The Association between Adverse Temperature Shocks and Schooling Outcomes in India: Impact Quantification and Mitigation Potentials

Sumit Agarwal[a], Pulak Ghosh[b], Francesco Scarazzato[c], and Sofie R. Waltl[d,c]

[a]National University of Singapore, Business School
[b]Indian Institute of Management Bangalore, Decision Sciences
[c]Vienna University of Economics and Business, Department of Economics
[d]University of Cambridge, Department of Land Economy

Preliminary Draft as of September 5, 2024

Abstract

Does extreme heat adversely affect the educational outcomes of kids in India? To address this question, we link results in the Indian Upper Primary Level Examination to information on local weather conditions, air pollution, and vegetation density sourced from remote sensing data. This link is established by precisely geo-coding school addresses. The panel structure of our data allows us to track the success of students attending the same school over four years, while accounting for any time-invariant school- or location-specific attributes. Our analysis reveals that both cumulative heat exposure and exposure to higher temperatures during examinations adversely affect students' performance. A constant increase in temperature by merely 0.5°C results in a drop in the number of students passing the exam by approximately 2% and a drop in the number of highest grades ("distinctions") of almost 15%, hinting towards a sizable potential loss in human capital. We find that the negative impact on exam pass rates intensifies with higher temperature ranges, with the effect being largest on days with maximum temperatures exceeding 40°C. Furthermore, we show that vegetation in the proximity of schools has a strong mitigating effect that increases with forest cover density: given an area of about 80 hectares around each school, 2.6 hectare of forest would offset the impact of an increase in temperature by 1°C on the probability of passing the exam, and 9.5 hectares would offset the effect of a 1°C increase, on the probability of achieving a distinction. These findings suggest that increasing vegetation in the vicinity of schools is a valid policy recommendation to soften the adverse impacts of extreme heat on adverse educational outcomes, and adapt to expected long-term changes in climate.

Keywords: Education, Extreme Weather Events, Heat, Adaptation, India
JEL codes: Q54, I24, I25

Notes and Acknowledgments: We are grateful for fruitful discussions with and comments by participants in the *2023 NOeG conference* at the *Paris Lodron University of Salzburg*, the *63rd ERSA Congress* in Portugal, as well as research seminars at the *Wegener Center for Climate and Global Change* at the *University of Graz*, the *Cambridge Centre for Environment, Energy and Natural Resources Governance (CEENRG)* at the *University of Cambridge* as well as *WU Department of Economics PhD seminar*. Further, we thank Jesús Crespo Cuaresma, Shailaja Fennell, Klaus Gugler, Lukas Brunninger, Nikolas Kuschnig, Lukas Vashold, Ulrich Wohak, Simon Heß and Johannes Boehm for valuable comments and suggestions.
This article uses data from the *District Information System for Education (DISE)*, the *Copernicus Climate Change Service* and *NASA Earth Observations (NEO)*. This article should not be reported as representing the views of the data providers. The views expressed are solely those of the authors.

"India will have the highest population of young people in the world over the next decade, and our ability to provide high-quality educational opportunities to them will determine the future of our country."

Indian National Education Policy 2020[1]

1 Introduction

In India, significant warming trends have been measured over the past decades and the frequency of warm-extreme events has increased. Further, the frequency and intensity of warm days and warm nights are projected to increase whereas those of cold days and cold nights are expected to decrease (Sanjay et al., 2020). UNICEF's Children's Climate Risk Index is a composite index linking child vulnerability to exposure to climate and environmental hazards. In-line with these projected trends, India ranks at the very low position 26 of 163 countries, mainly due to a very bad score in the category *"Climate and environmental shocks"*. Debnath et al. (2023) show that extreme heat and prolonged heatwaves expose India to particular risks with regard to its developmental efforts, and could slow or even reverse the country's progress towards the *Sustainable Development Goals (SDGs)*.

In the present study, we measure the consequences of exposure to extreme heat on early schooling success across India and by that provide a quantification of the adverse effects of hot temperatures and heat waves in combination with other meteorological and environmental characteristics. We also provide the first evidence that vegetation in the proximity of the schools may have a mitigating effect on cognitive performance, thus suggesting a possible adaptation strategy.

We do so by combining two types of data: First, we assess the official *District Information System for Education (DISE)* dataset providing longitudinal school- and class-level information on exam results as well as basic information on students' demographics, and descriptive information about schools and teachers. Second, we retrieve longitudinal information on meteorological conditions and land use information for fine locational areas from satellite data (see Donaldson and Storeygard, 2016, for a review of the use of remote sensing data in economics). We match these data to the DISE data via time and location using geo-coded school addresses.

We measure variation in schooling outcomes between 2015 and 2018 relying on results from a country-wide exam that concludes the first cycle of primary education. We observe whether a student 'fails', 'passes' or 'passes with distinction'. We therefore use logistic regressions to estimate the log-odds of passing the exam (or passing with distinction), given observed

[1] See `https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf` for details, last accessed in June 2024.

temperature, humidity and further local weather conditions measured during the school-year preceding the exam as well as the immediate period before and during the exam.

We find that higher measured temperature significantly decrease the odds of students passing the exam as well as the odds of students passing the exam with distinction. This effect becomes even larger when additionally controlling for further meteorological and other remote-sensing data.

To better understand the severity of the impact of temperature shocks, we express the effect sizes in terms of numbers of students affected. We compare the predicted number of students that fail or miss the distinction for the observed temperature and add simulated results for hypothetically increased temperatures. Our estimates predict that an increase in average temperature by merely 0.5°C implies a drop in the number of students passing the exam by almost 2% and a drop in distinctions of almost 15%. In absolute numbers, this means that of all the students in our sample 47,040 and more than 2 Mio., respectively, would achieve less favorable results and will thus likely not live up to their full potential. General results survive a battery of robustness checks.

A heterogeneity assessment reveals that temperature increases do not affect boys and girls differently: although precisely estimated, the magnitude of the differential effect is negligible. We also find schools located in rural areas are more severely affected than urban schools. This could be explained by uneven availability of coping strategies, possibly due to differences in income.

We further investigate how natural green areas affect exposure to heat and its consequences on schooling outcomes by exploiting information on tree cover density in the proximity of schools inspired by the cooling factor of vegetation and trees in particular Han et al. (2024). Our findings indeed hints towards a potential mitigation strategy: students attending schools that are located in areas with more vegetation — and thus experiencing a cooler microclimate -– are significantly more likely to pass the exam and pass it with distinction, respectively.

This study assesses adverse impacts of environmental conditions on schooling success and by that contributes to the broader fields studying the consequences of climate change, as well as economics of education, development and inequality. Related studies, for instance, found that increased air pollution may lead to worse students' performance and absenteeism (e.g., Currie et al., 2009; Ebenstein et al., 2016). Palacios et al. (2022) conduct a field experiment among primary schools in the Netherlands to show that insufficient classroom ventilation negatively affects children's cognitive performance and by that decreases standardized test results. Previous research has also explicitly shown the negative impact of high temperatures on cognitive performance and learning. Cho (2017) and Graff Zivin et al. (2020) study the effects of temperature on short-run cognitive performance during college admission examinations in South Korea and China, respectively, and Conte Keivabu (2024) derives a link between long-term exposure to extreme hot (or cold) temperatures and increased school absenteeism in China.

However, our study is most closely related to the work of Park et al. (2020, 2021), McCormack (2023) and Conte Keivabu (2024), who study how cumulative heat exposure (partly combined with the lack of air conditioning) directly affects education, as it increases absenteeism and inhibits learning both at home and at school.

Elevated heat may thus put at risk a country's educational goals. This is particularly problematic, as we measure the adverse effect for an exam concluding primary education, which may be perceived as an early negative signal of performance, affecting future educational outcomes and potentially overall success in life.

The remainder of this article is structured as follows: section 2 introduces the Indian primary education system, discusses the consequences of failing and describes the mechanism we aim to model. Section 3 presents the measurement concepts, data used and our identification strategy. Thereafter, section 4 assesses the effect of cumulative heat exposure on educational outcomes and section 5 discusses the effectiveness and potential of adaptation and mitigation measures. Finally, section 6 concludes. A comprehensive appendix provides additional details.

2 Context, Implications and Mechanisms

2.1 Schooling in India

Education in India is a fundamental right of every child and thus is generally free and compulsory according to the 2009 *Right of Children to Free and Compulsory Education Act* (see Srivastava and Noronha, 2014, for details). Precisely, Article 45 of the Indian Constitution states:

> **Provision for free and compulsory education for children.** *The State shall endeavor to provide, within a period of ten years from the commencement of this Constitution, for free and compulsory education for all children until they complete the age of fourteen years.*

The Indian education system hence commences with two cycles of *compulsory* primary education: Primary school (grades 1 to 5) and upper primary school (grades 6 to 8), which is followed by largely non-compulsory secondary education or vocational training (see Hill and Chalaux, 2011, Figure 1). Towards the end of each primary education cycle, a country-wide exam takes place: the *Primary Level (Class IV/V) Examination (PLE)* and the *Upper Primary Level (Class VII/VIII) Examination (UPLE)*. The necessary condition to pass these exams is to score at least 35%, while scoring at least 60% results in a distinction.

Following primary education, school continues to be free yet is non-compulsory. Secondary education lasts for four years and is split into two cycles each lasting for two years (General/Lower Secondary School and Upper/Senior Secondary School).

Overall, UDISE+ data[2] shows that there have been significant improvements in access to primary education. The share of students enrolled in upper primary (Class VI-VIII) education has increased from 67.77% for girls and 64.47% for boys in 2012-13 to 71.66% for girls and 71.00% for boys in 2021-2022. These shares still remain low for higher grades: the enrollment share for higher secondary education (Class XI-XII) increased from 24.71% for girls and 25.18% for boys in 2012-13 to only 34.95% and 33.54%, respectively, in 2021-22.

2.2 The Consequences of Failing

As per the 2009 Right To Free and Compulsory Education Act, students in principle cannot be detained even if failing the exam as the Act "prohibits holding back and expulsion of a child from school till the attainment of elementary education."[3]

Ever since then, there has been a heated debate across the country whether failing and repeating a class should be possible or not. In 2019, an amendment to this Act permitted states to hold examinations in Classes V and Class VIII and make children failing these exams to repeat them. This amendment overturned the previously in place "no-detention policy" and children failing the exam may be instructed to take a re-examination. If a child would fail again, according to the Amendment, the relevant state or central government may decide to allow schools to detain the child. States did implement and execute this amendment in different ways.

We assess here data from the intermediate period (2014–2018) when failing did not have immediate implications. However, assessing this period allows us to perform simulation exercises as school data was made publicly available. Consequences are much more severe in those states that have already implemented stricter implications or are debating doing so.

Irrespective of that, previous research has shown that just being associated with failure can already have negative long-lasting consequences: Papay et al. (2010) assess the prospects of students participating in the Massachusetts (US) high-school leaving exam. They adopt a regression discontinuity design to identify a causal effect of just passing or just failing the exam on the probability of graduation. While in the overall student population, there appears to be no significant difference between the two groups, however the heterogeneity analysis identifies a significant negative effect among low-income urban students.

Machin et al. (2020) add evidence from the English education system by investigating the consequences of just failing the national examinations (General Certificate of Secondary Education) to be taken at age 16 by all students at the end of compulsory education. Passing this exam with at least a grade C is often required to be eligible for high-quality upper secondary education and, by that, impacts the quality of preparation for tertiary education. Further, achieving at least a C appears to be also important for employers. Again, causal inference techniques are adopted to compare the outcomes of students who narrowly miss or pass the grade C threshold.

[2]See subsection 3.1 and `https://udiseplus.gov.in/#/home` for details, last accessed in May 2024.

[3]See the Clarification on Provisions of the 2009 Act `https://www.education.gov.in/sites/upload_files/mhrd/files/upload_document/RTE_Section_wise_rationale_rev_0.pdf`, last accessed in June 2024.

They find that students that just fail to achieve a C are more likely to drop out of education early and become classified as 'not in education, employment or training' at age 18.

These findings may thus negatively impact the achievement of the official Indian *National Education Policy 2020* that envisions "an education system rooted in Indian ethos that contributes directly to transforming India, that is Bharat, sustainably into an equitable and vibrant knowledge society, by providing high-quality education <u>to all</u>, and thereby making India a global knowledge superpower" [National Education Policy[1] 2020, paragraph 1.6].

2.3 Mechanisms

We aim to model a reduced form of a complex mechanism explaining the effect of elevated temperatures on cognitive performance in both the long and short run. Thereby, the long term means an impact on learning while the short-run means an impact of cognitive performance during an exam and associated immediate preparation before, i.e., acute cognitive performance.

Previous research has shown that both, acute (immediate) exposure and long-term (cumulative) exposure to extremely high temperatures are associated with cognitive decline. Yin et al. (2024) relate cognitive functionality measured via both a verbal and math test taken as part of a nationwide survey in China to measured temperatures at the date and location of test taking. While exposure to extremely high temperature was found to be significantly associated with cognitive function decline, it induces less cognitive decline for people used to higher temperatures.

A direct physiological effect of heat on reduced cognitive performance thus affects both, learning outcomes during the year (cumulative) and performance in exams (immediate) which may both explain measured variation in exam results. Importantly, next to an absolute measure of heat, deviation from typical conditions appears to be key which we adopt in our testing strategy.

Cumulative exposure to heat does not only affect students and their learning capacity, but also teachers and their quality of teaching provided. Together, these two factors mean a clear comparative disadvantage for students exposed to extended periods of heat during the school year as discussed in Park et al. (2020).

With regard to the impact of heat on learning, a second, indirect effect may be relevant: increased duration or frequency of absenteeism over an extended period of the school year. Conte Keivabu (2024) describes the link between cumulative extreme hot (or cold) temperatures on increased school absences, which in turn may lead to reduced learning outcomes. The study discusses four types of channels contributing to this link: health, behavior, educational institutions, and energy poverty. All these links have been separately studied in existing literature: The health channel focuses on a negative impacts of heat on children's health which leads to absenteeism. The behavioral channel describes changes of parental behavior who may keep children at home to protect them or parent's changes in time allocation during such periods. Further, the institutional channel links extreme weather conditions to school closure mandates which in turn reduce children's learning opportunities in comparison to schools not affected by the same measure. Finally, energy poverty associated with less reliable and less available cooling technologies is further associated with a lost opportunity to mitigate the negative effects of heat on absenteeism.

Conte Keivabu (2024) tests the channel focusing on hot and cold extremes in England and information on attendances in primary and secondary schools in the school year 2011–2012 to 2018–2019. These data are linked to meteorological information on a 0.1° resolution and assessed as cumulative exposure over a school year. These data are amended with information on precipitation and air pollution.

The study finds a significant positive association between absenteeism and increased numbers of warm and hot days. When separating absenteeism by its cause, the study finds that with more warm and hot days the number of authorized holidays and absenteeism due to illnesses increases while there is no systematic association between elevated temperature and both medical appointments nor unauthorized holidays. These findings support both a health and a behavioral channel leading to a causal association between increased temperatures and reduced learning success.

3 Measurement Concepts and Data

3.1 Measuring Schooling Outcomes

Outcomes for the *Primary Level Examination (PLE)* and the *Upper Primary Level Examination (UPLE)* are collected within the *District Information System for Education (DISE)*, a government database created by the Indian *Ministry of Education* and administered by the National University of Educational Planning and Administration (NUEPA). DISE data cover all registered primary and secondary schools. Due to its broadly applicable information content, DISE has been used by researchers for several purposes, e.g., for studying the effect of schools organized by NGOs on general access to education (Blum, 2009), private education (Kingdon, 2020), as well as the link between the provision of sanitary pads and girls' school attendance (Agarwal et al., 2022).

Our main outcome variables are the results of the *UPLE*, also called *Class VIII* exam, which concludes primary education. Results are reported in the categories 'failed', 'passed' or 'passed with distinction'. We focus on an exam taking place early in students' school career as failing may be perceived as a negative signal of performance and, by that, negatively impacting students' remaining educational career and hence also future success in life as concluded from administrative data and longitudinal cohort studies tracking groups of children from birth to old-age (see Machin et al., 2020, for an overview of this literature).

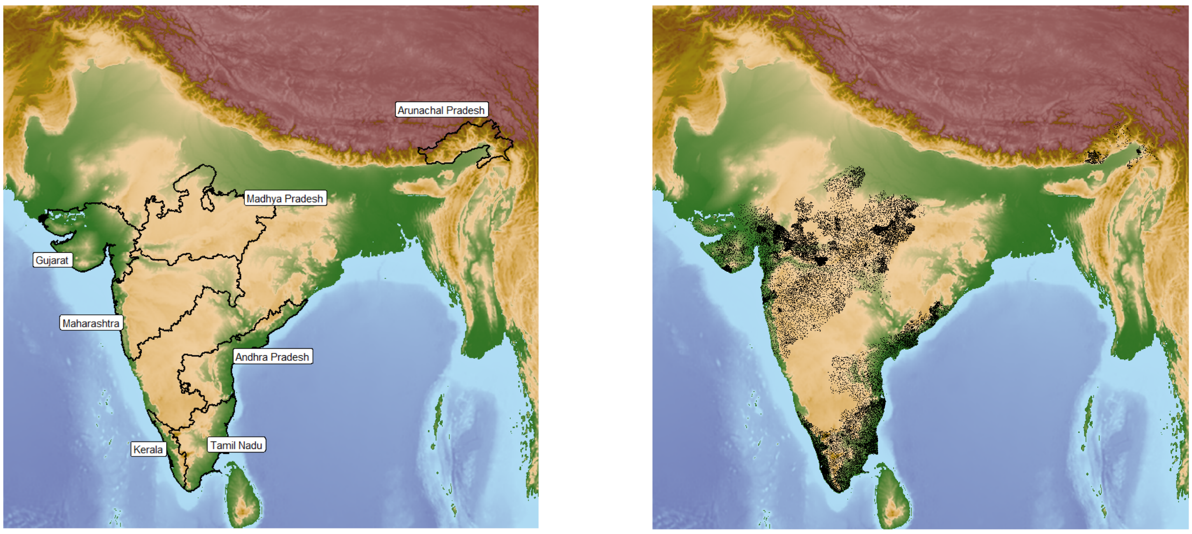
We assess exam results for public schools in seven Indian states (Andhra Pradesh, Arunachal Pradesh, Gujarat, Kerala, Madhya Pradesh, Maharashtra, and Tamil Nadu) between the school years 2014-15 and 2017-18. The exact selection procedure is explained in Appendix A.1 and aimed to ensure a high degree of comparability with regards to the principle type of school and the structure of the academic year. The latter is crucial to consistently link exposure to heat during the same phase of the school year and, thus, also the rough learning schedule.

Ultimately, 74,598 schools enter the analysis: Panel (a) of Figure 1 shows the states assessed and Panel (b) the exact location of individual schools part of our analyses.

We use school addresses to proxy all local weather conditions relevant for affecting the learning process, that comprises learning at school *and* learning at home. This approximation is supported by Rekha et al. (2020) assessing travel distances of students in Tiruchirappalli City (Tamil Nadu). They find that 39% of people need to travel less than 2km, 25% between 2 and 5 km, 33% between 5 and 10 km and only 3% travel larger distances (10–20 km) for school trips. This indicates that our approximation is expected to cover both locations where learning takes place.

Figure 1: Schools' Location.

(a) (b)



Notes: Topographic map of the subcontintent. Panel (a) highlights Indian states included in this study and Panel (b) the location of individual schools.

The data is available at grade-level, i.e., for each grade in each school characteristics of the student and teacher population as well as students' exam results are reported in the form of shares of the total number of students at this grade. For our models, we change the data format by projecting this wide table into a long table yielding information per student i, year t and school s.

To account for all unobserved factors related to time or space, we adopt a panel model approach by including school-year ('time') and school-address ('school') fixed-effects to our final generic model presented in subsection 4.1.

It is important to note that the vast majority of students pass these early exams. Failing is indeed rare yet we are interested in the increase in these rare events and their link with elevated temperature. Moreover, we assume that underlying the observable outcome of the exam, there exists an unobservable continuous variable measuring the amount of learning done by each student. Variations in this latent variable is what we are ultimately interested in.

Because of the school fixed-effects, only schools with reported variation in the outcome variable contribute to the estimation. For this reason, we have to rely on a sub-sample of students (about 10%), who are enrolled in schools where at least one student failed the exam over our period of observation. The same applies to the models predicting the odds of passing with distinction, but the selection criteria is only binding in less than 10% of schools: for this model, we are able to retain about 92% of the observations corresponding to all schools that do not pass all its students with distinction in all years of observation.

Summary statistics are reported as part of Table 1. The first column reports summary statistics for the whole sample; the second column for the sub-sample of students enrolled in schools where somebody fails the exam; the third column for the sub-sample of students enrolled in schools where somebody fails to get a distinction.

Figure 4 in Appendix C reports the distribution of the average temperature in March, by school year, for the whole sample and the two sub-samples.

3.2 Temperature

As the school data contains the exact address of each school, we can create therefrom a geolocation which is precisely linkable to a wide set of meteorologic data (see Appendix A.2 for details). We obtain these meteorological data from different sources whereas the main source constitutes the *Copernicus Climate Change Service*. Precisely, we use the ERA5-Land data set (Muñoz Sabater et al., 2021), which contains hourly values for various climate and weather related variables, from 1950 to the present. The data has a resolution of $0.1° \times 0.1°$ (approximately 100 km^2 at the equator).

We operationalize temperature exposure in two different ways. Our main approach simply relies on the average maximum temperature during schools days[4] (see Figure 6 in Appendix C). Results are reported in Table 2.

We use one more alternative way to measure temperature exposure reported as robustness checks in Table 8 in the appendix. We follow Graff Zivin et al. (2020) and Park et al. (2020) and count the number of school days for which the maximum temperature falls into several temperature bins. We use eight bins with a length of 5°C ranging from $(-\infty; 10°C)$ to $[40°C; +\infty)$. In our regression model we use a "comfortable" temperature, namely $[20°C; 25°C)$, as the reference category. Using bins means relaxing the assumption that the likelihood to pass/pass with distinction is linearly associated with temperature as coefficients to bins proxy any kind of linear, non-linear or even non-continuous functional form of the effect.

To identify the effect of adverse temperature shocks on learning, we regress exam results on the

[4] All days between June and March, with the exclusion of Saturdays, Sundays and the major holidays.

Table 1: Summary Statistics.

	All schools		Only schools with variation in outcome			
			Pass		Pass with Distinction	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Pass	0.988	0.110	0.909	0.288	0.988	0.111
Pass with distinction	0.710	0.454	0.610	0.488	0.691	0.462
Boy	0.518	0.500	0.520	0.500	0.517	0.500
Temperature, °C	30.657	1.513	30.480	1.534	30.650	1.519
Wet Bulb Temperature, °C	21.860	1.611	21.610	1.575	21.856	1.611
Heat Index, °C	31.996	1.958	31.710	1.954	31.992	1.953
Relative humidity, %pt	57.964	7.443	57.481	6.937	57.969	7.476
Wind speed, m/s	2.634	0.649	2.593	0.603	2.627	0.650
Precipitation, mm	2.022	1.007	2.024	0.979	2.028	1.013
Days below 10°C	0.032	1.342	0.045	1.683	0.033	1.358
Days 10-15°C	0.079	2.003	0.128	2.480	0.082	2.046
Days 15-20°C	0.996	4.700	1.355	5.657	1.045	4.786
Days 20-25°C	9.758	14.571	11.572	15.040	9.967	14.639
Days 25-30°C	87.698	34.181	89.146	33.041	87.565	33.974
Days 30-35°C	92.574	30.276	90.397	30.902	92.349	30.103
Days 35-40°C	24.281	17.733	22.464	15.351	24.330	17.785
Days above 40°C	2.352	3.382	2.656	3.501	2.402	3.416
PM 2.5	37.824	10.086	39.530	9.991	37.966	10.194
Carbon Monoxide	90.614	5.257	91.199	5.264	90.556	5.251
Nitrogen Dioxide	189.969	76.848	199.128	80.438	190.870	77.457
Rural	0.657	0.475	0.622	0.485	0.657	0.475
Public	0.507	0.500	0.508	0.500	0.519	0.500
Forest 1.00 km, ha	32.465	68.197	30.450	65.243	32.868	68.858
Forest 0.50 km, ha	8.057	17.463	7.502	16.659	8.156	17.625
Forest 0.25 km, ha	2.005	4.486	1.869	4.302	2.030	4.527
No. of Observations (students)	22,778,692		2,780,912		19,915,599	

Notes: Summary statistics refer to the intersection of all data used as described in detail in section 3. Meteorological data refer to the entire school year.

temperature measured at the school's location during the academic year preceding the exam ('cumulative heat exposure'). Each school year begins in late May or early June, and ends at the end of March, or beginning of April.

Moreover, we test the effect of variations in temperature measured during March (i.e., the last month of each school year). We include temperature at the school's location shortly before and during the exam period ('immediate heat exposure'), because hot temperature likely also affects studying success and adequately preparing for the exam.

3.3 Supplemental Concepts and Data

We rely on a variety of supplemental meteorological data potentially mitigating or intensifying the consequences of heat exposure as discussed in subsection 4.1 and section 5.

3.3.1 Relative Humidity

It has been well-studied that heat in combination with humidity can have even larger impacts on human's health and performance. Extreme heat in combination with humidity means elevated risks and discomfort. As Raymond et al. (2020) put it: "A normal internal human body temperature of 36.8°C ± 0.5°C requires skin temperatures of around 35°C to maintain a gradient directing heat outward from the core. [...] Once the air (dry-bulb) temperature (T) rises above this threshold, metabolic heat can only be shed via sweat-based latent cooling, and at WBT [wet-bulb temperature] exceeding about 35°C, this cooling mechanism loses its effectiveness altogether." They assess WBT hotspots between 1997 and 2017 globally and find that all measures of WBT occurrences show a clear upward trend over time. However, even lower temperatures can have – in combination with humidity – strong adverse effects particularly when adaptive measures are not taken. India is adversely affected by this dangerous combination (see for instance Wehner et al., 2016).

As heat is particularly dangerous in combination with humidity, we additionally make use of the concept of *relative humidity* measuring water vapor relative to air temperature. Therefore, we source temperature and dew point data from the ERA5-Land data set, and calculate relative humidity using the August-Roche-Magnus approximation yielding an easy algebraic relationship between relative humidity [%], the dew point Dp [°C] and temperature T [°C]. Using the coefficients suggested by Alduchov and Eskridge (1996) implies

$$\text{Relative Humidity} = 100 * \frac{\exp\left(\frac{17.625\,Dp}{243.04+Dp}\right)}{\exp\left(\frac{17.625\,T}{243.04+T}\right)}.$$

Again, we link each school's location to daily relative humidity, and average over all school days in each academic year.

3.3.2 Wet-Bulb Temperature and Heat Index

Next to measuring temperature and humidity separately, we can also draw on the concept of *wet bulb temperature* (WBT) that merges dry air temperature with humidity into a single measure as mentioned in subsubsection 3.3.1. This is achieved by measuring the temperature of a surface from which water has evaporated. Thus, it can be interpreted as the minimum temperature necessary for sweating being able to cool the body given current air temperature and humidity.

Physiologists have theorized a wet-bulb temperature of 35°C to be the limit to human adaptability to extreme heat for young and fit subjects (Sherwood and Huber, 2010; Raymond et al., 2020) and thus the threshold for survivability. While empirical work argues that thresholds need to be adjusted downwards and more heterogeneity needs to be accounted for (see, for instance, Vecellio et al., 2022), we take from this line of research that negative effects are already to be expected for WBT way below the extreme threshold of 35°C.

Furthermore, we use the simplified version of the index suggested by Steadman (1979) that relies only on air temperature and moisture, generally known as the "heat index". Steadman's index translates air temperature and air moisture into the temperature that humans would perceive if dew point temperature were 14°C. The heat index has become a widely used measure of thermal comfort in environmental health research (see Anderson et al., 2013), and it is used by the U.S. National Weather Service to indicate the level of threat to human health: "heat indices meeting or exceeding 103°F (39.4°C) [in the shade] can lead to dangerous heat disorders with prolonged exposure and/or physical activity in the heat".[5]

We compute the maximum daily heat index using hourly air temperature and relative humidity of each school day, then average over the school year.[6]

3.3.3 Wind Speed

Wind causes heat loss and thus helps to mitigate the adverse effects of heat on the micro climate and human thermal comfort. The mitigating effect of wind on heat has been particularly highlighted for urban areas where the loss of vegetation has led to higher level of air and surface temperature in comparison to rural areas. This leads to UHI as increased ambient temperature of urban areas due to warmer surfaces (see Synnefa et al., 2007). Previous research has empirically shown that particularly for tropical regions, wind does play an important role: For instance, Erell et al. (2012) measure that in Singapore a wind velocity of 1–1.5 m/s creates cooling effect equivalent to a 2°C decrease in measured temperature. Thus, we match our school data with daily wind speed from the ERA5-Land data set, and average over school days.

3.3.4 Precipitation

Precipitation (from ERA5-Land) is calculated as cumulative precipitation over each school day, then averaged over all school days. Precipitation is considered as a relevant climatic factor for assessing the impact of heat in hot climates such as India. This is due to its "potential to naturally mitigate heat excess in buildings and cities by evaporative cooling; and as a primary source of water to artificially reproduce this cooling mechanism, particularly in the humid tropics and subtropics" (Diaz et al., 2015).

3.3.5 Air Pollution

As air pollution may correlate with temperature and other meteorological variables, we control for the average level of air quality. We obtain grid data on particle pollution from fine particulates (PM 2.5) from Van Donkelaar et al. (2021) and grid data on Carbon Monoxide and Nitrogen Dioxide from the NASA Earth Observations database.[7]

[5] https://www.weather.gov/ama/heatindex. Accessed May 2024.

[6] We use the algorithm provided by Anderson et al. (2013) in the R package *weathermetics*.

[7] See `https://neo.gsfc.nasa.gov/`, last accessed in October 2023.

3.3.6 Tree Cover

The spatial resolution at which we measure temperature variation is rather low (≈ 100 km^2). However, there could be significant unobserved variation within each cell of the grid, driven by, for example, land use. This creates concerns as differences in vegetation and canopy height may systematically impact the micro-climate in the surrounding of schools.

The height and amount of the vegetation above ground level is a key component of the impact of temperature on the ground: forests function as a thermal insulator and thus, when ambient temperatures are hot, forests and in general trees cool the understory (see De Frenne et al., 2019). Furthermore, in urban areas tree canopy shading has been shown to alleviate the UHI effect and to increase thermal comfort (Iungman et al., 2023; Knight et al., 2021).

It is possible that the presence and extent of vegetation correlates with unobserved school characteristics that contribute to heat mitigation, or that influence cognitive performance. To address this threat to identification, we use data on *Global Forest Change (GFC)* from the *Global Land Analysis and Discovery Laboratory* (Hansen et al., 2013), to track variations in trees' density over time. The data, based on images from the Landsat satellite program, covers yearly gross forest cover loss since 2001, with a spatial resolution of 1 arc-second (about 30×30 meters at the equator).

India has the tenth largest forest area in the world FAO (2010) (see Figure 8), mostly concentrated in the Eastern and North-Eastern regions (mostly tropical moist and tropical dry forests) and in the Southern regions (tropical moist forest and rain forest) (FAO, 2012).

These forests mean a rich natural resource of global importance. The platform *Global Forest Watch (GFW)*[8] monitors various changes in forests globally and too use, inter alia, GFC data as input. According to these data, India has lost a sizable amount of its forest (2.29 Mha between 2001 to 2023 whereas only 38.1 kha thereof were lost due to fires but the vast majority due to any other reason including deforestation). This rate had been larger over the past 10 years as compared to the preceding decade. At the same time, however, India has also gained substantial amounts of forest over the same period. Overall, the net effect of changes in forest cover is positive between 2000 to 2020: 874 kha (1.3%) in tree cover.

Although the GFC data has been extensively used to study deforestation (see, for instance, Li et al., 2016; **?**), this is not the main concern in India as discussed above. Thus, we are not interested in measuring the implications of large-scale, fast-paced forest loss, yet we focus on heterogeneity within the country. The high spatial resolution of the data allows us to measure small-scale fluctuations in the extent of tree coverage in the proximity of schools.

As the GFC data is only available on a solar year basis, we cannot perfectly overlap the school years to the yearly changes in tree cover. We use the forest loss from the solar year that mostly overlaps with the school-year, and that precedes the end of the school-year: we match tree

[8] See `https://www.globalforestwatch.org` for details; last accessed in May 2024,

cover loss from 2014 with observations in the school year 2014-15, and so on.

We measure the extent of tree cover, in hectares, within a circle of 0.5 km (1.0 km; 0.25 km) centered around each school. We use the forest cover of 2012 as the benchmark, and compute yearly values for each school by subtracting the yearly loss of forest in the relevant area. Hansen et al. (2013) codify forest loss as binary, and they define it as a "stand-replacement disturbance, or a change from a forest to non-forest state".

4 The Effect of Heat Exposure on Test Results

4.1 Empirical Strategy

Exam results are reported in the broad categories of 'fail', 'pass' and 'pass with distinction.' A large majority of students does indeed pass the exam. We therefore estimate two separate logit models that distinguish 'fail' versus 'pass' and 'fail/pass' versus 'pass with distinction'.

To ensure clean measurement, we account for any time-invariant school-, or more generally speaking, location-specific aspects by estimating a fixed-effects model. Thus, we rely on temperature variation within each school-location over four academic-years (2014-15 to 2017-18).

This general set-up yields two logit regressions modeling the outcome of passing the exam y^p and passing the exam with distinction y^{pd}, respectively. For both outcomes $y \in \{p, pd\}$ following form:

$$Pr\{y_{ist} = 1|\text{temperature}_{st}, X_{ist}, \alpha_s, \lambda_t, \beta, \gamma\} = \textbf{logit}(\beta\ \text{temperature}_{st} + X_{ist}\gamma + \alpha_s + \lambda_t), \quad (1)$$

whereby y^p_{ist} is equal to 1 if student i in school s, year t passes the exam and 0 otherwise, and accordingly y^{pd}_{ist} for the outcome "pass with distinction." X_{ist} contains an optional set of further meteorological controls with associated parameters γ. The models include fixed-effects α_s to account for any school-specific (and by that also location-specific) effects as well as time fixed-effects λ_t relating to school-years. We use the algorithm suggested by Stammann et al. (2016) to estimate fixed effects logit models with large panel data, at low computational cost.

In both cases, the identifying assumption is

$$E(\varepsilon_{ist}|\text{temperature}_{st}, \alpha_s) = 0,$$

i.e., the error term is conditionally independent of our respective temperature measure and school-fixed effects.

4.2 Cumulative Heat Exposure

We start by assessing the impact of elevated heat during the school year ('cumulative heat exposure') on test results at the end of the year. For that, we measure the deviation from the long-term location- and season-specific average temperature during the entire school-year as detailed in subsection 3.2.

Table 2 reports results. Overall we find that with higher measured long-term average temperatures at the school location, the odds of students passing the exam as well as the odds of passing with distinction decreases significantly. In a model neglecting any other meteorological conditions, the decrease in log odds is -0.134 for passing and -0.468 for passing with distinction.

We add the battery of further meteorological indicators discussed in subsection 3.3 as controls. As expected, the likelihood to pass the exam (with distinction) decreases with increasing humidity. The effect of wind and precipitation, as well as the direction of the interacted effect of temperature and humidity are not unambiguously clear.

Our preferred models (4) and (8) also control for measured air pollution to cleanly disentangle adverse effects stemming from temperature shocks and pollution, respectively. By accounting for relative humidity, wind, precipitation and air pollution, the effect size of temperature remains negative, large in magnitude and highly statistically significant: the log ratios are -0.521 for passing and -1.160 for passing with distinction, respectively. This means that a 1°C increase in temperature is predicted to decrease the odds of passing by 40.6%, and the odds of getting a distinction by 68.7%.

Table 2: Cumulative Heat Exposure.

	Pass				Pass with Distinction			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Temperature	-0.134***	-0.425***	-0.472***	-0.521***	-0.468***	-1.190***	-1.444***	-1.160***
	(0.0094)	(0.0232)	(0.0503)	(0.0237)	(0.0020)	(0.0049)	(0.0101)	(0.0051)
Humidity		-0.063***	-0.088***	-0.084***		-0.169***	-0.309***	-0.170***
		(0.0042)	(0.0243)	(0.0043)		(0.0009)	(0.0050)	(0.0010)
Wind		-1.048***	-1.045***	-0.785***		-1.488***	-1.480***	-1.495***
		(0.0479)	(0.0479)	(0.0489)		(0.0103)	(0.0103)	(0.0105)
Precipitation		-0.267***	-0.266***	-0.249***		0.023***	0.032***	0.061***
		(0.0077)	(0.0079)	(0.0080)		(0.0016)	(0.0017)	(0.0017)
Temperature * Humidity			0.001				0.004***	
			(0.0007)				(0.0002)	
Air pollution	N	N	N	Y	N	N	N	Y
School FE	Y	Y	Y	Y	Y	Y	Y	Y
School year dummy vv.	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,780,912	2,780,912	2,780,912	2,778,323	19,915,599	19,915,599	19,915,599	19,900,270
BIC	1,386,808	1,384,958	1,384,972	1,382,881	21,910,394	21,865,591	21,864,784	21,835,802

Note: The table reports estimation results for Logit model with fixed effects expressed as log odds. Standard errors are reported in parentheses. Statistical significance is coded following the standard notation: *p<0.05; **p<0.01; ***p<0.001

To better understand the severity of the impact of temperature shocks on educational outcomes, we express this effect in absolute numbers of students affected. Note that we hypothetically increase average temperature over an extended period of time thus even small changes mean in practice a substantial temperature increase. These increases are thus also in-line with the concept of heat waves.

Table 3: Counterfactual Analysis: Increased Temperature

Not Passing

θ	Empirical Probability $\sum \hat{y}^{\theta}_{ist}$ / N^p	# Not Passing	$\Delta^{\neg p}(\theta)$	% Change [p.p.] Not passing (passing)
0	0.9074	257,190	-	-
0.25	0.8992	279,936	22,746	8.8 (-0.9)
0.5	0.8905	304,230	47,040	18.3 (-1.9)
1	0.8712	357,623	100,433	39.1 (-4.0)

Not Passing with Distinction

θ	Empirical Probability $\sum \hat{y}^{\theta}_{ist}$ / N^{pd}	# Not Passing w/ Distinction	$\Delta^{\neg pd}(\theta)$	% Change [p.p.] Not passing (passing)
0	0.6937	6,094,832	-	-
0.25	0.6441	7,081,670	986,838	16.2 (-7.1)
0.5	0.5918	8,123,025	2,028,193	33.3 (-14.7)
1	0.4832	10,284,193	4,189,361	68.7 (-30.3)

Notes: The table reports the simulated excess number of students not passing or nor passing with distinction in case of an increase in the temperature of θ degrees measured at the respective school locations. Shares relate to the total student populations of $N^p = 2,778,323$ and $N^{pd} = 19,900,270$, respectively.

Concretely, we compute the number of students that are expected to fail or miss a distinction due to adverse temperature conditions via a counterfactual analysis. Therefore, we test changes in outcomes for an increase in location-specific temperature during the academic year by θ degrees, i.e.,

$$T_{st} = \text{temperature}_{st} + \theta.$$

We use the estimated coefficients from models (4) and (8) and compute hypothetical outcomes $\hat{y}^{\theta}_{ist}$ via plug-in estimators.

We compare the predicted number of students that fail or miss the distinction for the observed temperature as well as hypothetically increased temperature. Overall, we observe outcomes for $N^p = 2,778,323$ and $N^{pd} = 19,900,270$ students, respectively, over the four school years under consideration.[9] We compute the excess numbers of students failing $\Delta^{\neg p}(\theta)$ and the excess number of students missing the distinctions $\Delta^{\neg pd}(\theta)$ given an increase in temperature of θ degrees via

$$\Delta^{\neg p}(\theta) = \sum_{i=1}^{N^p} \left(\hat{y}_{ist} - \hat{y}^{\theta}_{ist}\right) \quad \text{and} \quad \Delta^{\neg pd}(\theta) = \sum_{i=1}^{N^{pd}} \left(\hat{y}_{ist} - \hat{y}^{\theta}_{ist}\right). \tag{2}$$

[9] The numbers for passing p and passing with distinction pd differ as we observe a rather large number of schools where all students did pass the exam in all years considered. In case of no variation in the data, we are unable to retrieve the likelihood of failing from our data.

Table 3 reports the results for $\theta \in \{0, 0.25, 0.5, 1\}$. An increase in temperature by merely 0.5°C means a drop in the number of students passing the exam by 1.9% and a drop in distinctions of 14.7%. In absolute numbers, this means that 47,040 and 2,028,193, respectively, would achieve less favorable results and will thus likely not live up to their full potential.

We support these results with a battery of robustness and sensitivity checks, reported in Appendix B. First, we report the same results using a different approach to measuring heat exposure: while our main results rely on measuring elevated heat during the entire academic year preceding the exam, the alternative approach reported in subsection 4.4 tests whether elevated heat during the exam period itself has similar negative effects on preparing and taking the exam. Second, we test whether our results are sensitive to technical choices and specifications. Both types of robustness checks confirm our main results and conclusions.

4.3 Heterogeneity Analysis

While these overall results appear remarkable, we further investigate whether there are any groups that are particularly at risk to loose out due to exogenous adverse shocks during their educational careers. We do so by adding interactions between specific students or school characteristics and temperature. We assess potential heterogeneities by gender and school location (urban versus rural). We perform this analysis for elevated temperatures over the academic year.

First, Table 4 reports results for gender: the average effect of -0.521 we found in the main model (4) reported in Table 2 appears to be symmetrical for girls and boys. Boys are less strongly affected by hot temperatures than girls as indicated by the positive interaction effect. However, the magnitude of the coefficient is very small: an increase in temperature by 1°C is predicted to decrease the odds of passing for boys by 40.43%, and for girls by 40.84%. The probability of passing the exam is not significantly related to the student's gender.

However, relatively to the probabilities of passing with distinction, we find a general lower likelihood for boys to pass with distinction, and boys appear to be more negatively affected by elevated temperatures than girls, even though the effect does not appear to be economically significant.

Second, we check for differences by school location as we believe that the UHI effect is a major driver for the adverse heat effect on educational outcomes. Therefore, we make use of an identifier provided by schools themselves that classify their location as either 'rural' or 'urban'. Due to the fact that at higher temperatures urban areas heat up more than rural areas, we expect that the same observed increase in temperature effectively hits differently between urban and rural location.

Table 4 reports results for cumulative heat exposure: opposite to our expectations, elevated temperatures are more harmful in rural areas as indicated by the negative interaction effect between temperature and the rural identifier. An increase in temperature by 1°C is predicted

Table 4: Heterogeneity Analysis of Cumulative Heat Exposure.

	Pass		Pass with Distinction	
Temperature	-0.525***	-0.381***	-1.159***	-1.145***
	(0.0238)	(0.0254)	(0.0024)	(0.0054)
Boy	-0.180		-0.051*	
	(0.1066)		(0.0243)	
Temperature*Boy	0.007*		-0.006***	
	(0.0035)		(0.0008)	
Temperature*Rural		-0.227***		-0.028***
		(0.0150)		(0.0032)
Humidity	-0.083***	-0.085***	-0.170***	-0.170***
	(0.0043)	(0.0043)	(0.0009)	(0.0009)
Wind	-0.783***	-0.787***	-1.500***	-1.498***
	(0.0488)	(0.0489)	(0.0105)	(0.0105)
Precipitation	-0.249***	-0.247***	0.061***	0.061***
	(0.0080)	(0.0080)	(0.0017)	(0.0017)
Air pollution	Y	Y	Y	Y
School FE	Y	Y	Y	Y
School year dummy vv.	Y	Y	Y	Y
Observations	2,778,323	2,778,323	19,900,270	19,900,270
BIC	1,382,802	1,382,663	21,801,629	21,835,734

Notes: Standard errors are reported in round bracket. Statistical significance is coded following the standard notation: *p<0.05; **p<0.01; ***p<0.001

to decrease the odds of passing for students in urban schools by 31.7%, against the decrease in odds of 45.6% for students in rural areas. The effect on the odds of getting a distinction is also statistically significant, but not as relevant in magnitude: the decrease in odds is 68.2% for urban schools, against 69.1% for rural schools.[10]

Whether a school is located in a rural or urban area most likely correlates with unobserved attributes of school (or location). The disproportionate impact of temperature on rural schools might be explained by differences in economic status, that are reflected in uneven availability of coping strategies.

4.4 Cumulative versus Immediate Heat Exposure

So far, we have focused on cumulative exposure to heat over a school year. However, there might also be a negative effect of excess heat during the exam period. We thus test whether heat exposure exclusively in the month March has a similar effect on students' exam results ('immediate heat exposure').[11]

We find again that with higher measured temperature during the immediate period before and during the exam, the odds of students passing and passing with distinction decreases significantly. Neglecting any other meteorological conditions, the decrease in log odds is -0.028 for passing and -0.071 for passing with distinction (Table 5).

We add again meteorological controls (measure also for the same shortened period) and, as expected, the likelihood to pass and pass with distinction decreases with increasing humidity. The effect of wind and precipitation, as well as the direction of the interacted effect of temperature and humidity, however, are ambiguous.

When adding our measure for air pollution, the effect sizes for other meteorological measures loose in importance (suggesting that air pollution is not orthogonal to meteorological measures) yet the main effect of temperature remains negative, large in magnitude and highly statistically significant: the log ratios are -0.292 for passing and -0.312 for passing with distinction, respectively.

This clearly documents disadvantages for children being exposed to extreme heat during this crucial exam period.

We repeat the counterfactual analysis reported in subsection 4.2 yet rely on larger temperature increases as we assess a short time period only. Results are reported in Table 6.

An increase in temperature by merely 0.5°C during the end of the school year means a drop in the number of students passing the exam by 1.0% and a drop in distinctions of 3.8%. In absolute numbers, this means that 25,554 and 522,667, respectively, would achieve less favorable results

[10]Note that main effects are absorbed by school fixed effects as location is a time-invariant characteristic of schools.

[11]This, however, is not our preferred approach as we aim to measure the exposure to heat during the period of learning. Further, exams do not take place at the very same day in all schools and we only know the exam period overall.

Table 5: Immediate Heat Exposure.

	Dependent variable							
	Pass				Pass with Distinction			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Temperature	-0.028***	-0.330***	-0.280***	-0.292***	-0.071***	-0.235***	-0.331***	-0.312***
	(0.0040)	(0.0093)	(0.0113)	(0.0106)	(0.0009)	(0.0020)	(0.0026)	(0.0023)
Humidity		-0.070***	-0.007	-0.066***		-0.002***	-0.117***	-0.006***
		(0.0021)	(0.0085)	(0.0022)		(0.0005)	(0.0019)	(0.0005)
Wind		-0.016	-0.001	0.028**		-0.187***	-0.205***	-0.171***
		(0.0090)	(0.0092)	(0.0093)		(0.0020)	(0.0020)	(0.0020)
Precipitation		0.002	-0.025**	0.044***		-0.337***	-0.288***	-0.354***
		(0.0082)	(0.0090)	(0.0086)		(0.0019)	(0.0020)	(0.0019)
Temperature * Humidity			-0.002***				0.004***	
			(0.0003)				(0.0001)	
Air pollution	N	N	N	Y	N	N	N	Y
School FE	Y	Y	Y	Y	Y	Y	Y	Y
School year dummy vv.	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,780,912	2,780,912	2,780,912	2,770,243	19,915,599	19,915,599	19,915,599	19,865,438
BIC	1,386,962	1,385,701	1,385,656	1,379,877	21,960,918	21,910,480	21,906,649	21,837,523

Note: The table reports estimation results for Logit model with fixed effects expressed as log odds. Standard errors are reported in parentheses. Statistical significance is coded following the standard notation: *p<0.05; **p<0.01; ***p<0.001

Table 6: Counterfactual Analysis: Increased Temperature in March

Passing

θ	Empirical Probability $\sum \hat{y}^{\theta}_{ist}$ / N^p	# Not Passing	$\Delta^{\neg p}(\theta)$	% Change [p.p.] Not passing (passing)
0	0.9074	256,447	-	-
0.25	0.9029	268,984	12,537	4.9 (-0.5)
0.5	0.8982	282,001	25,554	10.0 (-1.0)
1	0.8883	309,504	53,057	20.7 (-2.1)

Passing with Distinction

θ	Empirical Probability $\sum \hat{y}^{\theta}_{ist}$ / N^p	# Not Passing	$\Delta^{\neg p}(\theta)$	% Change [p.p.] Not passing (passing)
0	0.6943	6,072,089	-	-
0.25	0.6813	6,330,951	258,862	4.3 (-1.9)
0.5	0.6680	6,594,756	522,667	8.6 (-3.8)
1	0.6408	7,135,872	1,063,782	17.5 (-7.7)

Notes: The table reports the simulated excess number of students not passing or nor passing with distinction in case of an increase in the temperature of θ degrees measured at the respective school locations.

and will thus likely not live up to their full potential.

Thus, we can conclude that heat exposure at the exam period itself has a similar negative effect of students' performance in the exam.

5 Adaptation and Mitigation

After documenting the adverse impact of heat on exam results, we discuss and test the effectiveness of a potential mitigation strategy. Previous research has highlighted the importance of air conditioning in mitigating the effect of heat on learning (Park et al., 2020). Here, we investigate the impact of an alternative "natural" strategy relying on vegetation that could mitigate the negative effects of heat on learning.

So far, we have relied on relatively low resolution data to detect the impact of heat exposure: as the temperature data has a spatial resolution of $0.1° \times 0.1°$, each cell of the grid covers an area of approximately 100 km^2. Averaging over such a large area clearly introduces a noisy measure of the effective temperature experienced at each specific location. We can refine this measure by relying on the well-known fact that artificial structures tend to absorb and reflect heat more than the natural landscape, and temperatures in urban areas can be significantly higher than in their rural surroundings. Moreover, vegetation and tree cover density may systematically impact the microclimate in the surrounding of schools (Knight et al., 2021). In the following, we therefore investigate the effect of tree cover variation on the probability of passing the exam.

We rely on the time varying GFC data, as described in subsubsection 3.3.6. Only a relatively small share of our sample is treated by loss of forest in its surroundings: if we measure tree loss within a radius of 1km from each school, only about 5% of schools are treated. To limit the degree of imbalance, we use statistical matching as a data pre-processing step (see Stuart, 2010, for an overview) and by that carefully prune our sample of students using a coarsened exact matching approach (Iacus et al., 2012). Besides enhancing balance, we match to address what we think is the main threat to our identification: tree cover loss is not strictly exogenous such as weather, and it may correlate with local economic development, or it may be caused by large scale events such as natural disasters, which by themselves could have an effect on education. By matching schools on a proximity basis, we can reasonably make sure that both treated and control schools would be affected by these latent drivers of tree cover loss.

Our identification strategy predominantly relies on observing schools over time. Thus, it is essential to keep students that – in any school-year we observe – attend schools that experience forest loss in their vicinity (precisely within a 0.25 km, 0.5 km or 1 km radius), i.e., all students that attend "treated" schools. Non-treated students, according to this definition, overwhelmingly outnumber treated ones. Thus, we can afford following a rather conservative approach in isolating the control group. The full (none-matched) control group consists of all the schools that never experience any tree cover loss within a radius of 2 km, in any of the school-years in our panel and also not in 2013-14 to avoid any potential lagged effects.

Summarizing, the control group is ultimately chosen such that each pair of treated and matched control schools

(i) is located in the same district;[12]

(ii) was surrounded by a similar amount of forest cover in 2014;[13]

(iii) has the same 'Public/Private' status; and

(iv) appears in the panel in the same school-years.

Although the matching characteristics are measured at the school level, we still match individual students ensuring an identical number of observations in the control and treatment groups. If there are multiple possible choices, ties are broken at random.

Table 7 reports the result for the 0.5 km radius. Since we are only able to match a small sample of students, we report results for both, matched observations (columns 1 and 3), as well as for the full sample (columns 2 and 4). Within-school tree cover loss over time has a statistically significant negative effect both on the probability of passing and the probability of passing with distinction.

From columns (1) and (3) we conclude that an additional hectare of forest increases the odds of passing by 37.2%, and of getting a distinction by 12.5%.

Based on the average partial effects from the full sample results, 2.6 hectares (i.e., about 3.5 football pitches) of forest over an area of 78.54 hectares would offset the impact of an increase in temperature by 1°C, on the probability of passing the exam. Analogously, 9.5 hectares of forest would be needed to offset the impact of the same temperature change on the probability of achieving a distinction.

Results for the 1 km and the 0.25 km radii are reported in Table 10 and Table 11 in the Appendix.

6 Summary and Conclusions

In this article, we empirically assess the link between exposure to heat and schooling success in India. Precisely, we link results in the exam that concludes compulsory education reported in the categories 'Fail', 'Pass' and 'Pass with Distinction' to temperatures measured during the preceding school-year or the month of the exam at the school location. For isolating this

[12]Districts are an administrative subdivision of a State. The largest state in our sample, Madhya Pradesh, has 55 district. The smallest state, Kerala, has 14.

[13]The distribution of tree cover is highly skewed, with most schools having none or almost no tree cover in their surroundings. We coarse the 2014 baseline forest cover over an area of 314.16 hectares (i.e., within 1 km from the school) into four categories: 0 to 1 hectares, 1 to 10 hectares, 10 to 100 hectares, 100 to 314.16 hectares. We use the same cut-off points for the 0.5 km (78.54 ha) and 0.25 km (19.63 ha) distances, scaled by area. We choose rather large intervals to avoid pruning an excessive number of observations. Even so, the matching largely reduces imbalance, as shown in Figure 5 in the appendix.

Table 7: Tree Cover Effect, 0.5 km radius.

	Pass		Pass with Distinction	
	(1)	(2)	(3)	(4)
Temperature	-6.386***	-0.437***	-0.192**	-1.156***
	(0.5869)	(0.0246)	(0.0645)	(0.0052)
	[-0.3360]	[-0.0036]	[-0.0356]	[-0.1766]
Forest 0.50 km	0.778***	0.167***	0.073***	0.121***
	(0.1188)	(0.0239)	(0.0177)	(0.0064)
	[0.0410]	[0.0014]	[0.0136]	[0.0185]
CEM	**Y**	**N**	**Y**	**N**
Humidity/wind/precipitation	Y	Y	Y	Y
Air pollution	Y	Y	Y	Y
School FE	Y	Y	Y	Y
School year dummy vv.	Y	Y	Y	Y
Observations	26,734	2,640,460	298,957	18,884,151

Note: Columns (1) and (3) refer to matched observations and columns (2) and (4) to the full sample. Average partial effects are reported in square brackets. Standard errors are reported in round bracket. Statistical significance is coded following the standard notation: *p<0.05; **p<0.01; ***p<0.001.

effect, we make use of the vast variation in weather conditions observed across India and over our period of study (2014 to 2018). Our identification strategy hence relies on the repeated observation of the same school (and thus the same location) over time. Baseline results not controlling for any amplifying or mitigating factors suggest a statistically highly significant and large negative impact of high temperatures on exam results. Students exposed to elevated heat are less likely to both, passing the exam and passing the exam with distinction.

As high temperatures are particularly dangerous when combined with high levels of humidity and in the absence of cooling factors such as wind speed or precipitation, we add further meteorological information, which leads to effects pointing in the same direction yet being even larger in magnitude. We translate log odds into the number of students concerned and find that 257,190 students fail the exam because of elevated heat exposure and even 6,094,832 miss the distinction.

We use our estimation results to simulate the impact of hypothetically increasing temperatures on the excess number of students with worse test results induced by this change. Even small increases are associated with large increases in the number of failing students. These dramatically increases in the number of students not performing well in this early formal examination which likely adversely affects their future school career and success in life.

Furthermore, we investigate the impact of a possible mitigation strategy relying on increased vegetation surrounding schools. For that, we quantify the effect of the current level of vegetation (measured via tree cover density in the direct neighborhood of a school) on the likelihood of passing or passing with distinction. We detect that relying on natural cooling in the form of increased vegetation can mitigate some of the advert heat effects: depending on the chosen outcome, between 2.6 and 9.5 hectares of forest within a radius of 0.5 km from a school would offset the consequences of an increase in temperature of 1°C on schooling outcomes. This suggests that policies incentivising increased vegetation in the proximity of schools could have a substantial positive impact on educational outcomes. This result suggests that such policies could mean a viable adaptation strategy to counteract the risks climate change poses to the development of human capital.

References

Agarwal, S., Chia, L. E., and Ghosh, P. (2022). Do sanitary pads alleviate period poverty and improve girls' educational outcomes? *Available at SSRN 4189740*.

Alduchov, O. A. and Eskridge, R. E. (1996). Improved magnus form approximation of saturation vapor pressure. *Journal of Applied Meteorology and Climatology*, 35(4):601 – 609.

Anderson, G. B., Bell, M. L., and Peng, R. D. (2013). Methods to calculate the heat index as an exposure metric in environmental health research. *Environ Health Perspectives*, 121(10):1111–1119.

Blum, N. (2009). Small NGO schools in India: Implications for access and innovation. *Compare*, 39(2):235–248.

Cho, H. (2017). The effects of summer heat on academic achievement: A cohort analysis. *Journal of Environmental Economics and Management*, 83:185–196.

Conte Keivabu, R. (2024). Temperature and school absences: evidence from England. *Population and Environment*, 46(6).

Currie, J., Hanushek, E. A., Kahn, E. M., Neidell, M., and Rivkin, S. G. (2009). Does pollution increase school absences? *The Review of Economics and Statistics*, 91(4):682–694.

De Frenne, P., Zellweger, F., Rodríguez-Sánchez, F., Scheffers, B. R., Hylander, K., Luoto, M., Vellend, M., Verheyen, K., and Lenoir, J. (2019). Global buffering of temperatures under forest canopies. *Nature Ecology & Evolution*, 3(5):744–749.

Debnath, R., Bardhan, R., and Bell, M. L. (2023). Lethal heatwaves are challenging india's sustainable development. *PLOS climate*, 2(4):e0000156.

Diaz, C. A., Osmond, P., and King, S. (2015). Precipitation and buildings: estimation of the natural potential of locations to sustain indirect evaporative cooling strategies through hot seasons. page 45–54. 49th International Conference of the Architectural Science Association 2015.

Donaldson, D. and Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4):171–198.

Ebenstein, A., Lavy, V., and Roth, S. (2016). The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics*, 8(4):36–65.

Erell, E., Pearlmutter, D., and Williamson, T. (2012). *Urban microclimate: designing the spaces between buildings*. Routledge, New York, USA.

FAO (2010). Global forest resources assessment 2010. Technical report, FAO Forestry Paper No. 163. Rome.

FAO (2012). Global ecological zones for fao forest reporting: 2010 update. Technical report, Forest Resources Assessment Working Paper 179.

Graff Zivin, J., Song, Y., Tang, Q., and Zhang, P. (2020). Temperature and high-stakes cognitive performance: Evidence from the national college entrance examination in China. *Journal of Environmental Economics and Management*, 104(102365).

Han, L., Heblich, S., Timmins, C., and Zylberberg, Y. (2024). Cool cities: the value of urban trees. *NBER Working Paper Series*, Nr. 32063.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., and Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160):850–853.

Hill, S. and Chalaux, T. (2011). Improving access and quality in the Indian education system. *OECD Economic Department Working Papers*, No. 885.

Iacus, S. M., King, G., and Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 20(1):1–24.

Iungman, T., Cirach, M., Marando, F., Barboza, E. P., Khomenko, S., Masselot, P., Quijal-Zamorano, M., Mueller, N., Gasparrini, A., Urquiza, J., Heris, M., Thondoo, M., and Nieuwenhuijsen, M. (2023). Cooling cities through urban green infrastructure: a health impact assessment of European cities. *The Lancet*, 401(10376):577–589.

Kingdon, G. G. (2020). The private schooling phenomenon in india: A review. *The Journal of Development Studies*, 56(10):1795–1817.

Knight, T., Price, S., Bowler, D., King, S., Konno, K., and Richter, R. L. (2021). How effective is ‘greening’ of urban areas in reducing human exposure to ground-level ozone concentrations, UV exposure and the ‘urban heat island effect’? an updated systematic review. *Environmental Evidence*, 10.

Li, Y., Zhao, M., Mildrexler, D. J., Motesharrei, S., Mu, Q., Kalnay, E., Zhao, F., Li, S., and Wang, K. (2016). Potential and actual impacts of deforestation and afforestation on land surface temperature. *Journal of Geophysical Research: Atmospheres*, 121(24):14–372.

Machin, S., McNally, S., and Ruiz-Valenzuela, J. (2020). Entry through the narrow door: The costs of just failing high stakes exams. *Journal of Public Economics*, 190:104224.

McCormack, K. (2023). Education under extremes: Temperature, student absenteeism, and disciplinary infractions. *Mimeo.*

Muñoz Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N. (2021). Era5-land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9):4349–4383.

Palacios, J., Eichholtz, P., Kok, N., and Duran, N. (2022). Indoor air quality and learning: evidence from a large field study in primary schools. *MIT Center for Real Estate Research Paper*, (22/13).

Papay, J. P., Murnane, R. J., and Willett, J. B. (2010). The consequences of high school exit examinations for low-performing urban students: Evidence from massachusetts. *Educational Evaluation and Policy Analysis*, 32(1):5–23.

Park, R. J., Behrer, A. P., and Goodman, J. (2021). Learning is inhibited by heat exposure, both internationally and within the United States. *Nature Human Behaviour*, 5:19–27.

Park, R. J., Goodman, J., Hurwitz, M., and Smith, J. (2020). Heat and learning. *American Economic Journal: Economic Policy*, 12(2):306–39.

Raymond, C., Matthews, T., and Horton, R. M. (2020). The emergence of heat and humidity too severe for human tolerance. *Science Advances*, 6(19):eaaw1838.

Rekha, R. S., Radhakrishnan, N., and Mathew, S. (2020). Spatial accessibility analysis of schools using geospatial techniques. *Spatial Information Research*, 28:699–708.

Sanjay, J., Revadekar, J., Ramarao, M., Borgaonkar, H., Sengupta, S., Kothawale, D., Patel, J., Mahesh, R., Ingle, S., and AchutaRao, K. (2020). Temperature changes in India. In *Assessment of climate change over the Indian region: a report of the Ministry of Earth Sciences (MoES), Government of India*, pages 21–45. Springer.

Sherwood, S. C. and Huber, M. (2010). An adaptability limit to climate change due to heat stress. *Proceedings of the National Academy of Sciences*, 107(21):9552–9555.

Srivastava, P. and Noronha, C. (2014). Institutional framing of the right to education act: Contestation, controversy and concessions. *Economic and Political Weekly*, pages 51–58.

Stammann, A., Heiß, F., and McFadden, D. (2016). Estimating fixed effects logit models with large panel data. Number G01-V3. ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft.

Steadman, R. G. (1979). The assessment of sultriness. part i: A temperature-humidity index based on human physiology and clothing science. *Journal of Applied Meteorology and Climatology*, 18(7):861 – 873.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, 25(1):1–21.

Synnefa, A., Santamouris, M., and Apostolakis, K. (2007). On the development, optical properties and thermal performance of cool colored coatings for the urban environment. *Solar energy*, 81(4):488–497.

Van Donkelaar, A., Hammer, M. S., Bindle, L., Brauer, M., Brook, J. R., Garay, M. J., Hsu, N. C., Kalashnikova, O. V., Kahn, R. A., Lee, C., et al. (2021). Monthly global estimates of fine particulate matter and their uncertainty. *Environmental Science & Technology*, 55(22):15287–15300.

Vecellio, D. J., Wolf, S. T., Cottle, R. M., and Kenney, W. L. (2022). Evaluating the 35 c wet-bulb temperature adaptability threshold for young, healthy subjects (psu heat project). *Journal of Applied Physiology*, 132(2):340–345.

Wehner, M., Stone, D., Krishnan, H., Achuta Rao, K., and Castillo, F. (2016). The deadly combination of heat and humidity in India and Pakistan in summer 2015. *Bulletin of the American Meteorological Society*, 97(12):30–32.

Yin, B., Fang, W., Liu, L., Guo, Y., Ma, X., and Di, Q. (2024). Effect of extreme high temperature on cognitive function at different time scales: A national difference-in-differences analysis. *Ecotoxicology and Environmental Safety*, 275:116238.

Appendix

A School Data

A.1 School Selection Procedure

The *DISE* database covers all registered primary and secondary schools in in 23 states and territories.[14] We use this database as the starting point, which then undergoes a step-wise selection process to ensure comparability and ultimately clean estimation. The step-wise selection procedure can be summarized as follows:

- Include public schools if managed by
  - + the Department of Education,
  - + the Tribal/Social Welfare Department,
  - + a local body, or
  - + the central government.
- Include private schools if managed by
  - + a private institution, or
  - + a private institution, yet aided by the government.
- Exclude schools
  - – not starting the school year in June,
  - – classified as Madrasa,
  - – classified as unrecognized,
  - – classified as 'other', or
  - – schools with inconsistent classification over time.

The rationales driving decisions to drop schools are the following: First, we exclusively keep schools that roughly follow the same academic year. This is important to precisely link meteorological data to the academic year and compare like with like. The majority of schools start in June, and ends in March or early April. These are schools located in the states Andhra Pradesh, Arunachal Pradesh, Gujarat, Kerala, Madhya Pradesh, Maharashtra, and Tamil Nadu. We thus exclude schools in other states. Further, schools located in these states yet reporting (at least in some years) a different beginning of the academic year.

[14] Andhra Pradesh, Arunachal Pradesh, Assam, Bihar, Chandigarh, Chhattisgarh, Gujarat, Haryana, Himachal Pradesh, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, West Bengal.

Next, we exclude institutions classified as Madrasa (islamic education) and drop schools that we cannot unambiguously classify as either managed by the government, or managed by a private institution with the aid or consensus of the government. Similarly, we exclude schools that over the observed period (2014-2018) are not consistently classified as either a private or a public institution. We exclude schools that are not consistently classified as either rural or urban schools, as “boys-only”, “girls-only” or “co-educational”, whose stated main language of instruction changes, or that change affiliation to the board administering the secondary school examinations. We drop schools that are not reported in the data every year of the panel.

This selection process yields ultimately a set of 374,598 well comparable schools, which means a drop from 1,567,741 in the original DISE database. Figure 1 shows the geographic distribution of schools entering our analysis.

A.2 Geo-coding School Data

We determine the location of schools via the address stated in the DISE data set. We geo-code this address using the Google Geocoding API to obtain the latitude and longitude coordinates.

We double check whether the geo-location matches the separately provided district information. Therefore, we retrieve district polygons from an official map of India and join it with the geo-coded coordinates. Accurate matches are labeled as “exact” while in case of a mismatch we assume that the district is correct and a random location within the district is assigned to the observation. The resulting geo-location is consequently labeled “random”. Only schools that have been “exactly” geolocated are used to study the mitigatory effect of vegetation (section 5).

B Robustness Checks and Sensitivity Analyses

This section performs a large battery of technical robustness checks and sensitivity analyses related to all parts of the study.

As a first check, we vary the model class used for measuring the effect of temperature on exam results. The main specification relies on logistic regression models that distinguishes between the categories ‘pass’ and ‘fail’, as well as ‘pass with distinction’ and ‘fail’ or ‘missing the distinction,’ respectively. We re-estimate the same models, but rely on ordinary least squares (i.e., a linear regression model) as estimation technique. The first and third columns in Table 8 report estimation results: the effect of temperature is still negative and highly significant.

Next, we assess the robustness of our results by changing how we incorporate the key variable of interest: temperature. Our main models include temperature as a continuous variable. Here, we allow for a more flexible functional form by including temperature as a sequence of temperature brackets. Models (2) and (4) include brackets of days with a maximum temperature within intervals of span 5°. Model (2) reveals that adverse effects on the probability of passing the exam are driven by high temperatures, with the strongest effect associated with temperatures above 40°C. The odds of passing the exam consistently decrease, for any additional day in a higher bracket, as clearly shown by Figure 2.

The effect on the probability of passing with distinction is very similar, as reported in Model (4) and Figure 2: the pattern is virtually identical, but the average marginal effects are larger by a factor of 10.

Table 9 reports the results for two alternative measures of thermal comfort: wet bulb temperature and the heat index. As detailed in subsection 3.3, these are often used as metrics for 'actual temperature', as they consider the combined effect of air temperature and air moisture. For both measures, the effect is similar, in magnitude and sign, to the effect of temperature alone, as expected. Additionally, Figure 3 reports the results from regressing the binary outcomes on the number of school days within a given heat index interval. We use four intervals that correspond to the U.S. National Weather Service classification of the level of risk associated with heat indices:

- Below 27°C (80°F): low risk.
- 27°C - 32°C (80°F - 90°F): "Caution: fatigue possible with prolonged exposure and/or physical activity."
- 32°C - 41°C (90°F - 105°F): "Extreme Caution: sunstroke, muscle cramps, and/or heat exhaustion possible with prolonged exposure and/or physical activity."
- 41°C - 54°C (105°F - 129°F): "Danger: sunstroke, muscle cramps, and/or heat exhaustion likely. Heatstroke possible with prolonged exposure and/or physical activity."[15]

Again, the adverse effects are driven by hot, humid days, with the strongest marginal effect identified for schooldays with temperatures above 41°C.

Overall, we find consistently negative effects for any additional hot day, and for any additional day that is hotter than usual. This is robust to the inclusion of different measures of thermal comfort, and is in-line with the results from our main specification.

B.1 Forest Cover Loss: 1 km and 0.25 km radii

Tables 10 and 11 report the results of using the same empirical approach of section 5, but measuring tree cover density within a 1 km and 0.25 km radius, respectively: that is, we select two alternative buffer areas around each school, one that four times larger, and one that is four times smaller than our baseline choice. The results are consistent, showing that loss of forest cover in the proximities of the schools has a negative effect on the probability of passing (with distinction). As expected, the magnitude of the average marginal effect of one additional hectare of forest (reported in square brackets) tends to decay with distance.

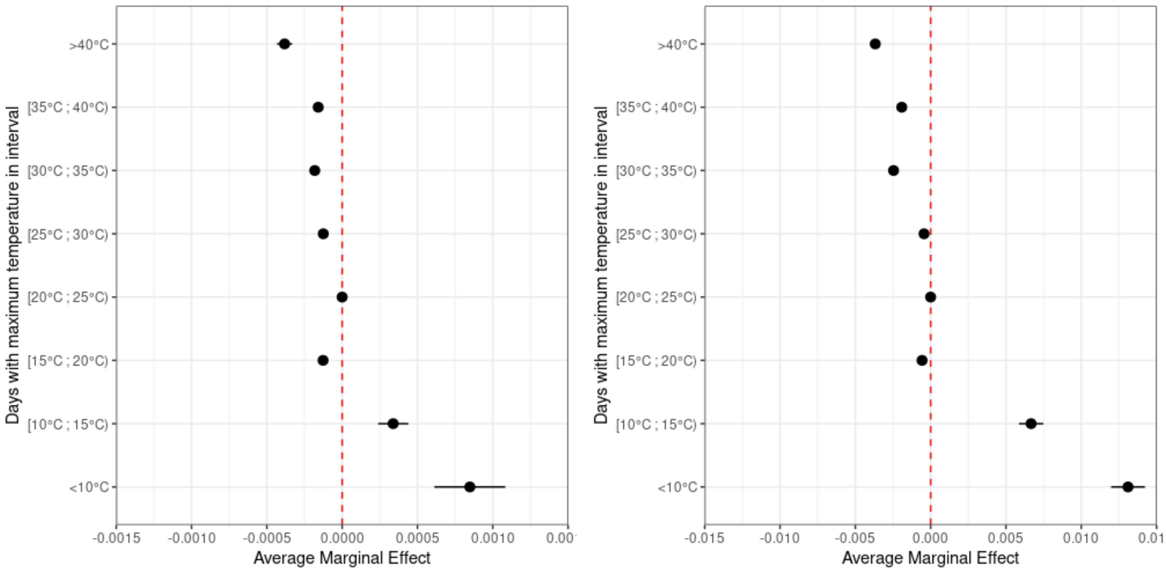
[15] See the Heat Index Chart of the U.S. National Weather Service: `https://www.weather.gov/ffc/hichart`, last accessed in June 2024. Heat indices above 54°C (130°F) are classified as "Extreme Danger", but such values are not observed in our sample.

Table 8: Robustness Checks – Cumulative Effects.

	Pass		Pass with Distinction	
	(1) *OLS*	(2) *Logit*	(3) *OLS*	(4) *Logit*
Temperature	-0.032* (0.0136)		-0.190*** (0.0055)	
Days below 10°C		0.105*** (0.0119)		0.085*** (0.0028)
Days 10-15°C		0.042*** (0.0052)		0.043*** (0.0019)
Days 15-20°C		-0.016*** (0.0013)		-0.004*** (0.0002)
Days 25-30°C		-0.016*** 0.0008		-0.003*** (0.0002)
Days 30-35°C		-0.023*** (0.0010)		-0.016*** (0.0002)
Days 35-40°C		-0.020*** (0.0014)		-0.012*** (0.0003)
Days above 40°C		-0.047*** (0.0024)		-0.024*** (0.0005)
Humidity/wind/precipitation	Y	Y	Y	Y
Air pollution	Y	Y	Y	Y
School FE	Y	Y	Y	Y
School year dummy vv.	Y	Y	Y	Y
Observations	2,780,912	2,554,345	19,915,599	18,493,322
Adjusted R^2	0.27		0.22	
BIC	345,636	1,264,838	22,716,787	21,916,805

Note: For the linear probability models, the standard errors are clustered at the school level. Standard errors are reported in parentheses. Statistical significance is coded following the standard notation: *p<0.05; **p<0.01; ***p<0.001

Figure 2: Marginal effect of an additional day in each temperature bin – Cumulative Effects.



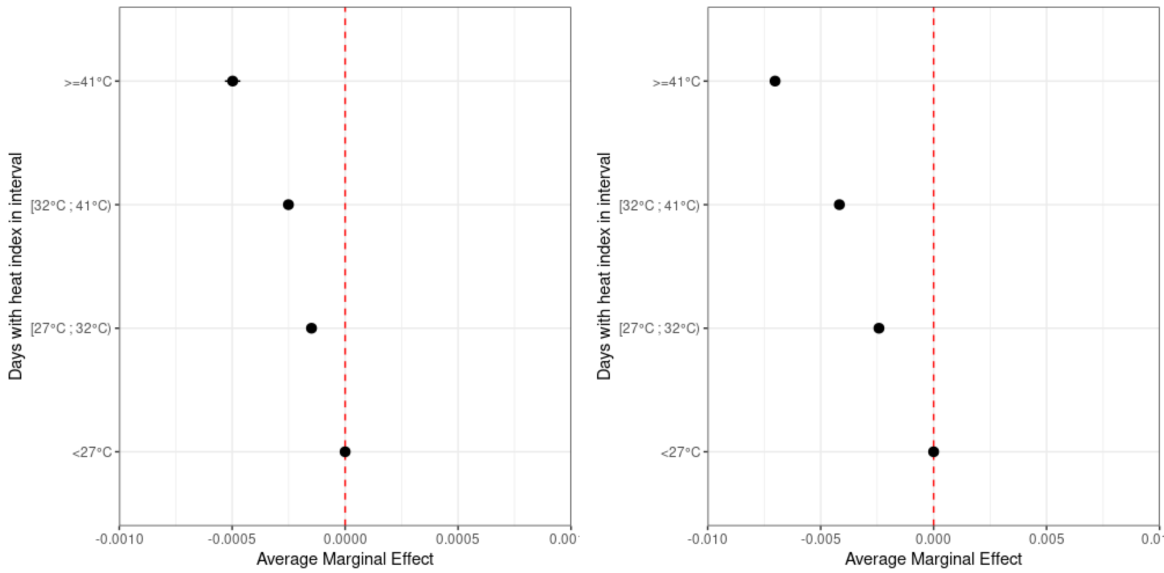
Notes: The figure shows the average marginal effect of an additional school day in each temperature bracket on the probability of passing the exam (left) and on the probability of passing with distinction (right).

Table 9: Wet Bulb Temperature and Heat Index.

	Pass		Pass with Distinction	
	(1)	(2)	(3)	(4)
Wet Bulb Temperature	-0.494***		-0.455***	
	(0.0219)		(0.0047)	
Heat Index		-0.258***		-0.332***
		(0.0139)		(0.0029)
Temperature	N	N	N	N
Humidity	N	N	N	N
Wind	Y	Y	Y	Y
Precipitation	Y	Y	Y	Y
Air pollution	Y	Y	Y	Y
School FE	Y	Y	Y	Y
School year dummy vv.	Y	Y	Y	Y
Observations	2,554,345	2,525,282	18,493,322	18,251,535
BIC	1,265,650	1,249,897	20,390,798	20,100,041

Notes: Standard errors are reported in round bracket. Statistical significance is coded following the standard notation: *p<0.05; **p<0.01; ***p<0.001

Figure 3: Marginal effect of an additional day in each heat index bin – Cumulative Effects.



Notes: The figure shows the average marginal effect of an additional school day in each heat index bracket on the probability of passing the exam (left) and on the probability of passing with distinction (right).

Table 10: Tree Cover Effect, 1 km Radius.

	Pass		Pass with Distinction	
	(1)	(2)	(3)	(4)
Temperature	-4.871***	-0.441***	-0.269***	-1.156***
	(0.3212)	(0.0247)	(0.0454)	(0.0052)
	[-0.2542]	[-0.0036]	[-0.0490]	[-0.1766]
Forest 1.00 km	0.210***	0.001	0.026***	0.023***
	(0.0368)	(0.0061)	(0.0060)	(0.0018)
	[0.0110]	[0.0000]	[0.0046]	[0.0035]
CEM	**Y**	**N**	**Y**	**N**
Humidity/wind/precipitation	Y	Y	Y	Y
Air pollution	Y	Y	Y	Y
School FE	Y	Y	Y	Y
School year dummy vv.	Y	Y	Y	Y
Observations	58,274	2,637,871	608,936	18,870,185

Note: Columns (1) and (3) refer to matched observations and columns (2) and (4) to the full sample. Average partial effects are reported in square brackets. Standard errors are reported in round bracket. Statistical significance is coded following the standard notation: *p<0.05; **p<0.01; ***p<0.001

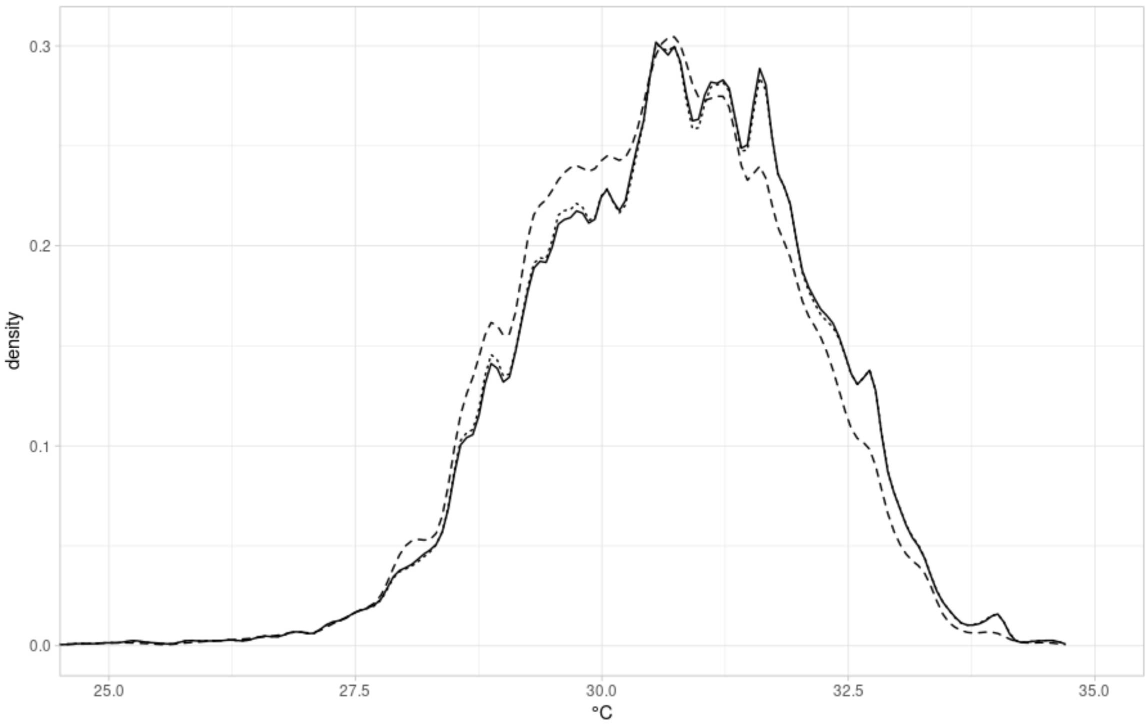
Table 11: Tree Cover Effect, 0.25 km Radius.

	Pass		Pass with Distinction	
	(1)	(2)	(3)	(4)
Temperature	-1.999*	-0.440***	-0.599***	-1.157***
	(0.8319)	(0.0247)	(0.1036)	(0.0052)
	[-0.0970]	[-0.0036]	[-0.0380]	[-0.1767]
Forest 0.25 km	1.244***	0.293***	-0.019	0.101***
	(0.2507)	(0.0061)	(0.0344)	(0.0143)
	[0.0604]	[0.0024]	[-0.0035]	[0.0155]
CEM	**Y**	**N**	**Y**	**N**
Humidity/wind/precipitation	Y	Y	Y	Y
Air pollution	Y	Y	Y	Y
School FE	Y	Y	Y	Y
School year dummy vv.	Y	Y	Y	Y
Observations	12,433	2,637,871	141,008	18,870,185

Note: Columns (1) and (3) refer to matched observations and columns (2) and (4) to the full sample. Average partial effects are reported in square brackets. Standard errors are reported in round bracket. Statistical significance is coded following the standard notation: *p<0.05; **p<0.01; ***p<0.001

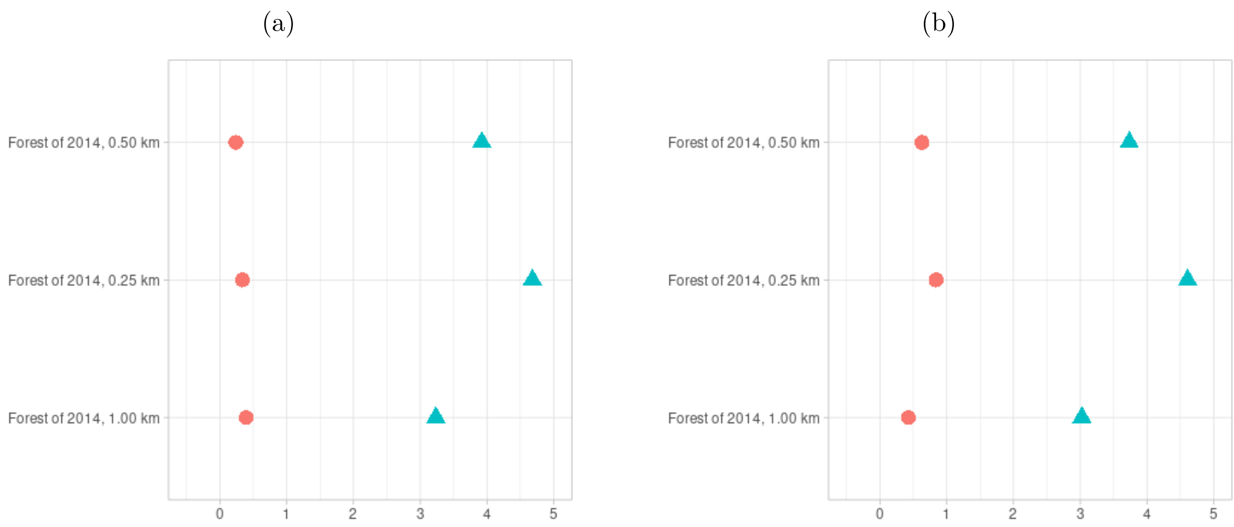
C Additional Figures

Figure 4: Distribution of Observed Temperature, °C – Cumulative.



Notes: The density plot shows the distribution of the average maximum temperature during school hours over the entire school year among all 22,778,692 schools (solid line); among the 2,780,912 schools with variation in outcome **Pass** (dashed); and among the 20,986,057 schools with variation in outcome **Pass with distinction** (dotted). The distribution is shown as density plots.

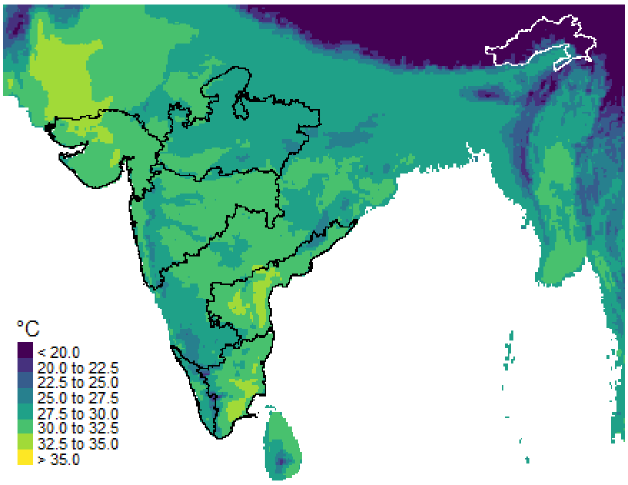
Figure 5: Standardized mean differences of the matching variables.



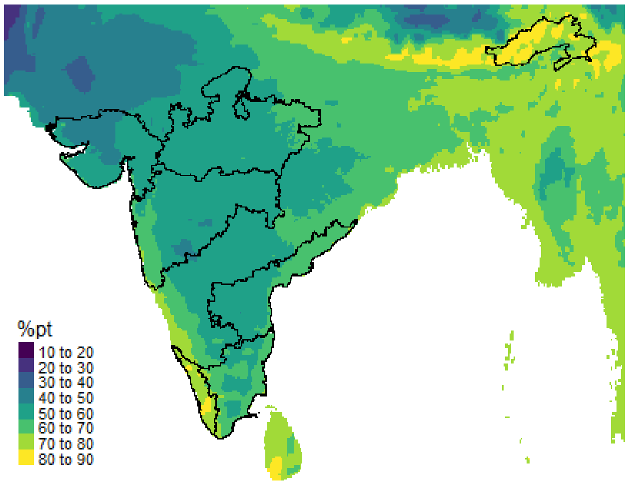
Notes: The figures show the standardized mean differences of the matching variables before (triangle) and after (circle) matching. Panel (a) shows balance in the samples used to predict the probability of passing; panel (b) shows balance in the samples used to predict the probability of earning a distinction. We exclude the categorical ("District") and binary variables ("Public", school-year dummy vv.), because post-matching standardized mean differences are zero by design.

Figure 6: Temperature, relative humidity, wind and precipitation - Cumulative.

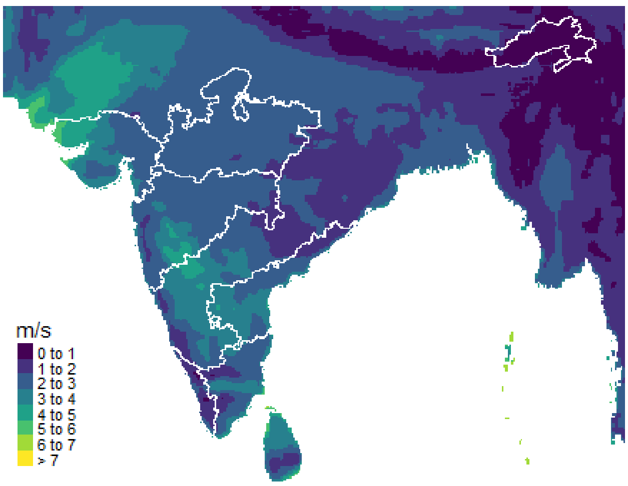
(a)



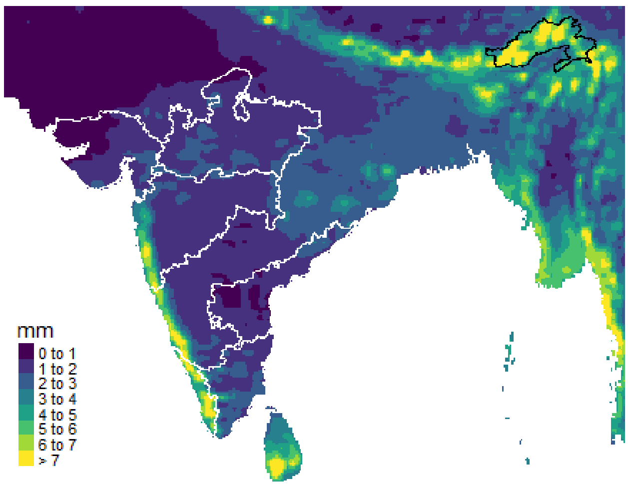
(b)



(c)



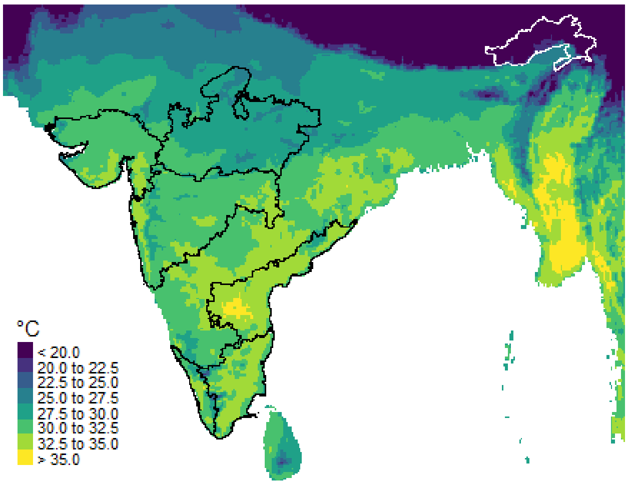
(d)



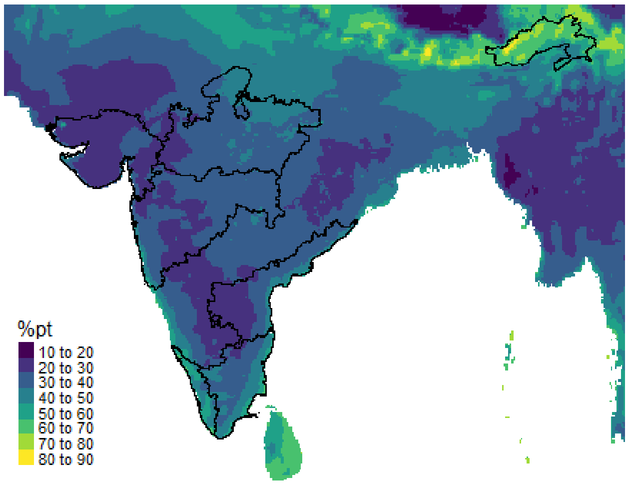
Notes: (a) temperature, (b) relative humidity, (c) wind, (d) precipitation. Aggregated over the school-year 2014-15.

Figure 7: Temperature, relative humidity, wind and precipitation - Immediate.

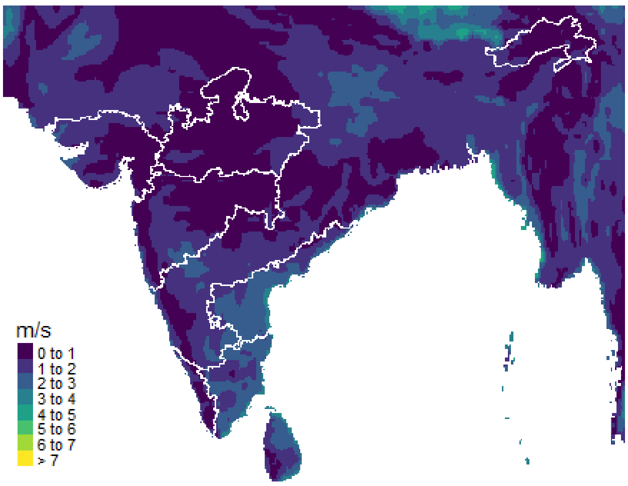
(a)



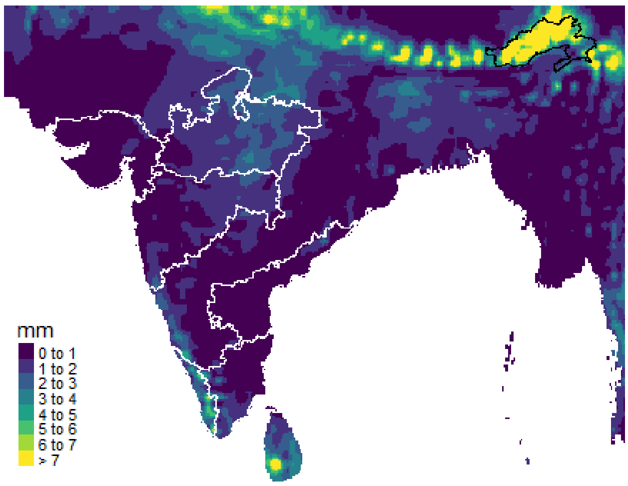
(b)



(c)

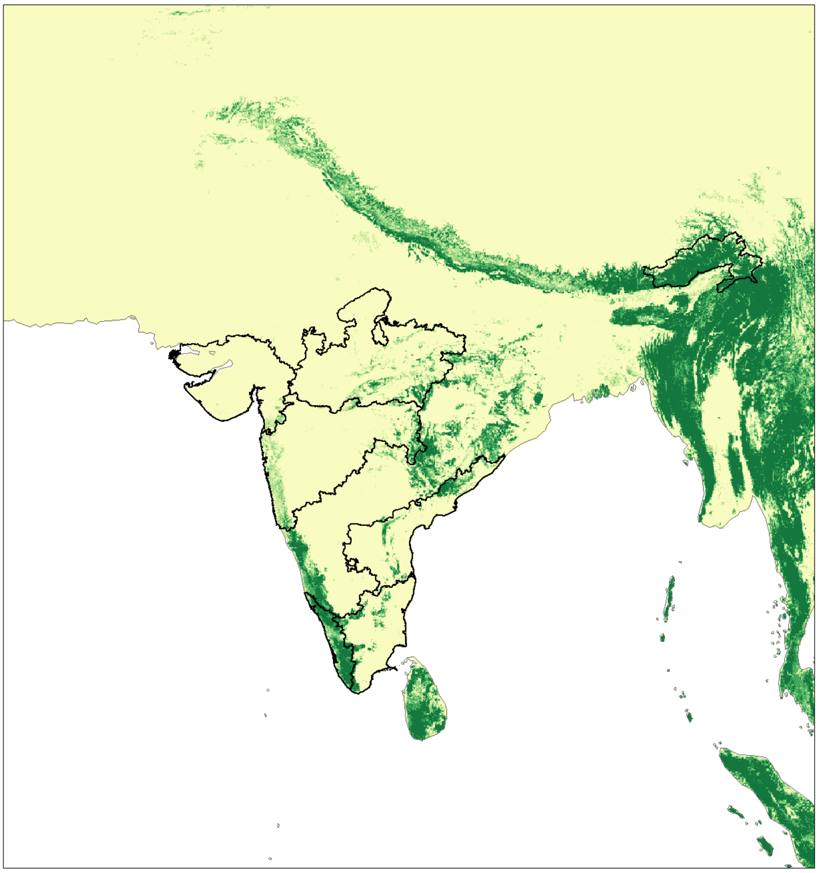


(d)



Notes: (a) temperature, (b) relative humidity, (c) wind, (d) precipitation. Aggregated over March 2015.

Figure 8: Forest Extent.



Notes: The figure depicts the geographical distribution of forests in 2014 across India.
Source: GFC