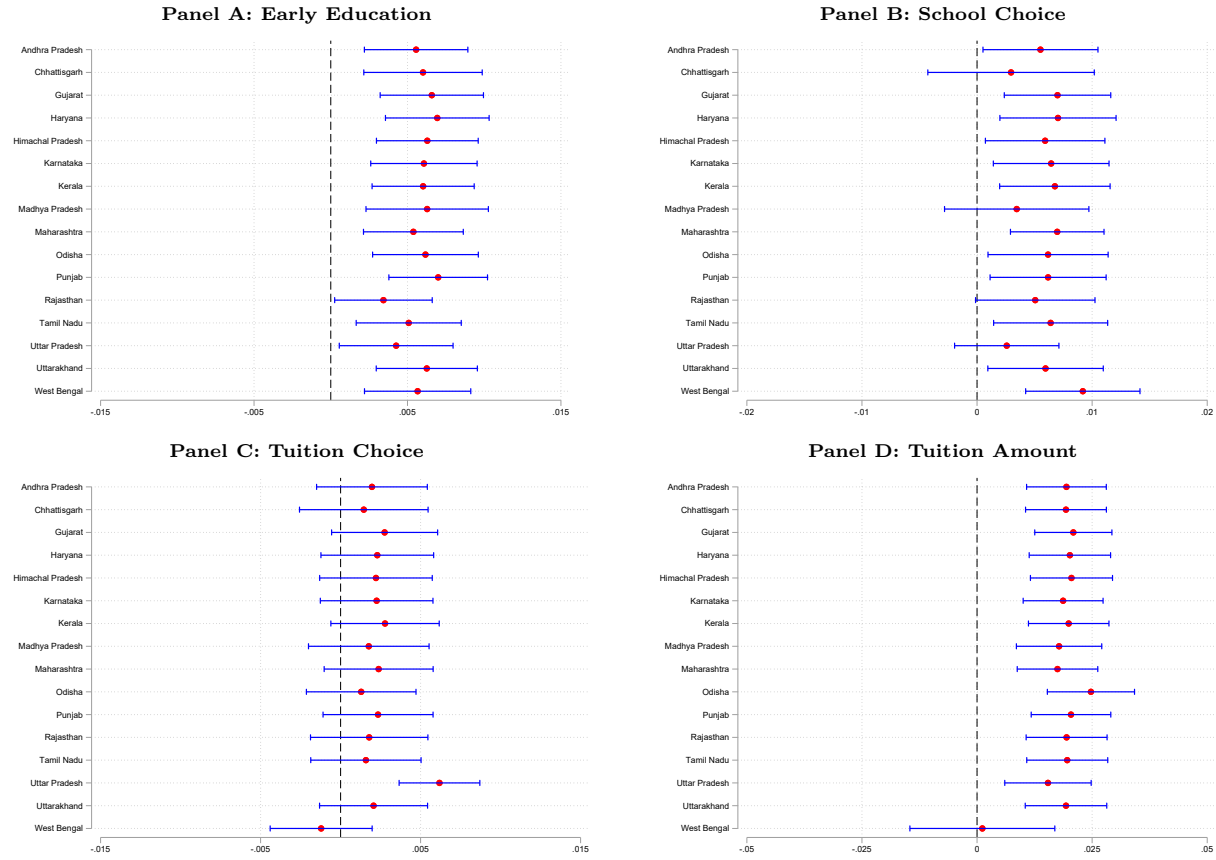
	Educational Investment			
	Schooling		Tuitions	
	Early	Public-Private	Extensive	Intensive
	(1)	(2)	(3)	(4)
Exposure	0.007***	-0.002	0.009***	0.012**
	(0.002)	(0.002)	(0.003)	(0.005)
Controls	✓	✓	✓	✓
Survey Year FE	✓	✓	✓	✓
Father Birth Year FE	✓	✓	✓	✓
Mother Birth Year FE	✓	✓	✓	✓
State FE	✓	✓	✓	✓
Observations	229074	1392989	1396276	115603
Outcome Mean (Control)	0.3863	0.366	0.149	0.044

Notes: Standard errors clustered at the state × birth-cohort level are reported in parentheses. The dependent variables are grouped into two categories: *Schooling* and *Tuitions*. The exposure variable measures the potential years of compulsory English education to which fathers were exposed, with mothers' exposure included as a control. All specifications employ a difference-in-differences design as described in equation 1. The sample is drawn from ASER survey rounds 2009–2022 (excluding 2015 and 2017).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2 State-Specific Sensitivity

Given that our analysis covers a large number of states that arguably possess significant heterogeneity in their respective regional political ideology, historical institutional constraints, and the contemporary socio-economic conditions, we assess whether the inclusion of any particular state drives our results. The effectiveness of policy implementation can vary across states and over time, for instance, states in the Hindi Belt such as Uttar Pradesh and Rajasthan has historically remained a laggard in terms of its exposure to the English language, and therefore, might be unable to initiate the introduction of the English language in the school education system to generate the gains observed in the preceding sections. To rule out the possibility that such differences drive our findings, we conduct a robustness check where we drop one state at a time and estimate a set of equations (with each state excluded once) using the specification given in equation 1.

Figure 2: Robustness to Specific States



In Figure 2, in the Appendix, we present graphical evidence for this leave-out-one analysis, dropping one state at a time from our estimation sample and re-estimating the baseline DID effects for all our outcome variations mentioned in Table 1. The vertical axis of the figure represents the dropped state. Our results remain largely stable across all the measures of educational investment except for tuition choice, and this is consistent with our main results. However, in some cases, the estimate becomes statistically insignificant but is mainly positive for the rest of the outcomes.

6 Spillover on Leaning Outcomes

In the previous sections, we examined the impact of parental exposure to the English language on educational investments made in their children. However, such investments are only meaningful if they translate into improvements in learning outcomes. To test this, we use children's reading and math scores reported in the ASER survey as outcome variables and estimate the same specification given in equation 1. To simplify interpretation, we standardize both reading and math scores. The results are presented in Table 5.

Table 5: Effect of Parents' Exposure to English on Children's Leaning Outcome

	Educational Investment							
	Maximum Exposure				Father's Exposure			
	Reading		Maths		Reading		Maths	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure	-0.022***	-0.013***	-0.016***	-0.008*	-0.015***	-0.008**	-0.008*	-0.001
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Controls	×	✓	×	✓	×	✓	×	✓
Survey Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Father Birth Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Mother Birth Year FE	✓	✓	✓	✓	✓	✓	✓	✓
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1273214	1260284	1270061	1257237	1273214	1260284	1270061	1257237

Notes: Standard errors clustered at the state × birth-cohort level are reported in parentheses. The dependent variables are children's standardized test outcomes in *Reading* and *Maths*, as measured in the ASER surveys. Columns (1)–(4) use the maximum parental exposure (father or mother), while columns (5)–(8) focus on father's exposure with mother's exposure included as a control. Odd-numbered columns report estimates without controls, and even-numbered columns include controls. All specifications employ a difference-in-differences design as described in equation 1. The data is from ASER.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 5, the first four columns use the exposure of the parent with the maximum number of years of compulsory English schooling as the treatment, while the last four columns use the father's exposure as the treatment, while controlling for the mother's exposure. Across both treatment definitions, we consistently find a negative effect of parental exposure to English on children's learning outcomes. Specifically, under the maximum exposure specification, a one-year increase in exposure is associated with a decline of 0.013 SD in reading scores with controls. For math scores, the decline is 0.008 SD with controls. Using the father's exposure as treatment yields very similar results: reading scores fall by 0.015 SD without controls and 0.008 SD with controls, while math scores decline by 0.008 SD

without controls but are statistically insignificant once controls are added.

Prima facie, the negative effects look counterintuitive, as one might expect that parental exposure to English would enhance children's learning outcomes. However, the fact that parental exposure to English increases educational investment without necessarily translating into effective learning is theoretically plausible. We provide three potential mechanisms: first, private schools and low-cost tuition centers in rural India are often of heterogeneous quality, which may create a gap between the quantity of investment and actual skill acquisition (Muralidharan and Kremer, 2006, Singh, 2015); second, parents exposed to English may emphasize English-medium schooling, which could disadvantage children in the short run if instruction quality is poor or if children struggle to acquire foundational skills in a non-native language (Banerjee et al., 2010, Andrabi et al., 2012b). Third, the misalignment between parental aspirations and their own limited ability to provide direct educational support, i.e., substitution of inputs in the education production function, may further widen the gap between investment and learning outcomes.

7 Potential Mechanism

In this section, we try to pinpoint the potential mechanism that can drive the results on educational investment and learning outcomes that we documented in the preceding section. Essentially, we argue that the introduction of English as a compulsory subject at some level could have had an adverse impact on the probability of passing high school for the treatment group parents, since English acted as an additional cognitive load. For the counterfactual, parents who were not exposed to compulsory English, passing high school may have been relatively easier. Therefore, we examine the impact of parental exposure to compulsory English education on the probability of completing both primary education and high school. We estimate these probabilities separately for fathers, mothers, and the parent with the maximum years of exposure to compulsory English.

Table 6 presents the results. The first three columns report effects on the probability of completing primary school, while the last three columns report effects on high school completion. We find that treated parents were less likely to complete high school relative to the control group. Specifically, parents with maximum exposure to compulsory English were 1.5 percentage points less likely to complete primary school and 1.7 percentage points less likely to complete high school. For mothers, we find a 0.5 percentage point reduction in the probability of completing high school, while the effect on primary school completion is a tight null. For fathers, the coefficients are small and statistically indistinguishable from zero at both levels of schooling.

Table 6: Effect of Parents' Exposure to English on Own Education

	Completion of Primary School			Completion of High School		
	Max	Mother	Father	Max	Mother	Father
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure	-0.015***	0.000	0.000	-0.017***	-0.005**	0.004
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Controls	✓	✓	✓	✓	✓	✓
Survey Year FE	✓	✓	✓	✓	✓	✓
Father Birth Year FE	✓	×	✓	✓	×	✓
Mother Birth Year FE	✓	✓	×	✓	✓	×
State FE	✓	✓	✓	✓	✓	✓
Observations	1673538	1669453	1673532	1673538	1669453	1673532
Outcome Mean (Control)	0.436	0.409	0.165	0.376	0.335	0.067

Notes: Standard errors clustered at the state × birth-cohort level are reported in parentheses. The dependent variables are indicators for whether the parent completed *primary school* (columns 1–3) or *high school* (columns 4–6). Columns labeled "Max" use the maximum years of potential exposure between father and mother; columns labeled "Mother" and "Father" use each parent's exposure separately, while controlling for the other parent's exposure. All specifications include controls, survey year fixed effects, state fixed effects, and parental birth year fixed effects as indicated. Estimates are based on a difference-in-differences design as described in equation 1. The sample is drawn from ASER.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The negative effects on parental educational attainment provide an important intergenerational channel. Parents exposed to compulsory English were less likely to complete primary and, more significantly, high school. This implies that they entered adulthood with relatively lower levels of formal human capital. In the education production function framework, this reduction in parental education lowers the quality and quantity of inputs that parents can directly provide to their children's learning. Faced with their own educational limitations, such parents may attempt to substitute their own inputs by investing

in private schooling and tuition for their children. However, these investments typically function as complements rather than substitutes to parental human capital. Parents with lower education may struggle to evaluate quality, support homework, or effectively monitor learning, thereby weakening the returns to private investments. This intergenerational mismatch helps explain why we observe higher spending on children's education by exposed parents, yet weaker actual learning outcomes for their children.

8 Conclusion

Educational policies that focus on increasing access to quality education, such as compulsory English, are generally aimed at expanding human capital and have a positive intergenerational impact. These reforms can encourage households to invest more in their own skill-building and to invest more in their children's education by changing parental aspirations. However, according to he education production function theory, the cognitive outcomes do not depend just on the level of investment but also on the quality of these inputs and complements between these inputs. When reforms unintentionally reduce parents' own attainment, the capacity to provide effective home-based inputs is weakened, potentially distorting the balance between aspirations and actual learning.

Our findings illustrate precisely this trade-off. Using the staggered roll-out of compulsory English across Indian states, we show that parental exposure to the policy increased the likelihood of children attending private schools and pre-schools and raised spending on tutoring. Yet, these higher investments did not translate into stronger skills. Instead, children of exposed parents performed worse in standardized reading and math tests. We provide a potential explanation for this outcome via a substitution mechanism: parents whose own human capital was negative, affected by compulsory English, were less able to support learning directly and compensated through market-based inputs of heterogeneous quality. From a policy standpoint, the results suggest that language reforms must be

carefully sequenced with investments in pedagogy and teacher training

References

Andrabi, T., Das, J., and Khwaja, A. I. (2012a). What did you do all day?: Maternal education and child outcomes. *Journal of Human Resources*, 47(4):873–912.

Andrabi, T., Das, J., and Khwaja, A. I. (2012b). What did you do all day?: maternal education and child outcomes. *Journal of Human Resources*, 47(4):873–912.

ASER, P. E. F. (2017). Annual status of education report (aser) 2016. Technical report, ASER Centre.

Baird, S., McIntosh, C., and Özler, B. (2011). Cash or condition? evidence from a cash transfer experiment. *The Quarterly Journal of Economics*, 126(4):1709–1753.

Banerjee, A. V., Banerji, R., Duflo, E., Glennerster, R., and Khemani, S. (2010). Pitfalls of participatory programs: Evidence from a randomized evaluation in education in india. *American Economic Journal: Economic Policy*, 2(1):1–30.

Banerjee, A. V. and Duflo, E. (2011). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. PublicAffairs.

Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of development economics*, 104:184–198.

Blimpo, M. P. (2014). Team incentives for education in developing countries: A randomized field experiment in benin. *American Economic Journal: Applied Economics*, 6(4):90–109.

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

Chakraborty, T., Schüller, S., and Zimmermann, K. F. (2019). Beyond the average: Eth-

nic capital heterogeneity and intergenerational transmission of education. *Journal of Economic Behavior & Organization*, 163:551–569.

Deccan Herald (2010). 'english should be made compulsory at the primary level'. *Deccan Herald*. "They also felt that English should be made compulsory at the primary level.".

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, 50(4):1051–1079.

EPW Engage (2021). English language education in india: How aspirations for social mobility shape pedagogy. *Economic & Political Weekly Engage*. Published April 23, 2021.

Farahzadi, S. (2024). Migration for marriage. *Available at SSRN 4982438*.

Glewwe, P. and Muralidharan, K. (2016). Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. *Handbook of the Economics of Education*, 5:653–743.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics*, 125(2):515–548.

Kremer, M., Brannen, C., and Glennerster, R. (2013). The challenge of education and learning in the developing world. *Science*, 340(6130):297–300.

Loyalka, P., Liu, C., Song, Y., Yi, H., Huang, X., Wei, J., Zhang, L., Shi, Y., Chu, J., and Rozelle, S. (2013). Can information and counseling help students from poor rural areas go to high school? evidence from china. *Journal of Comparative Economics*, 41(4):1012–1025.

Maluccio, J. A., Hoddinott, J., Behrman, J. R., Martorell, R., Quisumbing, A. R., and Stein, A. D. (2009). The impact of improving nutrition during early childhood on education among guatemalan adults. *The Economic Journal*, 119(537):734–763.

Muralidharan, K. and Kremer, M. (2006). Public and private schools in rural india. *Harvard University, Department of Economics, Cambridge, MA*, 9:10–11.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2):2218–2244.

Serneels, P. and Dercon, S. (2021). Aspirations, poverty, and education: Evidence from india. *Journal of Development Studies*, 57(1):163–183.

Singh, A. (2015). Private school effects in urban and rural india: Panel estimates at primary and secondary school ages. *Journal of Development Economics*, 113:16–32.

Appendix (for publication)

A Tables and Figures

A.1 Descriptive