Impact of Residential Schools on Educational Attainment of Indigenous Women: Evidence from India

Aashay Tripathi

03 November 2024

Abstract

While residential schools in North America have long been dismantled, India continues to expand its own residential school system, with a stated aim of "closing the gap" in education between Indigenous students and their peers. I provide the first causal evidence of the effect of enrollment in a residential school on the educational attainment of Indigenous women in India. Applying triple difference and instrumental variable strategies to a newly constructed dataset, I find that school exposure reduces educational attainment by up to four years. The result is driven by disruptions to family dynamics. Crowding out of day-school options by residential schools, along with mandatory residence at these schools, forces girls to cancel enrollment to fulfill their domestic work obligations.

1 Introduction

From the late 19^{th} to mid-20^{th} century, residential schools in the U.S. and Canada aimed to assimilate Indigenous populations by isolating children from their families, culture, and language. These schools exposed Indigenous children to poor nutrition, overcrowding, and cultural suppression, with many suffering abuse and inadequate healthcare (Meriam (1971), Feir (2016), Feir and Auld (2021)). The Truth and Reconciliation Commission of Canada (2015) (TRC) concluded that these schools sought to destroy Indigenous cultures, and the discovery of unmarked graves heightened calls for justice (Feir and Auld (2021), Jones (2021)).

In contrast, India's 2024 budget allocated 38,800 teachers and INR 63.99 billion ($\approx$ CAD 1.05 billion) for 740 *Eklavya Model Residential Schools* (EMRS), a 150% year-on-year increase, to improve education access for the Scheduled Tribes (STs)[1]. While North American studies have provided valuable insights into the effects of residential schools, limited evidence exists on the educational impact of operational residential schools, making India's case an important contemporary example. In this paper, I leverage a novel individual-level dataset to provide the first causal evidence of the impact of EMRS on the educational attainment of ST women in India.

The EMRS were sanctioned in phases, with central and state governments deciding the villages and the operational dates in every Integrated Tribal Development Project (ITDP)[2] area. These schools, which follow state or central education board curricula, offer competitive admissions with provisions for tribal and first-generation students. Once admitted, the students reside at the school premises for the entire academic year, with the cost of

[1]Scheduled Tribes (STs), often referred to as *Adivasis*, meaning Indigenous peoples or original inhabitants, are groups recognized and protected by the Indian Constitution. They predominantly reside in remote and under-served areas and receive special protections and benefits to address their historical, social and economic marginalization.

[2]These are contiguous large areas of the size of one or more Development Block in which the ST population is 50% or more of the total population.

education and lodging covered by the government. However, estimating the causal effect of residential schools on education is not straightforward, as unobservable factors, including test performance, may influence school attendance and educational attainment. This leads to potentially biased estimates when comparing attendees and non-attendees. I use the staggered opening dates and locations as a source of exogenous variation to estimate the average treatment effect on the treated.

As a first step, I exploit differences in ST status and the presence of a school in the village to show that school-eligible ST women report approximately 1.2 fewer years of education than their ineligible and non-ST counterparts. Since the schools catered to grades VI to XII, I also consider the effect on the probability of completing secondary education. I find an approximate 8 percentage point decline in completing secondary education.

However, villages with and without schools may differ systematically, potentially confounding the results. To address this, I restrict the analysis to villages with schools, controlling for unobservable factors[3] that might affect both school establishment and educational outcomes. Within this restricted sample, I exploit the difference in cohort[4] eligibility and ST status to find that eligible ST cohorts still have 1.4 fewer years of education than their non-ST and ineligible ST peers. Not surprisingly, there is no impact on the probability of completing secondary education for the ST women exposed to the residential schools.

These results rely on the assumption that, in the absence of residential schools, the difference between the educational outcomes of ST and non-ST populations would remain constant over time. However, it is reasonable to believe that this assumption may not hold. For example, ST populations in villages without residential schools may already face lower levels of education due to geographic isolation or cultural preferences for traditional knowledge, creating baseline level differences between ST and non-ST groups. Over time, government

[3]Such as community attitudes toward education or infrastructure quality

[4]Women belong to the school-eligible cohort if they were of school-going age when the school was operational in their village

policies or local economic shocks may disproportionately benefit non-ST populations, and these pre-existing disparities could further widen, leading to different trends in educational outcomes even without the schools.

I employ a triple difference approach to address the potential violation of the common-trend assumption by leveraging cohort eligibility within STs as the third difference. The first difference compares educational outcomes between STs and non-STs, while the second compares educational outcomes between villages with and without a school. The third difference relies on whether an individual's age places them in the school-going cohort when the school was operational in their village. This interaction of ST status, school locations, and cohort eligibility allows me to control for pre-existing differences between ST and non-ST groups that may have otherwise influenced educational outcomes. For instance, if the gap between eligible ST and non-ST cohorts widens beyond what pre-exists in the ineligible cohorts[5], it suggests the schools created this gap. The results show that ST women who lived in villages with operational schools during their school-going years experienced fewer years of education than non-ST and older peers. Specifically, the eligible ST cohort reported 1.25 fewer years of education.

To further check the robustness of the results, I use an instrumental variable approach. As schools were predominantly sanctioned in areas with higher ST populations, which tend to have higher forest cover, as measured by Vegetation Continuous Fields (VCF)[6] values, I use

[5] To identify the causal parameter δ, the difference between the difference of age-eligible ST and non-ST and age-ineligible ST and non-ST should trend similarly in the absence of schools.

$$= \delta + \{(\lambda_1^{ST=1} - \lambda_0^{ST=1})_{eligible} - (\lambda_1^{ST=0} - \lambda_0^{ST=0})_{eligible}\}$$
$$-\{(\lambda_1^{ST=1} - \lambda_0^{ST=1})_{non-eligible} - (\lambda_1^{ST=0} - \lambda_0^{ST=0})_{non-eligible}\}$$

[6] The measure for forest cover comes from Vegetation Continuous Fields (VCF), a Moderate Resolution Imaging Spectroradiometer (MODIS) product that measures tree cover at 250m resolution from 2000 to 2019. VCF is predicted from a machine learning algorithm based on broad-spectrum satellite images and trained with human-categorized data, which can distinguish between crops, plantations and primary forest cover.

this cross-sectional variation in the forest cover as an exogenous variation of school presence. I interact this variation with cohort eligibility to use as an instrument. The argument is that the women will have a higher probability of being exposed to the school if they reside in a village with high VCF *and* were in the school-going age when the school was functional in their village.

The findings show a 4-year decline in education for ST women exposed to residential schools, with average education levels in affected villages also falling by approximately 4 years. The summary statistics illustrate the relative magnitude of the effect. For instance, eligible ST women in villages with an EMRS average 7.364 years of education. If this 4-year decline is applied, it represents more than half of their total educational attainment lost due to exposure, effectively doubling the disadvantage compared to eligible non-ST cohorts, who average 8.907 years. These estimates assume the instrument affects education solely through the likelihood of attending a residential school. The results are robust across model specifications, with district fixed effects controlling for time-invariant district characteristics and state-year fixed effects accounting for state-specific trends. The spatial and temporal variation introduced by forest cover and cohort eligibility remains intact, ensuring the instrument isolates the effect of residential school exposure. Placebo tests and alternative channels are ruled out, reinforcing the validity of the instrumental variable results.

Finally, I provide evidence[7] of the underlying mechanism driving the decline in educational attainment. As students in EMRS must reside at the schools for the entire academic year, families lose access to their labor for domestic or agricultural tasks, increasing the household burden. This disruption of family dynamics and the crowding out of day-schooling options by the residential schools raise the opportunity cost of sending children to EMRS, leading to ST girls cancelling their enrollment due to their domestic work obligations. While these schools aim to provide formal education, the findings highlight how such policies can

[7]Due to data limitations, the estimates are aggregated to the district level.

unintentionally harm the communities they seek to support when they fail to consider the socioeconomic realities and cultural norms of tribal communities.

Historically, residential schools have been criticized for their role in the "cultural genocide" of Indigenous populations, as they systematically stripped away the cultural identity of the children who attended them (Truth and Reconciliation Commission of Canada (2015)). However, the actual outcomes of residential schools have long been debated. Scholars argue that the schools left Indigenous populations culturally stranded, uneducated, and impoverished (Adams (1995); Milloy (2017)), leading to socioeconomic disparities and deep cultural scars not limited to loss of traditional practices and family connections (Bombay, Matheson and Anisman (2014); Bougie and Senécal (2010)). Others suggest that the schools resulted in higher high school graduation rates, higher per capita income, lower poverty rates, a more significant proportion of English-only speakers, and smaller family sizes in the present day. These schools produced a culturally connected and educated elite that later advocated for Indigenous rights (Reyhner and Eder (2017); Szasz (2006); Glenn (2011); Gregg (2018)).

Feir (2016) finds that neither of these positions fully captures the reality; instead, the effects of residential schools are more nuanced, with both economic benefits and cultural losses. Feir (2016) shows that while residential school attendance increased the likelihood of high school graduation and employment, it also significantly reduced cultural ties, such as speaking an Aboriginal language at home or participating in traditional activities. Furthermore, Jones (2021) finds that residential school attendance is linked to lower educational attainment in subsequent generations. This negative association suggests that the trauma and cultural disconnection caused by these schools have disrupted the typical intergenerational transmission of human capital. The findings challenge the narrative that parents' education positively correlates with their children's educational attainment (Black, Devereux and Salvanes (2005); Oreopoulos, Page and Stevens (2006)).

My paper contributes to this literature in two main ways. First, to the best of my knowledge, it is the first to study the educational impacts of residential schools in India, where attendance is voluntary and based on entrance tests, unlike the North American system, where Indigenous populations were forced to attend. Second, the residential school system in India is rapidly expanding, and this study focuses on its immediate effects rather than historical, long-term impacts.

This study contributes to the literature on student aid and educational outcomes(Dynarski (2004); Deming and Dynarski (2009); Dynarski and Scott-Clayton (2013)). Jones (2023) finds that funding cuts reduced completion rates among Indigenous students in the U.S., especially on reserves where post-secondary education was less accessible. Residential schools in India offer monetary relief by covering all educational costs, and STs and non-STs live in the same villages, eliminating *geographic* barriers tied to reserves. However, STs still face lower economic returns due to cultural discrimination, suggesting that their outcomes are driven by socioeconomic marginalization, not geography.

This study also contributes to research on state-backed development missions. EMRS in India can be seen as a state-led "mission" for assimilation, similar to Catholic missions in South America (Valencia Caicedo (2019)) and Christian missionary efforts in Africa (Jedwab, Meier zu Selhausen and Moradi (2022)). While these missions improved human capital, they often imposed cultural assimilation. Likewise, EMRS, while providing education, risks pushing tribal communities into a standardized socio-cultural model, potentially disregarding their unique cultural contexts.

Lastly, this study adds to the large literature on school enrollment and attainment in developing countries, which has focused on school quality (Banerjee et al. (2007)), gender-friendly schools (Kazianga et al. (2013)), incentives (Barrera-Osorio et al. (2011)), infrastructure (Duflo (2001); Breierova and Duflo (2004)), and scholarships (Kremer, Miguel and Thornton (2009)). However, there is limited knowledge on the effect of a government-funded residential school on a minority, more so when that minority is the Indigenous population.

2 Context and Data

2.1. Context

2.1.1. Scheduled Tribes

India is home to the world's second-largest Indigenous population, constituting 8.6% of India's total population, referred to as *Adivasis* or original inhabitants. The 705 ethnic groups classified as Scheduled Tribes (STs) are among the most marginalized communities, granted special legal protections and benefits. According to Mehta (1953), tribes are defined by kinship, common ancestry, region, and culture, with STs being the least integrated into mainstream society. These criteria—region, language, economic life, and cultural practices—determine a tribe's inclusion in the Schedule.

STs have endured centuries of marginalization, struggling with poverty, unemployment, illiteracy, and lack of basic amenities. Traditionally involved in subsistence agriculture or hunting and gathering, their livelihoods have been severely impacted by state oppression and development projects that displace them, perpetuating a cycle of poverty and violence. One significant challenge for STs is limited access to education. Article 46 of the Constitution mandates special attention to the education of ST children, focusing on primary education. However, the formal education system often fails to meet the unique needs of tribal communities.

2.1.2. Eklavya Model Residential Schools (EMRS)

The Eklavya Model Residential School (EMRS) program, launched by the Ministry of Tribal Affairs, aims to provide quality education to Scheduled Tribe (ST) children in remote villages. Sanctioned under Article 275(1) of the Constitution, it is funded by the Ministry of Tribal Affairs. By August 2024, 728 schools were sanctioned, with 409 functional. The program gained momentum with the 2010 target to establish an EMRS in every Integrated Tribal Development Project (ITDP) area and a 2018 directive expanding it to all sub-districts. As of 2022, 462 new schools were sanctioned in 564 sub-districts identified by the 2011 census.

EMRS schools can affiliate with State or Central Boards of Secondary Education and are managed by a Society comprising local, State, and Central Government representatives. Some states use a public-private partnership (PPP) model involving participation from the private sector and NGOs. Admission is competitive, with provisions for primitive tribal groups and first-generation students, ensuring equal seats for girls and boys (up to 60 students per class). The curriculum focuses on English, Hindi, and the student's mother tongue, incorporating tribal culture, tradition, and history. All expenses, including tuition, books, uniforms, food, and boarding, amount to CAD 700 per child annually. The schools aim to foster skill development through extracurricular activities.

However, the emphasis on formal education and achieving parity with non-ST populations may overlook tribal children's vocational interests and socio-cultural connections. Limited annual funding and reserved seats for non-tribal students also raise questions about the program's alignment with its goals, and the lack of mechanisms to maintain family and community ties poses a risk to the children's socio-cultural environment.

2.2. Data

I develop a novel dataset by integrating multiple data sources. The first dataset encompasses all Eklavya Model Residential Schools (EMRS) sanctioned across India till 2024. This includes the complete addresses of these schools, covering the state, district, block, and village locations. I geo-coded these addresses to obtain the exact coordinates of each school. The second dataset is the National Family Health Survey (NFHS), which provides individual-level data on education and health across various states and union territories in India. This dataset includes variables such as years of education, the highest level of education achieved, income status, malnutrition, anemia, hypertension, HIV, and high blood glucose levels. A key feature of this dataset is the use of GPS-based clustering to ensure the anonymity of villages by randomly displacing their locations within specified limits. In rural areas with lower population density, clusters are displaced up to 5 kilometres, with 1% of clusters randomly displaced up to 10 kilometres, while ensuring all points remain within the country, district, and survey region. These GPS coordinates are essential for merging NFHS data with additional datasets.

Using the EMRS coordinates along with the GPS locations from the NFHS clusters, I calculated the Haversine[8] distance – the angular distance between two points on a sphere – between the two pairs of coordinates. This calculation helps me precisely map the schools to the nearest NFHS village clusters. Figure 1 shows the distribution of sanctioned EMRS and the percentage of the Scheduled Tribe (ST) population in NFHS clusters, demonstrating that these schools are predominantly located in regions with higher ST concentrations.

The third dataset comprises Forest Cover data for all villages in India, measured as tree cover at a 250-meter resolution. I geo-coded the $\approx 565,000$ village names to determine their coordinates and then calculated the Haversine distances between the NFHS clusters and the village coordinates. I successfully mapped the villages to their respective NFHS clusters by identifying the shortest distance between each village and cluster. I merge this set with the NFHS data using the cluster ID, adding the average forest cover for each NFHS cluster. Figure 2 illustrates the density of forest cover and the percentage of STs in NFHS clusters, highlighting that ST populations are concentrated in remote, forested areas.

Finally, I combined the NFHS-EMRS-Forest Cover dataset with the NFHS-Covariates data using NFHS cluster IDs. This integrated dataset provides information on education, income status, health indicators, school details, and geographical features. To the best of my knowledge, this is the first such comprehensive dataset. Table I presents the summary statistics for the variables used in the paper's main section.

[8] $D(x,y) = 2\arcsin\sqrt{\sin^2((x_1-y_1)/2) + \cos(x_1), \cos(y_1), \sin^2((x_2-y_2)/2)}$

Table 1: Summary Statistics

Villages with an EMRS						
Variable	**Eligible Cohort (ST)**		**Ineligible Cohort (ST)**		**Eligible Cohort (Non-ST)**	
	Mean	**Std Dev**	**Mean**	**Std Dev**	**Mean**	**Std Dev**
Years of Education	7.364	4.617	3.551	4.551	8.907	4.434
Age	22.591	4.592	39.242	5.366	22.730	4.550
Hindu	0.6765	0.4678	0.6761	0.4680	0.8557	0.3514
Married	0.5364	0.4987	0.8681	0.3384	0.5830	0.4931
Male head of HH	0.8355	0.3707	0.8385	0.3679	0.8580	0.3491
Wealth Index	2.234	1.284	2.253	1.306	3.011	1.399

Villages with no EMRS						
Variable	**Eligible Cohort (ST)**		**Ineligible Cohort (ST)**		**Eligible Cohort (Non-ST)**	
	Mean	**Std Dev**	**Mean**	**Std Dev**	**Mean**	**Std Dev**
Years of Education	7.747	4.505	3.999	4.647	9.184	4.430
Age	22.531	4.667	39.510	5.395	22.396	4.634
Hindu	0.6750	0.4683	0.6555	0.4751	0.8008	0.3993
Married	0.5383	0.4985	0.8852	0.3186	0.5524	0.4972
Male head of HH	0.8447	0.3621	0.8378	0.3685	0.8477	0.3592
Wealth Index	2.429	1.327	2.420	1.335	3.144	1.372

Note: Forest cover is the mean percentage of tree cover detected in the polygon. The wealth index spans from 0=poorest to 5=richest. The variables Hindu, Married, and Male head of HH are indicators that equal 1 if the individual is a Hindu, is married, and the head of that HH is a male. An individual belongs to the eligible cohort if they were of the school-going age when the school was operational in their village.

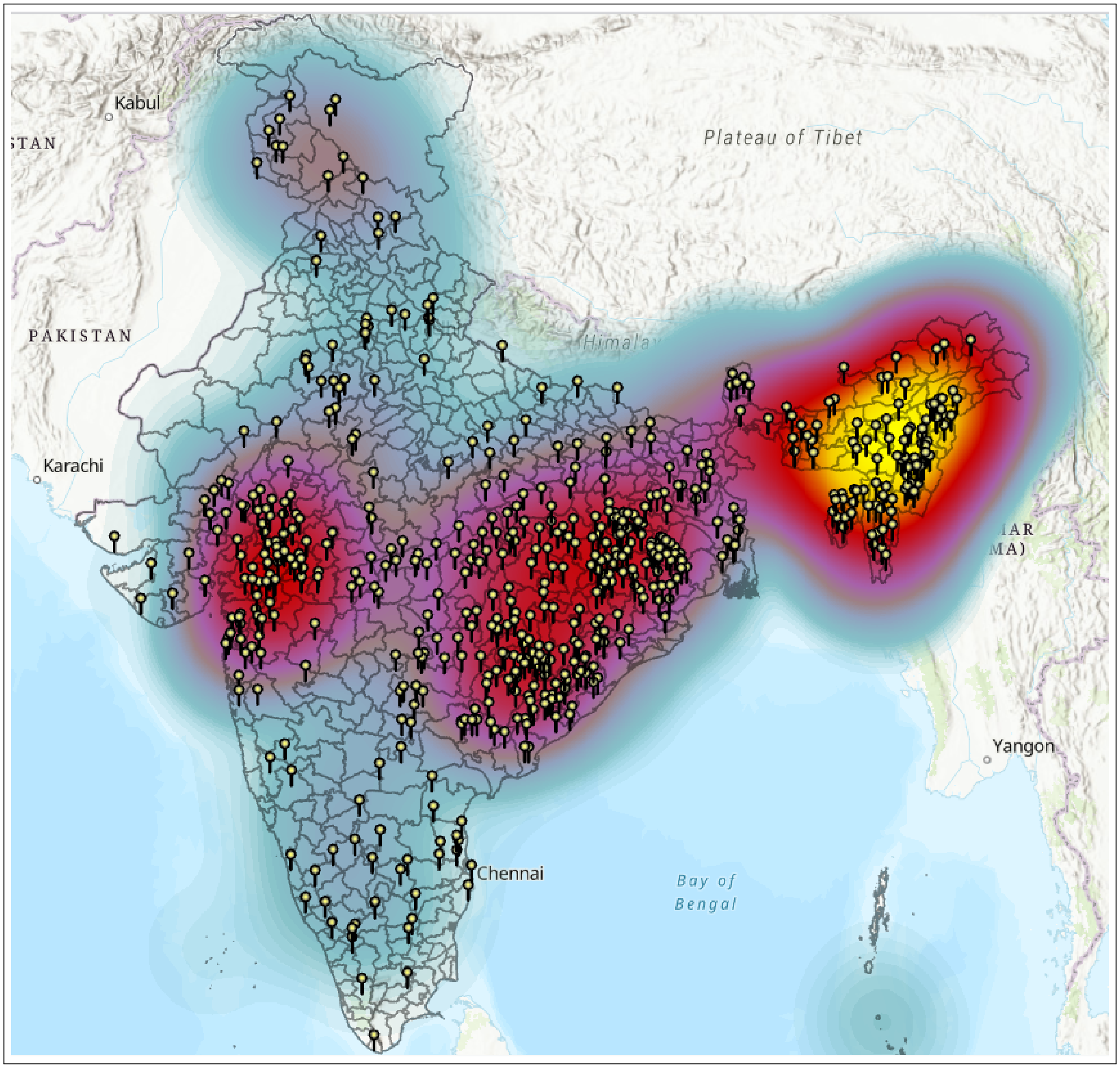

Figure 1: Overlap of EMRS on the percentage of ST population in the NFHS village clusters.

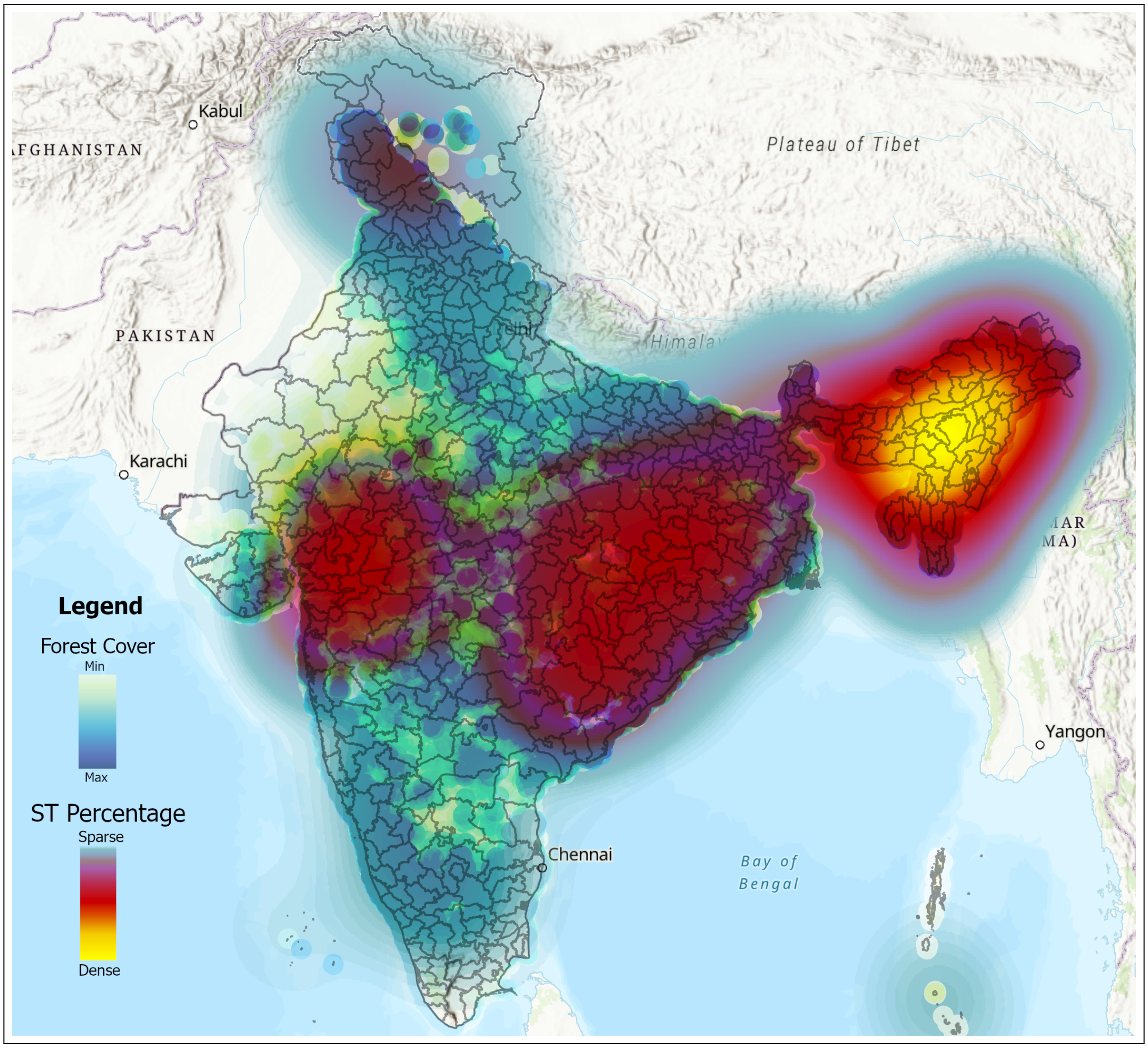

Figure 2: Overlap of Forest Cover on the percentage of ST population in the NFHS village clusters.

3 Identification I: Difference-in-Difference

I capitalize on a policy change related to sanctioning government residential schools for scheduled tribes across villages. Post the 2010 policy modification, every Integrated Tribal Development Agency (ITDA) with at least 50% Scheduled Tribe (ST) population was mandated to establish a residential school. ITDAs encompass at least one district with multiple blocks and villages. Although the selection of ITDAs may follow specific criteria, the choice of villages for sanctioning the schools is arbitrary. Thus, the first difference stems from whether a residential school was in the village. Considering that these schools exclusively cater to scheduled tribes, the second difference arises from the individual's ST status. Even if a village has a residential school, eligible children from non-ST groups would not be exposed to the policy. The third difference emerges from cohort eligibility within the scheduled tribe, depending on whether an individual's age fell within the school-going age bracket when the school became operational in the village. Consequently, I can compare outcomes between eligible women attending these schools and those residing in the same village, belonging to a Scheduled Tribe, but not exposed to the policy due to age constraints.

The baseline difference-in-difference equation is:

$$Years\ Education_{iv} = \beta_0 + \beta_1 ST_i + \beta_2 School_v + \beta_3 ST_i \times School_v + \epsilon_{iv} \tag{1}$$

where $Years\ Education_{iv}$ corresponds to the years of education of individual i, in village v. ST is 1 if the individual belongs to a scheduled tribe, and $School_v$ takes the value 1 if the village has a residential school. ϵ_{iv} is the idiosyncratic error term which satisfies $E(\epsilon_{iv}|i, v) = 0$. The conditional mean function $E(Years\ Education_{iv}|i, v)$ takes on four possible values. The parameter of interest is β_3, which is calculated as the difference between the potential outcomes as below:

$$[E(Years\ Education_{iv}|ST = 1, School_v = 1) - E(Years\ Education_{iv}|ST = 1, School_v = 0)] \tag{2}$$

$$[E(Years\ Education_{iv}|ST = 0, School_v = 1) - E(Years\ Education_{iv}|ST = 0, School_v = 0)] \tag{3}$$

The difference in potential outcomes between a village having a school or not is given by the difference in equation (2). The difference in the potential outcomes for an individual belonging to a scheduled tribe compared to their non-ST counterpart is given by the difference in equation (3). The difference of these differences:

$$\begin{aligned}
&[E(Years\ Education_{iv}|ST = 1, School_v = 1) - E(Years\ Education_{iv}|ST = 1, School_v = 0)] \\
&-[E(Years\ Education_{iv}|ST = 0, School_v = 1) - E(Years\ Education_{iv}|ST = 0, School_v = 0)] \\
&\quad = \beta_3 + [E(\epsilon_{iv}|ST = 1, School_v = 1) - E(\epsilon_{iv}|ST = 1, School_v = 0)] \\
&\quad\quad -[E(\epsilon_{iv}|ST = 0, School_v = 1) - E(\epsilon_{iv}|ST = 0, School_v = 0)]
\end{aligned} \tag{4}$$

To isolate the causal impact, we need some assumptions. One, the error terms are independent of the treatment. This implies that the unobserved factors affecting the outcome are not systematically related to whether an individual is in the treatment or control group. Two, the error terms follow a similar distribution for both groups, ensuring that any random fluctuations or shocks in the error terms affect both groups similarly over time. If the error terms are independent of the treatment and follow a similar distribution in both groups, the term added to β_3 will average out to zero, leaving β_3 as the causal impact on the years of education of an individual belonging to a scheduled tribe and residing in a village with a residential school.

Equation (1) captures the differential impact of residential schools on the educational outcomes of the ST compared to the non-ST. It takes the difference of the difference in potential outcomes of ST between a village having a school or not, and the difference in potential outcomes between ST and non-ST in villages with a school. However, the validity can be a concern if residential schools have within-village spillovers from non-ST to ST students. Or if villages with residential schools have different underlying socioeconomic conditions. This means ST students in villages with schools would have trended differently from ST students in villages without schools, regardless of the presence of residential schools. Therefore, the identification assumes that the difference between the ST and non-ST would have remained constant over time without the treatment. In other words, the ST and non-ST would have experienced the same time trend in education without the schools. To study the plausibility of this assumption empirically, Figure 3 shows the result of an event study. The policy does not affect the cohorts not exposed to it. The coefficients are negative for women born in 1993[9] and later. This shows that the presence of schools impacted the outcomes of women eligible to attend the schools.

Recall the three sources of differences: the presence of a residential school in a village, the tribal status of children, and cohort eligibility within the scheduled tribe, contingent on an individual's age falling within the school-going bracket when the school was operational. I use cohort eligibility to add a third dimension and compare eligible ST students to eligible non-ST and non-eligible ST and non-ST students. Cohort eligibility further randomizes the exposure to treatment based on the individual's age when the school was operational. By including interaction terms that compare the educational outcomes of different groups (ST vs. non-ST, eligible vs. non-eligible) across villages with and without schools, the approach addresses the issue that villages with residential schools might have different socioeconomic characteristics that could affect educational outcomes. The baseline equation is:

$$\begin{aligned} Years\ Education_{igv} = {} & \beta_0 + \beta_1 ST_i + \beta_2 Eligible_g + \beta_3 School_v + \beta_4 ST_i \times Eligible_g \\ & + \beta_5 Eligible_g \times School_v + \beta_6 School_v \times ST_i \\ & + \beta_7 ST_i \times Eligible_g \times School_v + \epsilon_{igv} \end{aligned} \tag{5}$$

[9]An individual belongs to a school-going age cohort if their age in 2021, the last year of the data, is 29 or less. This means the age of the individual exposed to the earliest school would be 18 years when it was functional. So, 1993 is the earliest birth year exposed to the schools.

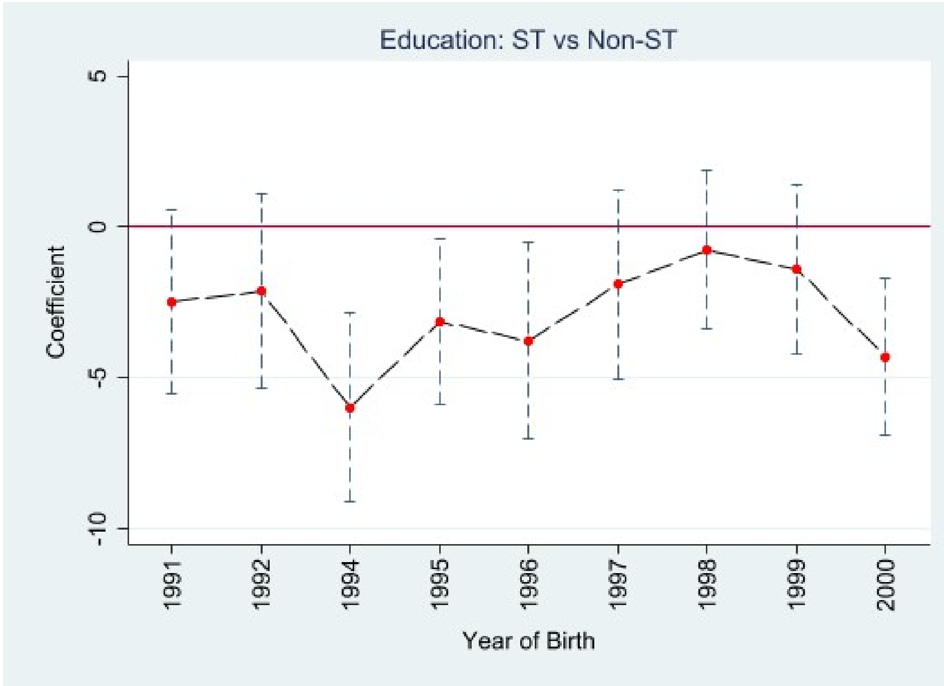

Figure 3: The impact on years of education is negative for women born in 1993 or later. Anyone born in 1993 would be 18 when the earliest school was sanctioned. They would have been exposed for at least 1 year. 1993 is the base year

where $Years\ Education_{igv}$ corresponds to the years of education for individual i, belonging to cohort $g \in \{School - Age, Non - School - Age\}$, in village v. $Eligible_g$ is 1 if the individual belongs to the school-age cohort, ST_i is 1 if the individual belongs to a scheduled tribe, and $School_v$ is 1 if the village has a school. β_7 captures the effect of residential schools on the years of education of eligible ST students compared to other students in the same village or villages without schools.

4 Results I: Difference-in-Difference

Table 1(a) shows the impact of the schools on educational attainment by comparing ST and non-ST women of school-going age across villages. The coefficient for the interaction term $ST \times School$ in column (6) is -1.199, indicating that eligible ST women who were exposed to the school experienced a reduction of approximately 1.2 years of education compared to their eligible non-ST peers. Since EMRS is for grades VI to XII, I additionally consider the impact of the school on the probability of completing secondary education. Table 1(b) presents the results. The eligible ST cohort reports a decline in completing secondary education by almost 8 percentage points.

Table 2(a) presents the results from the difference-in-difference analysis for villages with a residential school. The analysis reveals a significant and adverse effect of EMRS exposure on the educational outcomes of these women. Specifically, the coefficient for the interaction term $ST \times Eligible$ is -1.424, which implies that ST women in villages with a school experience a reduction of approximately 1.4 education years compared to their non-ST counterparts in villages with a school. Table 2(b) presents the results for completing secondary education. Although the results are insignificant, ST women report a decline of 7.2 percentage points in completing secondary education.

Table 1(a): Education: Eligible ST vs Eligible Non-ST

Dependent Variable:	**Years of Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
$ST \times School$	-1.241**	-1.134**	-1.253**	-1.280**	-1.120**	**-1.199*** **
	(0.5860)	(0.0623)	(0.5419)	(0.5526)	(0.4627)	(0.4489)
Observations	2,942	2,942	2,941	2,941	2,941	2,941
R-squared	0.0547	0.1154	0.1675	0.2891	0.3338	0.4175
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is years of education. The estimation controls for state-year and district-fixed effects. Controls include religion, marital status, gender of the household head, and years of education. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 3(a) presents the results from the triple difference estimation. Comparing the

Table 1(b): Education: Eligible ST vs Eligible Non-ST

Dependent Variable:	**Completing Secondary Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
$ST \times School$	-0.0322	-0.0321	-0.0448	-0.0711	-0.0563	**-0.0785***
	(0.0454)	(0.0433)	(0.0396)	(0.0486)	(0.0366)	(0.0448)
Observations	2,941	2,941	2,941	2,941	2,941	2,941
R-squared	0.0214	0.0435	0.0886	0.1764	0.1965	0.2670
Year FE	-	✓	-	-	-	-
State FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is an indicator which equals 1 if the individual completes secondary education. Estimation controls for state-year and district-fixed effects. Controls include religion, marital status, gender of the household head, and income status. Standard errors are clustered at the village level and shown in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

outcomes of ST women who were of school-going age when the school was operational in their village with those who were either too old or too young at the time of the school's opening and resided in a village without a school helps to further isolate the causal impact of the school on the educational outcomes of STs. The coefficient for the triple interaction term $ST \times School \times Eligible$ is -1.248, indicating a significant and negative impact of the EMRS policy on the educational attainment of the targeted ST population. Specifically, this result suggests that ST women eligible to attend an EMRS during their school years completed approximately 1.24 fewer years of education than their non-ST and non-eligible counterparts in villages without the school. Table 3(b) reports the results for completing secondary education. As before, although insignificant, ST women eligible to attend an EMRS during their school years report a 3.4 percentage point decline in completing secondary education.

The magnitude of this effect is particularly concerning, given that the primary objective of the EMRS policy was to enhance educational opportunities for marginalized tribal populations. The robustness of this finding across various model specifications—including those that control for state-year and district-fixed effects—indicates that a residential school in a village is strongly associated with diminished educational attainment among the ST population. Contrary to its intended purpose, the significant decrease in years of education among those directly exposed to the policy indicates a profound disconnection between the policy's objectives and actual outcomes. The negative educational impact mirrors the educational disadvantages in the Canadian and North American context. Focusing on the long-term impacts, Feir (2016) and Jones (2021) also show that Indigenous children who attended residential schools often emerged with lower educational attainment than their peers who did not attend such schools. This consistent outcome across different cultural and policy contexts illustrates a recurring pattern: residential school systems, whether in North America or India, by removing students from their cultural and social environments and imposing a one-size-fits-all approach to education, are fundamentally misaligned with the needs and values of Indigenous populations, resulting in significant educational setbacks for Indigenous populations.

Table 2(a): Education: ST vs Non-ST (Villages with a school)

Dependent Variable:	**Years of Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
ST × School	0.6466	0.4017	-0.3521	-0.5625	-0.9675**	**-1.424*** **
	(0.4940)	(0.4885)	(0.4817)	(0.5382)	(0.4601)	(0.5181)
Observations	1,926	1,926	1,913	1,913	1,350	1,350
R-squared	0.1393	0.2124	0.3113	0.4767	0.4972	0.5947
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is years of education. The estimation controls for state-year and district-fixed effects. Controls include religion, marital status, gender of the household head, and years of education. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 2(b): Completing Secondary Education: Eligible ST vs Eligible Non-ST

Dependent Variable:	**Completing Secondary Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
ST × Eligible	-0.0414	-0.0345	-0.0517	-0.0500	-0.0633	-0.0726
	(0.0225)	(0.0404)	(0.0420)	(0.0451)	(0.0410)	(0.0461)
Observations	1,826	1,826	1,913	1,913	1,350	1,350
R-squared	0.0404	0.0717	0.1657	0.3236	0.2600	0.3494
Year FE	-	✓	-	-	-	-
State FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is an indicator which equals 1 if the individual completes secondary education. Estimation controls for state-year and district-fixed effects. Controls include religion, marital status, gender of the household head, and income status. Standard errors are clustered at the village level and shown in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3(a): Education: Triple Difference

Dependent Variable:	**Years of Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
ST × School × Eligible	-0.0950	-0.1462	-0.3913	-0.6477	-0.9559*	**-1.248****
	(0.6006)	(0.5816)	(0.5493)	(0.5643)	(0.5367)	(0.5409)
Observations	8,310	8,310	8,305	8,305	7,042	7,042
R-squared	0.1239	0.1994	0.2488	0.3287	0.4110	0.4658
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is the total years of education. Estimation controls for state-year and district-fixed effects. Controls include religion, marital status, gender of the household head, and income status. Standard errors are clustered at the village level and shown in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3(b): Completing Secondary Education: Triple Difference

Dependent Variable:	**Completing Secondary Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
$ST \times Eligible \times School$	0.0320	0.0317	0.0117	0.0059	-0.0230	-0.0341
	(0.0458)	(0.0448)	(0.0403)	(0.0429)	(0.0399)	(0.0424)
Observations	8,310	8,310	8,305	8,305	7,042	7,042
R-squared	0.0404	0.0573	0.1058	0.1608	0.1800	0.2144
Year FE	-	✓	-	-	-	-
State FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is an indicator which equals 1 if the individual completes secondary education. Estimation controls for state-year and district-fixed effects. Controls include religion, marital status, gender of the household head, and income status. Standard errors are clustered at the village level and shown in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5 Identification II: Instrumental Variable

There are a few concerns while studying the impact of residential schools on educational outcomes. Consider the causal equation:

$$Years\ Education_{iv} = \gamma_0 + \gamma_1 School\ Exposure_v + \gamma_2 X_{iv} + \epsilon_{iv}$$

where $Years\ Education_{iv}$ correspond to the years of education of individual i in village v, $School\ Exposure_v$ is 1 if the village v has a school, and X_{iv} are the controls. One concern is the potential for reverse causality. The level of education might be influenced by the presence of schools in a village, and the sanctioning of additional schools in that village is contingent on the average years of education within the community. Another concern is the selection on unobservables, meaning some unobserved factors affect both the selection into treatment and the potential outcomes. In the present context, for example, unobservable factors might include parental motivation, community attitudes toward education, or the inherent abilities of the students[10]. So, we have reasons to abandon the assumptions $[Y_0,\ Y_1 \perp D \mid X]$ and $\mathbb{E}[u \mid X, D] = 0$, where Y_0, Y_1 are the potential outcomes, D is the treatment indicator, X are the observable characteristics and u are the unobservables.

To reduce the impact of unobservables[11], instead of using a broad instrument like the forest cover in isolation, I use the interaction of cohort eligibility and the forest cover as an instrumental variable for the presence of a school and, consequently, school exposure. Residential schools are specifically designated for STs who predominantly inhabit secluded, hilly, and forested regions. Consequently, schools are strategically established in remote and rugged areas inhabited mainly by the STs, exposing them to the treatment. If such school placement is followed, it ensures that the density of forest cover captures the exogenous variation in school exposure.

I define the instrument as $Z_{ivt} = Forest\ Cover_v \times Post_{it}$; where $Post_{it}$ is an indicator that equals 1 if the ST individual was in the school going age when the school was operational in the village, and $Forest\ Cover_v$ is an indicator that equals 1 if the forest cover of the village is more than the average. I consider the following setup with Z_{ivt} as the instrument:

Causal Equation:

$$Years\ Education_{ivt} = \delta_0 + \delta_1 School\ Exposure_{ivt} + \epsilon_{ivt}$$

First Stage:

$$School\ Exposure_{ivt} = \gamma_0 + \gamma_1 Post_{it} + \gamma_2 Forest\ Cover_v + \gamma_3 Z_{ivt} + \epsilon_{ivt}$$

[10] Since the students were admitted based on their performance in the entrance tests.

[11] I provide further evidence in section 8.

Reduced Form:

$$Years\ Education_{ivt} = \beta_0 + \beta_1 Post_{it} + \beta_2 Forest\ Cover_v + \beta_3 Z_{ivt} + \epsilon_{ivt}$$

The parameter of interest, δ_1, can be calculated as the ratio of reduced form and first stage coefficients. The first-stage and reduced form coefficients can be calculated by the first-differenced regressions of $Y_{iv1} - Y_{iv0}$ and $S_{iv1} - S_{iv0}$ on Z_{ivt}

$$\gamma_{IV} = \frac{E[Y_{iv1} - Y_{iv0} \mid Z_{ivt} = 1] - E[Y_{iv1} - Y_{iv0} \mid Z_{ivt} = 0]}{E[S_{iv1} - S_{iv0} \mid Z_{ivt} = 1] - E[S_{iv1} - S_{iv0} \mid Z_{ivt} = 0]}. \tag{6}$$

where Y_{ivt} is the years of education, S_{ivt} is the exposure to the schools, and Z_{ivt} is the instrument. Figure 3 plots the sanctioned EMRS schools on the density of forest cover. Regions with thicker covers received more schools, confirming that the forest cover is correlated with the causal endogenous variable of interest. For the instrument to be valid, it should satisfy the exclusion restriction:

$$Cov(Z_{ivt},\ \epsilon_{ivt}) = 0$$

implying that the instrument is uncorrelated to any other determinants of the years of education.

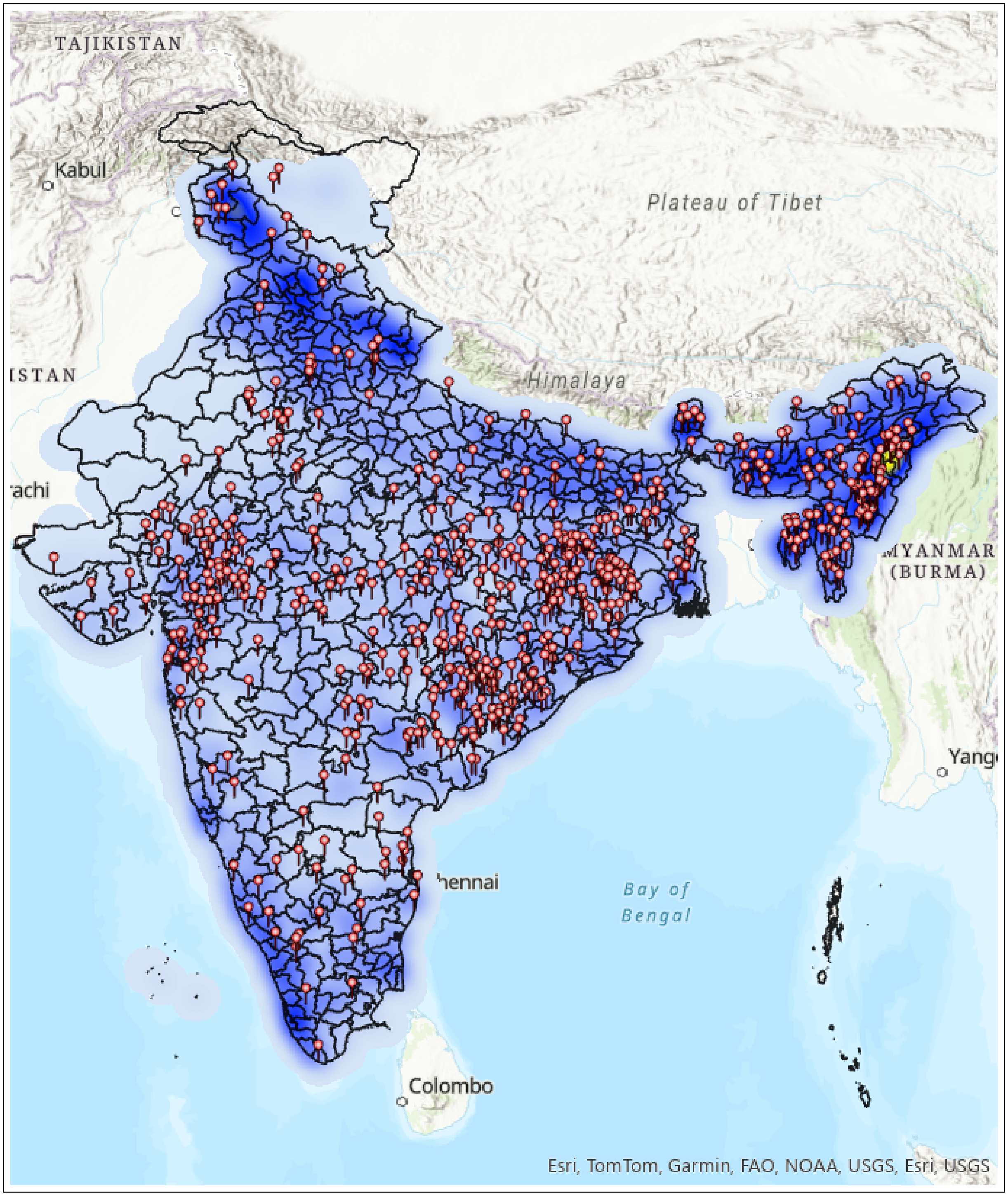

Figure 4: Overlap of EMRS on the Forest Cover of villages across India.

While forest cover in isolation may affect factors that also impact the years of education, it can be argued that the interaction of forest cover with cohort eligibility ensures that the instrument affects the years of education only through the residential schools. In section 8, I consider other potential channels through which forest cover and eligibility might affect the outcomes and provide evidence against their effect. Finally, the assumption of parallel trends should be satisfied. The growth paths of outcomes and treatment are independent of the assignment of the instrument. This means that the way the instrument influences who gets treated should not affect the outcome trends of the groups before the treatment is applied. So, the difference in years of education of the two groups in our setting - $Eligible,\ Not - Eligible$ - should not be affected in regions with less than average forest cover.

6 Results II: Instrumental Variable

The first stage of tables 4(a) and 4(b) presents the relationship between the interaction of forest cover and cohort eligibility as an instrument for school exposure. The interaction term $Forest\ Cover\ \times\ Post$ coefficients are consistently positive and statistically significant across specifications. Column (6) is the column of interest, as it includes the controls and the state-year and district-fixed effects. These estimates suggest that eligible cohorts in districts with higher forest cover are more likely to be exposed to the residential schools. The Cragg-Donald Wald F statistic is greater than 10, implying that the instrument used is not weak.

The IV estimates in Table 4(a) highlight the impact of school exposure on years of education. The coefficients on Years of Education are negative across all models. Including controls, state-year, and district fixed effects does not substantially alter the magnitude of these coefficients, suggesting a robust negative impact of school exposure on education. Focusing on column (6), the estimate of -4.145 suggests that increased exposure to schools in forested districts significantly reduces the years of education of the affected cohorts by almost 4 years. Table 4(b) shows the results for completing secondary education. Although statistically insignificant, there is a signal that the probability of completing secondary education declined by almost 12 percentage points. This supplements the evidence on declining years of education and confirms that due to the decline in years of education, ST women exposed to the school were unable to complete their secondary education.

The results in Tables 4(a) and 4(b) consider all districts. I narrowed the sample and focused only on districts with at least one residential school. By restricting the analysis to districts with residential schools, I aim to eliminate the noise that might come from including all districts, as the variation in outcomes could be driven by factors unrelated to the presence of residential schools. Furthermore, if students in non-school districts are fundamentally different from those in school districts, different characteristics would influence the establishment of schools and educational outcomes. By focusing on districts with residential schools, I reduce the selection bias that might occur. Table 5(a) presents these results. As before, the estimates for the instrument are positive and statistically significant across specifications. The Cragg-Donald Wald F statistic greater than 10 again ensures that the instrument is not weak. The IV estimates in column (6) suggest approximately 4 fewer years of education for the eligible cohort in heavily forested regions. Table 5(b) shows the results for completing secondary education. Again, although statistically insignificant, the probability of completing secondary education declined by almost 12 percentage points. This also shows that due to the decline in years of education, ST women exposed to the school were unable to complete their secondary education.

Lastly, I consider the impact of these schools on the average village-level education. These schools were sanctioned at the village level. If the village-level factors fed into the sanctioning decision, it is safe to assume that the policy's impact evaluation would also be based on village-level statistics. Additionally, the presence of a residential school might not only affect those directly exposed but also indirectly influence peers, families, or community

Table 4(a): Education: All Districts

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.7550***	-0.0277	-0.0251	-0.0463	-0.0175	-0.0682**
	(0.0463)	(0.0499)	(0.0496)	(0.0325)	(0.0539)	(0.0325)
Forest Cover	0.7540***	-0.0312	-0.0289	0.0184	-0.0467	-0.0018
	(0.0537)	(0.0642)	(0.0634)	(0.0767)	(0.0719)	(0.0871)
Forest Cover × Post	-0.6977***	0.0694	0.0638	0.1488***	0.1062	0.1753***
	(0.0747)	(0.0727)	(0.0721)	(0.0548)	(0.0789)	(0.0565)
Cragg-Donald Wald F statistic	1651.402	2.717	2.196	29.160	6.239	36.038
IV Estimation	**Dep. Variable: Years of Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
School	6.035***	-4.036	-2.434	-4.907	-5.051	**-4.145****
	(0.3561)	(9.499)	(9.598)	(3.264)	(4.235)	(2.084)
Observations	4,231	4,231	4,209	4,209	3,604	3,604
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is years of education. The interaction of the forest cover and cohort eligibility is the instrument. The estimation controls for state-year and district-fixed effects. Controls include income status, marital status, religion, the gender of the household head, and ST status. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 4(b): Completing Secondary Education: All Districts

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.7550***	-0.0277	-0.0251	-0.0463	-0.0175	-0.0682**
	(0.0463)	(0.0499)	(0.0496)	(0.0325)	(0.0539)	(0.0325)
Forest Cover	0.7540***	-0.0312	-0.0289	0.0184	-0.0467	0.0018
	(0.0537)	(0.0642)	(0.0634)	(0.0767)	(0.0719)	(0.0871)
Forest Cover × Post	-0.6977***	0.0694	0.0638	0.1488***	0.1062	0.1753***
	(0.0747)	(0.0727)	(0.0721)	(0.0548)	(0.0789)	(0.0565)
Cragg-Donald Wald F statistic	1651.402	2.717	2.196	29.160	6.239	36.038
IV Estimation	**Dep. Variable: Completing Secondary Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
School	0.0748***	0.6637	0.7807	0.1054	0.3045	0.1219
	(0.0085)	(0.7379)	(0.9178)	(0.1127)	(0.2137)	(0.0906)
Observations	4,231	4,231	4,209	4,209	3,604	3,604
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is an indicator which equals 1 if the individual completes secondary education. The interaction of the forest cover and cohort eligibility is the instrument. The estimation controls for state-year and district-fixed effects. Controls include income status, marital status, religion, the gender of the household head, and ST status. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 5(a): Education: Districts with a Residential School

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.8059***	-0.0746	-0.0739*	-0.0478	-0.0911*	-0.0736**
	(0.0449)	(0.0464)	(0.0457)	(0.0339)	(0.0486)	(0.0342)
Forest Cover	0.8084***	-0.0747	-0.0738	0.0165	-0.1088*	-0.0068
	(0.0502)	(0.0599)	(0.0592)	(0.0767)	(0.0656)	(0.0874)
Forest Cover × Post	-0.7489***	0.1355**	0.1359**	0.1522***	0.1967***	0.1840***
	(0.0677)	(0.0662)	(0.0656)	(0.0564)	(0.0711)	(0.0586)
Cragg-Donald Wald F statistic	1751.413	12.760	12.286	26.949	20.492	33.580
IV Estimation			**Dep. Variable: Years of Education**			
	(1)	(2)	(3)	(4)	(5)	(6)
School	5.675***	11.924	12.945	-5.035	1.762	**-4.177****
	(0.3287)	(7.644)	(7.989)	(3.286)	(2.330)	(2.051)
Observations	3,860	3,862	3,844	3,844	3,238	3,238
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is years of education. The interaction of the forest cover and cohort eligibility is the instrument. The estimation controls for state-year and district-fixed effects. Controls include income status, marital status, religion, the gender of the household head, and ST status. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 5(b): Completing Secondary Education: Districts with a Residential School

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.8059***	-0.0772*	-0.0764*	-0.0487	-0.0950**	-0.0750
	(0.0451)	(0.0464)	(0.0455)	(0.0341)	(0.0477)	(0.0344)
Forest Cover	0.8074***	-0.0776	-0.0771	0.0115	-0.1130*	-0.0092
	(0.0505)	(0.0596)	(0.0590)	(0.0765)	(0.0653)	(0.0873)
Forest Cover × Post	-0.7458***	0.1418**	0.1426**	0.1560***	0.2063**	0.1871***
	(0.0678)	(0.0659)	(0.0654)	(0.0562)	(0.0706)	(0.0584)
Cragg-Donald Wald F statistic	1751.467	13.916	13.451	27.127	22.536	34.365
IV Estimation		**Dep. Variable: Completing Secondary Education**				
	(1)	(2)	(3)	(4)	(5)	(6)
School	0.0714***	0.7158*	1.914	0.0986	0.2837**	0.1197
	(0.0084)	(0.3876)	(0.3966)	(0.1124)	(0.1411)	(0.0897)
Observations	3,858	3,858	3,840	3,840	3,238	3,238
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is an indicator which equals 1 if the individual completes secondary education. The interaction of the forest cover and cohort eligibility is the instrument. The estimation controls for state-year and district-fixed effects. Controls include income status, marital status, religion, the gender of the household head, and ST status. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 6(a): Average Education: Village Level

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.7550***	-0.0104	-0.0277	-0.0503	-0.0251	-0.0463
	(0.0463)	(0.0546)	(0.0499)	(0.0332)	(0.0496)	(0.0325)
Forest Cover	0.7540***	-0.0092	-0.0312	0.0182	-0.0289	0.0184
	(0.0537)	(0.0698)	(0.0642)	(0.0772)	(0.0634)	(0.0767)
Forest Cover × Post	-0.6977***	0.0660	0.0694	0.1504***	0.0638	0.1488***
	(0.0747)	(0.0800)	(0.0727)	(0.0554)	(0.0721)	(0.0648)
Cragg-Donald Wald F statistic	1651.402	2.793	2.717	31.651	2.196	29.160
IV Estimation	**Dep. Variable: Average Years of Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
School	7.654***	-3.392	-2.994	-3.835	-3.519	-3.824
	(0.3895)	(7.430)	(7.300)	(2.450)	(8.096)	(2.438)
Observations	4,231	4,231	4,231	4,231	4,209	4,209
State FE	-	✓	✓	-	-	-
Year FE	-	-	✓	-	-	-
State-Year FE	-	-	-	-	✓	✓
District FE	-	-	-	✓	-	✓

Notes: The dependent variable is the average years of education. The interaction of forest cover and cohort eligibility is the instrument. The estimation controls for state-year and district-fixed effects. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 6(b): Average Completion of Secondary Education: Village Level

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.7550***	-0.0104	-0.0277	-0.0503	-0.0251	-0.0463
	(0.0463)	(0.0546)	(0.0499)	(0.0332)	(0.0496)	(0.0325)
Forest Cover	0.7540***	-0.0092	-0.0312	0.0182	-0.0289	0.0184
	(0.0537)	(0.0698)	(0.0642)	(0.0772)	(0.0634)	(0.0767)
Forest Cover × Post	-0.6977***	0.0660	0.0694	0.1504***	0.0638	0.1488***
	(0.0747)	(0.0800)	(0.0727)	(0.0554)	(0.0721)	(0.0648)
Cragg-Donald Wald F statistic	1651.402	2.793	2.717	31.651	2.196	29.160
IV Estimation	**Dep. Variable: Average Completion of Secondary Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
School	0.0994***	0.2086	0.2258	-0.0136	0.2081	-0.0293
	(0.0089)	(0.3279)	(0.3395)	(0.0845)	(0.3542)	(0.0840)
Observations	4,231	4,231	4,231	4,231	4,209	4,209
State FE	-	✓	✓	-	-	-
Year FE	-	-	✓	-	-	-
State-Year FE	-	-	-	-	✓	✓
District FE	-	-	-	✓	-	✓

Notes: The dependent variable is the village average of an indicator, which equals 1 if the individual completes secondary education. The interaction of forest cover and cohort eligibility is the instrument. The estimation controls for state-year and district-fixed effects. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

attitudes toward education. Village-level analysis would help capture these community-wide effects. Table 6(a) presents the village-level results for years of education. The estimates, although statistically insignificant, show a decline in the average years of education by almost 4 years, mirroring the individual-level findings. A robust first stage, suggesting that the instrument is not weak, persists. The village- and individual-level estimates complement each other, providing a comprehensive view of how residential school exposure affects educational outcomes. Table 6(b) presents the results for the average secondary school completion at the village level. A statistically insignificant result supports the idea that women exposed to the residential school reported a decline in years of education and were, therefore, unable to complete their secondary education.

7 Robustness

7.1. Difference-in-Difference

In this subsection, I provide evidence to support the robustness of the difference-in-difference and triple difference estimates. An identifying assumption of this approach is the parallel trend assumption. For an unbiased estimate, the average difference between the treated and control groups would remain constant over time without treatment. This translates to having the treatment effects close to or equal to 0 for the pre-treatment period. Figure 3 plots the impact on the years of education of attending the residential schools. The cohorts not eligible to attend the school show no impact as they were beyond the school-going years, supporting the assumption of parallel trends.

Another concern is whether the estimates truly identify the causal impact or whether we estimate the effect by chance, which some other factors may drive. One way to check is to apply the methodology to the group not exposed to the treatment. If we find any statistically significant estimates, we can assume that other factors drive the results. I do this in two ways. First, I run equation (1) for the ineligible cohort - women who have passed the school-going age. It is difficult to think of why these women would be exposed to the treatment. Second, I consider the villages without any residential schools. These villages were not treated, so the years of education should not be impacted. To test this, following Roth et al. (2023), I randomly assign *placebo schools* to these villages and run equations (1) and (5) with these placebo schools. Third, the non-ST women were not exposed to the treatment either. So, I compare the years of education of eligible and ineligible *non-ST* cohorts in villages with a placebo school.

Tables 7 to 11 present the results for the above alternate specifications. None of the groups or combinations of groups not exposed to the treatment show statistically significant results. The triple difference estimates in Table 11 for a placebo school also remain insignificant. These alternate estimates show that the negative impact observed on the years of education because of the residential schools is, in fact, a robust causal effect.

Table 7: Education: Ineligible Cohort

Dependent Variable:	**Years of Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
$ST \times School$	1.130*	-0.9611*	-0.8885*	-0.5626	-0.2572	0.2593
	(0.5868)	(0.5384)	(0.5159)	(0.5147)	(0.4312)	(0.0951)
Observations	5,449	5,449	5,443	5,443	4,605	4,605
R-squared	0.0392	0.1473	0.1919	0.2815	0.3365	0.4029
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is the total years of education. The estimation controls for state-year and district-fixed effects. Controls include religion, marital status, gender of the household head, and income status. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 8: Education: Ineligible Cohort and Placebo School

Dependent Variable:	**Years of Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
ST × Placebo School	-1.025	-1.003	-0.9098	-0.2878	-0.5216	-0.0076
	(0.8482)	(0.7195)	(0.6644)	(0.7541)	(0.4864)	(0.5357)
Observations	4,063	4,063	4,053	4,053	4,053	4,053
R-squared	0.0280	0.1238	0.1893	0.4328	0.4595	0.5550
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is the total years of education. The estimation controls for state-year and district-fixed effects. Controls include religion, marital status, gender of the household head, and income status. women above 25 years are included. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 9: Education: Villages with Placebo Schools

Dependent Variable:	**Years of Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
ST × Placebo School	-1.112	-0.9061	-0.7790	0.6558	-0.7804	-0.7403
	(0.8532)	(0.7008)	(0.6903)	(0.7329)	(0.6752)	(0.6983)
Observations	1,921	1,921	1,919	1,914	1,403	1,400
R-squared	0.0411	0.1230	0.1918	0.3981	0.3957	0.5179
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is the total years of education. The estimation controls for state-year and district-fixed effects. Controls include religion, marital status, gender of the household head, and income status. women above 25 years are included. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 10: Education: Age-Eligible Non-ST and Placebo Schools

Dependent Variable:	**Years of Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
$Age - Eligible \times Placebo\ School$	0.0885	0.0141	-0.2156	-0.1328	0.3438	0.3221
	(0.4651)	(0.4450)	(0.4488)	(0.4060)	(0.4277)	(0.4234)
Observations	4,414	4,414	4,392	4,386	2,966	2,959
R-squared	0.0644	0.1278	0.2071	0.4282	0.4437	0.5194
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is the total years of education. The estimation controls for state-year and district-fixed effects. Controls include religion, marital status, gender of the household head, and income status. women above 25 years are included. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 11: Education: Triple Difference (Placebo Schools)

Dependent Variable:	**Years of Education**					
	(1)	(2)	(3)	(4)	(5)	(6)
$ST \times Eligible \times Placebo\ School$	-0.0868	0.1380	0.2244	0.0055	-0.7737	-0.8730
	(0.7466)	(0.7110)	(0.7372)	(0.7589)	(0.7489)	(0.7435)
Observations	5,984	5,984	5,975	5,975	4,339	4,339
R-squared	0.0878	0.1680	0.2342	0.4278	0.4606	0.5288
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is the total years of education. The estimation controls for state-year and district-fixed effects. Controls include religion, marital status, gender of the household head, and income status. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Instrumental Variable

In this subsection, I provide evidence to rule out possible reasons that might violate the causal identification assumptions of the instrumental variable strategy. I look into the potential indirect channels through which the interaction of cohort eligibility with forest cover could influence years of education. The assumption here is that the instrument (forest cover interacted with cohort eligibility) should only affect the dependent variable (years of education) through the endogenous variable (exposure to the residential school). The existence of other channels would violate the instrument's exclusion restriction and could lead to biased estimates.

Following Nunn and Qian (2014), I test the link between other geographical characteristics and years of education. If the identification strategy is valid, these geographical characteristics should not have the same relationship with the sanctioning of schools and, therefore, years of education as the forest cover. The results of this test are reported in Table 12. Columns (1) to (3) compare the baseline estimates to elevation and terrain ruggedness for all districts. Columns (4) to (6) show the regional elevation and ruggedness results for only districts with at least one residential school. The coefficients for the placebo characteristics suggest no significant relationship between sanctioning schools and years of education. This means that forest cover is not picking up broader environmental effects.

A violation of the exclusion restriction could occur if areas with denser forest cover present specific health challenges that impact educational attainment through increased absenteeism. If these health challenges disproportionately affect eligible cohorts, they could directly influence education outcomes independent of residential school attendance. Following Magesan and Swee (2018), I control for several health-related variables in the baseline model to address this concern. While it is impossible to determine the exact health status of ST women during their exposure to the schools, the health measures I include are those that develop over time and could potentially have begun in the early years. Tables 13 and 14

Table 12: Other Geographical Characteristics

	All Districts			Districts with a School		
First Stage: School Exposure	Forest Cover (1)	Terrain Ruggedness (2)	Elevation (3)	Forest Cover (4)	Terrain Ruggedness (5)	Elevation (6)
Post	-0.0682	0.0003	-0.0765	-0.0736**	0.0027	-0.0856
	(0.0325)	(0.0504)	(0.0786)	(0.0342)	(0.0530)	(0.0879)
Geographic Characteristic	-0.0018	-0.0859	0.0737	-0.0068	-0.0821	0.0696
	(0.0871)	(0.0752)	(0.1051)	(0.0874)	(0.0760)	(0.1040)
Geographic Characteristic × Post	0.1753***	0.0150	0.0995	0.1840***	0.0116	0.1081
	(0.0565)	(0.0621)	(0.0851)	(0.0586)	(0.0655)	(0.0945)
Cragg-Donald Wald F statistic	36.038	6.742	15.960	33.580	5.795	14.674
IV Estimation	**Dep. Variable: Years of Education**					
School	**-4.145****	-6.130	-0.2917	**-4.177****	-6.308	-0.9918
	(2.084)	(7.341)	(2.042)	(2.051)	(1.653)	(2.081)
Observations	3,604	3,604	3,604	3,238	3,238	3,238
State-Year FE	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓

Notes: The dependent variable is the total years of education. The interaction of the geographic characteristic and cohort eligibility is the respective instrument. The estimation controls for state-year and district-fixed effects. Controls include income status, marital status, religion, and the gender of the household head. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

present the results with these additional controls. Columns (4) and (5) show that including the health controls does not alter the impact on years of education. This confirms that the instrument affects the years of education *only* through the channel of the residential schools.

Next, I test whether the instrument affects educational outcomes before the schools are sanctioned. Any statistically significant relationship between the years of education and the instrument for this subpopulation would imply that the interaction of eligibility and forest cover affects the years of education from some other channel. I report the estimates in Tables 15 and 16. The coefficients remain statistically insignificant across specifications.

I further consider a subpopulation that should not be exposed to the residential schools — the non-ST women — and therefore should have no impact on their years of education because of these schools. Tables 17 and 18 show the results. As expected, the residential schools did not affect the non-ST population. These tests confirm the instrument's validity.

Table 13: Education: All Districts

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)
Post	-0.0251	-0.0503	-0.0463	-0.0682**	-0.0679
	(0.0496)	(0.0332)	(0.0325)	(0.0325)	(0.0321)
Forest Cover	-0.0289	0.0182	0.0184	0.0018	0.0010
	(0.0634)	(0.0772)	(0.0767)	(0.0871)	(0.0869)
Forest Cover × Post	0.0638	0.1504***	0.1488***	0.1753***	0.1742***
	(0.0721)	(0.0554)	(0.0548)	(0.0565)	(0.0564)
Cragg-Donald Wald F statistic	2.196	31.651	29.160	36.038	35.737
IV Estimation	**Dep. Variable: Years of Education**				
School	-2.434	-5.516	-4.907	**-4.145****	**-4.301****
	(9.598)	(3.387)	(3.264)	(2.084)	(2.105)
Abnormal Glucose Level	-	-	-	-	-1.500
	-	-	-	-	(1.298)
Stunting	-	-	-	-	-0.3786
	-	-	-	-	(0.2692)
Anemia	-	-	-	-	0.0052
	-	-	-	-	(0.2104)
Hemoglobin	-	-	-	-	-0.1415
	-	-	-	-	(0.1981)
Hypertension	-	-	-	-	-0.0212
	-	-	-	-	(0.0829)
Respiratory Disease	-	-	-	-	-0.1970
	-	-	-	-	(0.2141)
Heart Disease	-	-	-	-	0.0449
	-	-	-	-	(0.1852)
Observations	4,209	4,231	4,209	3,604	3,604
State-Year FE	✓	-	✓	✓	✓
District FE	-	✓	✓	✓	✓
Controls	-	-	-	✓	✓

Notes: The dependent variable is the total years of education. The interaction of the forest cover and cohort eligibility is the instrument. The estimation controls for state-year and district-fixed effects. Controls include ST status, income status, marital status, religion, and the gender of the household head. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 14: Education: Districts with a Residential School

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)
Post	-0.0751*	-0.0519	-0.0481	-0.0739	-0.0736**
	(0.0457)	(0.0350)	(0.0339)	(0.0344)	(0.0340)
Forest Cover	-0.0733	0.0156	-0.0157	-0.0077	-0.0069
	(0.0589)	(0.0774)	(0.0767)	(0.0874)	(0.0871)
Forest Cover × Post	0.1410**	0.1529***	0.1524***	0.1842***	0.1830***
	(0.0654)	(0.0569)	(0.0562)	(0.0585)	(0.0584)
Cragg-Donald Wald F statistic	13.007	28.622	26.780	33.558	33.226
IV Estimation	**Dep. Variable: Years of Education**				
School	11.497	-5.479	-4.870	**-4.082****	**-4.200****
	(7.225)	(3.404)	(3.256)	(2.041)	(2.054)
Abnormal Glucose Level	-	-	-	-	-1.399
	-	-	-	-	(1.313)
Stunting	-	-	-	-	-0.3050
	-	-	-	-	(0.2912)
Anemia	-	-	-	-	0.0643
	-	-	-	-	(0.2234)
Hemoglobin	-	-	-	-	0.0407
	-	-	-	-	(0.2118)
Hypertension	-	-	-	-	-0.0355
	-	-	-	-	(0.1096)
Respiratory Disease	-	-	-	-	-0.2146
	-	-	-	-	(0.2264)
Heart Disease	-	-	-	-	0.1816
	-	-	-	-	(0.1323)
Observations	3,839	3,858	3,839	3,234	3,234
State-Year FE	✓	-	✓	✓	✓
District FE	-	✓	✓	✓	✓
Controls	-	-	-	✓	✓

Notes: The dependent variable is the total years of education. The interaction of the forest cover and cohort eligibility is the instrument. The estimation controls for state-year and district-fixed effects. Controls include ST status, income status, marital status, religion, and the gender of the household head. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 15: Education Before Schools: All Districts

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.7606***	-0.0780	-0.0885	0.0703	-0.0986	0.0651
	(0.0370)	(0.0642)	(0.0790)	(0.0560)	(0.0872)	(0.0573)
Forest Cover	0.7865***	0.0074	0.0021	0.0892	0.0083	0.0801
	(0.0409)	(0.0529)	(0.0528)	(0.0673)	(0.0564)	(0.0720)
Forest Cover × Post	-0.7648***	-0.0072	0.0014	0.0011	0.0007	0.0076
	(0.0404)	(0.0198)	(0.0208)	(0.0133)	(0.0292)	(0.0119)
Cragg-Donald Wald F statistic	2317.043	1.088	0.681	15.412	0.855	12.380
IV Estimation	**Dep. Variable: Years of Education**					
School	6.986***	-4.213	-1.719	-8.119	0.8116	-7.404
	(0.3535)	(8.581)	(9.938)	(6.982)	(8.582)	(6.464)
Observations	5,898	5,898	5,868	5,868	5,048	5,048
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Table 16: Education Before Schools: Districts with a Residential School

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.8497***	-0.0420	-0.0665	0.0737	-0.0688	-0.0657
	(0.0321)	(0.0589)	(0.0748)	(0.0590)	(0.0782)	(0.0612)
Forest Cover	0.8425***	-0.0011	-0.0084	0.0870	-0.0018	0.0777
	(0.0376)	(0.0502)	(0.0497)	(0.0668)	(0.0521)	(0.0710)
Forest Cover × Post	-0.8489***	-0.0045	0.0086	0.0067	0.0057	0.0141
	(0.0359)	(0.0194)	(0.0217)	(0.0154)	(0.0224)	(0.0142)
Cragg-Donald Wald F statistic	2576.071	0.392	0.533	14.373	0.396	11.465
IV Estimation	**Dep. Variable: Years of Education**					
School	6.432***	26.078	24.634	-8.271	2.173	-7.229
	(0.3106)	(49.577)	(44.195)	(7.034)	(11.506)	(6.241)
Observations	5,371	5,371	5,345	5,345	4,528	4,526
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Table 17: Education For Non-ST: All Districts

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.7599***	-0.0501	-0.0889	0.0558	-0.1424	0.0291
	(0.0457)	(0.0685)	(0.0879)	(0.0407)	(0.1047)	(0.0323)
Forest Cover	0.7815***	-0.0008	0.0159	0.0528	-0.0019	-0.0027
	(0.0598)	(0.0703)	(0.0697)	(0.0782)	(0.0849)	(0.0956)
Forest Cover × Post	-0.7930***	-0.0297	-0.0467	-0.0297	-0.0512	-0.0254
	(0.0592)	(0.0435)	(0.0433)	(0.0261)	(0.0495)	(0.0225)
Cragg-Donald Wald F statistic	2065.250	1.338	1.628	1.361	3.176	0.802
IV Estimation	**Dep. Variable: Years of Education**					
School	9.441***	21.152	18.525	-5.747	0.9432	16.486
	(0.4844)	(39.252)	(24.083)	(12.842)	(4.472)	(42.014)
Observations	2,949	2,949	2,920	2,920	2,276	2,271
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is the total years of education. The interaction of the forest cover and cohort eligibility is the instrument. The estimation controls for state-year and district-fixed effects. Controls include income status, marital status, religion, and the gender of the household head. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

Table 18: Education For Non-ST: Districts with a Residential School

First Stage: School Exposure	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.8878***	0.0083	0.0094	0.0530	-0.0252	0.0327
	(0.0346)	(0.0570)	(0.0687)	(0.0448)	(0.0880)	(0.0378)
Forest Cover	0.8518***	-0.0490	-0.0300	0.0600	-0.0601	0.0095
	(0.0514)	(0.0599)	(0.0560)	(0.0767)	(0.0672)	(0.0933)
Forest Cover × Post	-0.9363***	-0.0240	-0.0472	-0.0370	-0.0685*	-0.0361
	(0.0484)	(0.0376)	(0.0374)	(0.0291)	(0.0415)	(0.0270)
Cragg-Donald Wald F statistic	2816.844	8.545	8.398	1.355	17.438	0.753
IV Estimation	**Dep. Variable: Years of Education**					
School	8.531***	11.371	11.846	-5.686	1.352	12.649
	(0.3903)	(11.806)	(11.160)	(11.763)	(2.605)	(27.234)
Observations	2,614	2,614	2,587	2,587	1,941	1,938
State FE	-	✓	-	-	-	-
Year FE	-	✓	-	-	-	-
State-Year FE	-	-	✓	✓	✓	✓
District FE	-	-	-	✓	-	✓
Controls	-	-	-	-	✓	✓

Notes: The dependent variable is the total years of education. The interaction of the forest cover and cohort eligibility is the instrument. The estimation controls for state-year and district-fixed effects. Controls include income status, marital status, religion, and the gender of the household head. Standard errors are clustered at the village level and shown in brackets. *p<.10; **p<.05; ***p<.01.

8 Mechanism

In this section, I discuss and provide evidence for the potential underlying mechanism. To find the reasons for dropping out of school, I use schedule 25.2 of the 75^{th} round of the National Sample Survey. This round, conducted by the National Sample Survey Office (NSSO), includes district-level data on educational attainment and services, including information on literacy rates, educational levels, types of institutions attended, enrolment status, reasons for non-attendance, and educational expenditures.

I estimate the following equation to understand why women drop out of school:

$$Dropout\ Reason_{idt} = \beta_0 + \beta_1 EMRS_{dt} + \beta_2 X_i + \gamma_d + \epsilon_{idt}$$

where $Dropout\ Reason_{idt}$ is the reason individual i in district d dropped out of school, $EMRS_{dt}$ is an indicator of whether the district has at least one residential school, X_i are the individual controls, and γ_d is the district fixed effects. β_1 is the coefficient of interest as the coefficient shows the correlation between the reason for dropout and whether the district had a residential school.

Table 19 reports the estimates of dropout reasons among women residing in districts with at least one EMRS. Notably, three key reasons emerge as statistically significant. First, dropout after completing the desired class increases by 3.81 percentage points, which suggests that after achieving basic educational milestones, students may feel less incentivized to continue. However, the significant decline in dropouts due to financial reasons, by 20.37 percentage points, indicates the effect of EMRS in alleviating the financial strain on households. This is expected, given that EMRS covers tuition and related costs, thus reducing the financial barriers that traditionally lead to school dropouts.

Conversely, there is a notable 27.13 percentage point increase in dropout due to engagement in domestic activities, which points to a critical mechanism underlying the observed decline in educational attainment for two reasons. First, the structure of EMRS requires students to reside in the school for the entire academic year, limiting their ability to assist with household domestic work. This is particularly true for families who may have relied on the child to contribute to household chores or income-generating activities. In socio-economically disadvantaged ST households, the burden on the household increases without the child's contribution to daily domestic activities. Thus, the absence of the child increases the opportunity cost of school enrollment for the household.

Second, if these students had instead enrolled in day schools, they could still contribute to household activities after school hours, reducing this strain. However, as demonstrated in Figure 5, there were little to no day schools in districts with at least one EMRS. This indicates that students in these regions face limited educational alternatives to the residential school model. As a result, the decision for many families becomes binary: either enroll their children in the residential EMRS and lose valuable household labor or choose to have them drop out of school entirely. Given the ST background, the results indicate that the households chose the latter. The evidence underscores the need for a more tailored educational system that accounts for household dynamics in socio-economically disadvantaged communities.

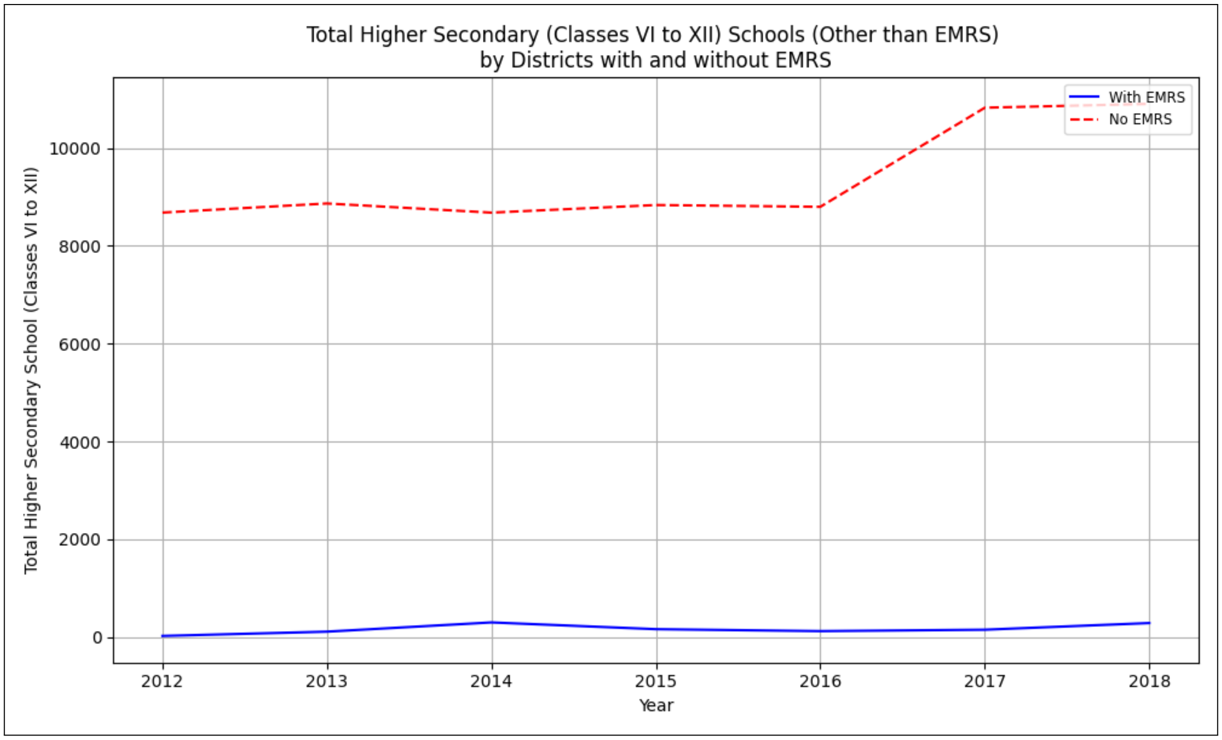
Total Higher Secondary (Classes VI to XII) Schools (Other than EMRS)
by Districts with and without EMRS
With EMRS
No EMRS
Total Higher Secondary School (Classes VI to XII)
0
2000
4000
6000
8000
10000
2012
2013
2014
2015
2016
2017
2018
Year

Figure 5

Table 19: Major Reasons for Dropping Out of School

Major Reason:	**Economic Activity**	**Domestic Activity**	**Financial Reason**	**Completed Desired Class**	**Unfamiliar Language**	**Inadequate Teachers**	**Teacher Quality**
EMRS Present in District	0.1354	0.2713***	-0.2037***	0.0381**	-0.0007	-0.0002	-0.0399
	(0.0944)	(0.0985)	(0.0688)	(0.0185)	(0.0006)	(0.0003)	(0.0397)
Observations	8,906	8,906	8,906	8,906	8,906	8,906	8,906
R-squared	0.1029	0.0928	0.1117	0.1150	0.0486	0.0294	0.0376

Major Reason:	**School Far**	**Academic Failure**	**Atmosphere Unsuitable**	**No Female Teacher**	**No Interest**	**Marriage**
EMRS Present in District	-0.0033	-0.0428	0.0001	0.0002	-0.0004	0.0382
	(0.0022)	(0.0416)	(0.0001)	(0.0001)	(0.0007)	(0.0252)
Observations	8,906	8,906	8,906	8,906	8,906	8,906
R-squared	0.0888	0.0838	0.0140	0.0325	0.0657	0.0509

Notes: Dependent variable is an indicator for the respective reason. The estimation controls for district-fixed effects. Controls include religion, household size, and monthly expenditure. The sample was restricted to only ST households. The sample includes women last enrolled in primary, upper primary, middle, secondary or higher secondary education. Standard errors are clustered at the household level and shown in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

9 Conclusion

This paper presents the first comprehensive study of the impact of *Eklavya Model Residential Schools* (EMRS) on the educational outcomes of the Indigenous[12] women in India. Against the backdrop of a global decline and phase-out of residential schools, India's commitment in the 2024 budget to allocate substantial resources and INR 63.99 billion to expanding EMRS across the country calls for an evaluation of the policy's effectiveness. Using a newly constructed dataset with a triple difference and instrumental variables strategies, I estimate the causal effect of residential school enrollment on years of education for Indigenous women in India.

Three key differences were helpful in the triple difference identification: a residential school in a village, the tribal status of children, and cohort eligibility, contingent on an individual's age falling within the school-going bracket when the school was operational. The findings show that exposure to EMRS led to unintended negative consequences. ST women residing in villages with operational schools during the school-going age experienced significantly fewer years of education than their non-ST counterparts. Specifically, eligible cohorts report 1.248 fewer years of education than non-ST women.

To check the robustness of these results, I use an instrumental variable (IV) analysis to further substantiate these findings. I use the interaction of forest cover and cohort eligibility as an instrument for school exposure. The idea is that the STs reside in very remote areas of India, often having dense forest covers. Since these schools catered to the STs, areas with higher forest cover were more likely to have these schools sanctioned. The robust first stage across specifications supports the argument that the instrument is not weak. The IV results demonstrate consistent declines in years of education at both the individual and village levels. Specifically, I observe a decrease of 4 years of education at the individual and

[12] Scheduled Tribes (ST) of India

village levels. The results, robust to multiple checks, reinforce the conclusion that exposure to these schools did not lead to improved educational outcomes.

Finally, I explore the underlying mechanisms driving these results. I find a significant 27.13 percentage point increase in dropout rates due to engagement in domestic activities in districts with EMRS. This suggests that the residential nature of EMRS disrupts traditional household dynamics, increasing the opportunity cost of education for families who rely on their children's labor for domestic and economic support. Furthermore, the lack of alternative schooling options in districts with EMRS leaves families with a binary choice: enroll their children in EMRS or have them drop out entirely. This forced trade-off may explain the decline in educational attainment observed in these regions.

While we have robust evidence of residential schooling's negative impact on STs, avenues for further investigation exist. This study focuses on short-term educational outcomes. Future work can examine the broader economic, social and psychological effects of residential schooling on ST communities. Issues such as mental health, community cohesion, and cultural preservation are equally crucial in understanding the long-term effects of these schools. Studies in the North American context have shown that the impact of residential schools extends beyond educational attainment, affecting various aspects of life, including health behaviours, mental health, cultural engagement, and labor market outcomes. A similar exploration in the Indian context would be beneficial.

While the EMRS program was implemented to improve educational outcomes for STs, its failure to account for the unique socio-cultural dynamics of tribal communities may have resulted in unintended negative consequences. The findings underscore the need for more tailored and culturally sensitive educational policies that better align with the realities of Indigenous populations.

References

Adams, David Wallace. 1995. "Education for extinction." *Lawrence: University Press of Kansas* 23:164.

Banerjee, Abhijit V, Shawn Cole, Esther Duflo and Leigh Linden. 2007. "Remedying education: Evidence from two randomized experiments in India." *The Quarterly Journal of Economics* 122(3):1235–1264.

Barrera-Osorio, Felipe, Marianne Bertrand, Leigh L Linden and Francisco Perez-Calle. 2011. "Improving the design of conditional transfer programs: Evidence from a randomized education experiment in Colombia." *American Economic Journal: Applied Economics* 3(2):167–95.

Black, Sandra E, Paul J Devereux and Kjell G Salvanes. 2005. "Why the apple doesn't fall far: Understanding intergenerational transmission of human capital." *American economic review* 95(1):437–449.

Bombay, Amy, Kimberly Matheson and Hymie Anisman. 2014. "The intergenerational effects of Indian Residential Schools: Implications for the concept of historical trauma." *Transcultural psychiatry* 51(3):320–338.

Bougie, Evelyne and Sacha Senécal. 2010. "Registered Indian children's school success and intergenerational effects of residential schooling in Canada." *International Indigenous Policy Journal* 1(1):1–41.

Breierova, Lucia and Esther Duflo. 2004. "The impact of education on fertility and child mortality: Do fathers really matter less than mothers?".

Deming, David and Susan Dynarski. 2009. Into college, out of poverty? Policies to increase the postsecondary attainment of the poor. Technical report National Bureau of Economic Research.

Duflo, Esther. 2001. "Schooling and labor market consequences of school construction in

Indonesia: Evidence from an unusual policy experiment." *American economic review* 91(4):795–813.

Dynarski, Susan. 2004. The new merit aid. In *College choices: The economics of where to go, when to go, and how to pay for it.* University of Chicago Press pp. 63–100.

Dynarski, Susan and Judith Scott-Clayton. 2013. "Financial aid policy: Lessons from research.".

Feir, Donna L. 2016. "The long-term effects of forcible assimilation policy: The case of Indian boarding schools." *Canadian Journal of Economics/Revue canadienne d'économique* 49(2):433–480.

Feir, Donna L and M Christopher Auld. 2021. "Indian residential schools: Height and body mass post-1930." *Canadian Journal of Economics/Revue canadienne d'économique* 54(1):126–163.

Glenn, Charles. 2011. *American Indian/First Nations schooling: From the colonial period to the present.* Springer.

Gregg, Matthew T. 2018. "The long-term effects of American Indian boarding schools." *Journal of Development Economics* 130:17–32.

Jedwab, Remi, Felix Meier zu Selhausen and Alexander Moradi. 2022. "The economics of missionary expansion: Evidence from Africa and implications for development." *Journal of Economic Growth* 27(2):149–192.

Jones, Maggie EC. 2021. "The intergenerational legacy of indian residential schools." *University of Victoria, Victoria, British Columbia. At https://maggieecjones. files. wordpress. com/2021/02/intergenerationalrs. pdf* .

Jones, Maggie EC. 2023. "Post-secondary funding and the educational attainment of indigenous students." *Economics of Education Review* 97:102475.

Kazianga, Harounan, Dan Levy, Leigh L Linden and Matt Sloan. 2013. "The effects of" girlfriendly" schools: Evidence from the BRIGHT school construction program in Burkina Faso." *American Economic Journal: Applied Economics* 5(3):41–62.

Kremer, Michael, Edward Miguel and Rebecca Thornton. 2009. "Incentives to learn." *The Review of Economics and Statistics* 91(3):437–456.

Magesan, Arvind and Eik Leong Swee. 2018. "Out of the ashes, into the fire: The consequences of US weapons sales for political violence." *European economic review* 107:133–156.

Mehta, BH. 1953. "Historical background of tribal population." *Indian Journal of Social Work* 14(3):236–244.

Meriam, Lewis. 1971. *The problem of Indian administration.* Number 17 Johnson Reprint Corporation.

Milloy, John S. 2017. *A national crime: The Canadian government and the residential school system.* Vol. 11 Univ. of Manitoba Press.

National Family Health Survey (NFHS), India, 2019-21. 2021. "International Institute for Population Sciences (IIPS) and ICF.".

Nunn, Nathan and Nancy Qian. 2014. "US food aid and civil conflict." *American economic review* 104(6):1630–1666.

Oreopoulos, Philip, Marianne E Page and Ann Huff Stevens. 2006. "The intergenerational effects of compulsory schooling." *Journal of Labor Economics* 24(4):729–760.

Reyhner, Jon and Jeanne Eder. 2017. *American Indian education: A history.* University of Oklahoma Press.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski and John Poe. 2023. "What's trending in difference-in-differences? A synthesis of the recent econometrics literature." *Journal of Econometrics* 235(2):2218–2244.

Szasz, Margaret Connell. 2006. "Through a wide-angle lens: Acquiring and maintaining power, position, and knowledge through boarding schools." *Boarding School Blues: Revisiting American Indian Educational Experiences* pp. 187–201.

Truth and Reconciliation Commission of Canada. 2015. *Honouring the Truth, Reconciling for the Future: Final Report of the Truth and Reconciliation Commission of Canada: Volume One: Summary.* James Lorimer Limited, Publishers.

Valencia Caicedo, Felipe. 2019. "The mission: Human capital transmission, economic persistence, and culture in South America." *The Quarterly Journal of Economics* 134(1):507–556.