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], Sofie R. Waltl[c,d]

[a] National University of Singapore

[b]Indian Institute of Management Bangalore

[c]Vienna University of Economics and Business

[d]University of Cambridge

Motivation

- India often experiences unusually hot periods and even **heatwaves** over the summer months (March to June).
- **Frequency and duration of heat waves** in the sub-continent **are increasing** (Rohini, Rajeevan, and Srivastava, 2016).
- **Goal of our study:** measure the **association between heat and exam results** and explore **adaptation strategies**.

Literature

Well known physiological effect of extreme temperatures on learning and cognitive performance:

- Cho (2017): Summer heat negative effect on student test scores in South Korea.
- Graff Zivin et al. (2020): higher temperatures during exam period decrease test scores in China
- Park et al. (2020): extreme heat inhibits learning and affects the results of standardized tests in the U.S.
- McCormack (2023): extreme temperatures increase students' absenteeism and disciplinary infractions in the U.S.

Research Question & Identification

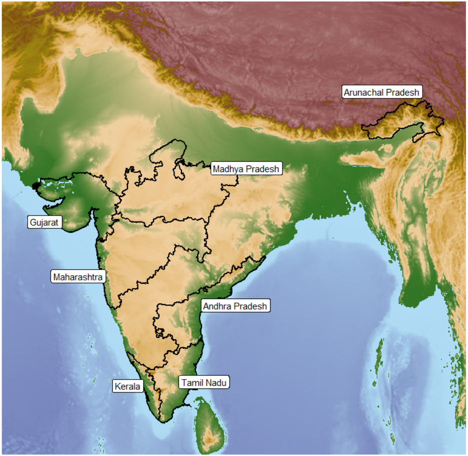
Do average temperatures over the school year affect students' learning?

- We exploit exogenous weather variations by year and location.
- Identification relies on **panel structure of the data**: the same school is observed over multiple (school-)years: we account for any time-invariant school-, or location-specific aspects by estimating a **fixed-effects model**.

School Data

- **District Information System for Education**, a database maintained by the Indian *Ministry of Education*.
- **Class VIII** exam results for schools in the 7 States where the school year runs from June to March, from 2014-15 to 2017-18.
- We observe the number of students **passing, passing with distinction and failing** the exam for each school.
- Additionally: school characteristics and location (transform to geo-references)
- Summary Statistics

School Data

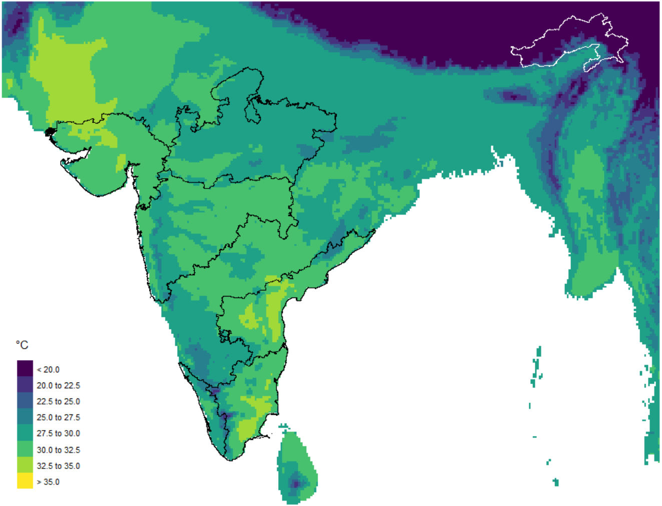


Notes: Location of the schools used for the analysis.

Meteorological Data

- Rely on the ERA5-Land **reanalysis dataset**, from the *Copernicus Climate Change* Service (Muñoz Sabater et al., 2021)
- Available **hourly** for small locational grid cells ($0.1° \times 0.1° \approx 100$ km^2)
- **Temperature:** 2 concepts
  - **average max temperature** during school days (we exclude weekends and major holidays);
  - number of days for which the maximum temperature falls into several **temperature bins**;

Average Maximum Temperature



Supplemental Data: Controls and Interaction (Mediation) Effects

- **Relative Humidity**
- **Wind**
- **Precipitation**
- **Air Pollution:** carbon monoxide, nitrogen dioxide and PM 2.5 from NASA Earth Observations database. Air quality may correlate with temperature and has been linked to students' performance (e.g. Currie et al. (2009); Ebenstein, Lavy, and Roth (2016)).

Econometric Strategy

Generic Model

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

whereby y^p_{ist} is equal to 1 if student i in school s and 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 meteorological controls.

Identifying Assumption

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

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

Main Results

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

Counterfactual Analysis

Set Up

Predict from estimated regression models hypothetical failure rates.
Counterfactual temperature scenarios:

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

for $\theta > 0$ and simulate excess numbers of students failing $\Delta^{\neg p}(\theta)$ and the excess number of students missing the distinctions $\Delta^{\neg pd}(\theta)$ via

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

Counterfactual Analysis: Temperature

Not Passing

θ	Empirical Probability $\sum \hat{y}_{ist}^{\theta}$ / 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 $\sum \hat{y}_{ist}^{\theta}$ / 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)

Counterfactual Analysis: 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)

Additional Results

- Combined effect of heat and humidity: interacted effect / wet bulb temperature / heat index. ▸ Heat & humidity
- Heterogeneity Analysis, by gender and rural-urban location. ▸ Heterogeneity

Technical robustness checks:

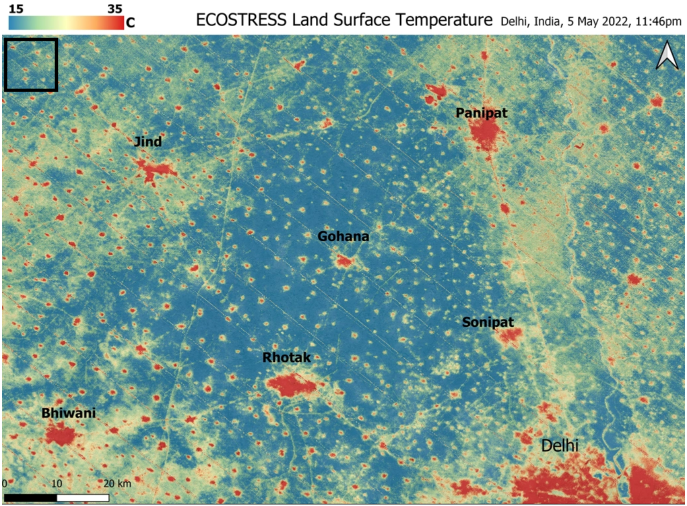
- Instead of including temperature directly, count the **number of days per temperature bracket** ▸ Robustness
- Vary model class: OLS
- Focus on the months during which the exams take place ▸ March

Adaptation

- We observe temperature at a spatial resolution of ≈ 100 km^2 and we know there can be significant unobserved differences within each cell of the grid, driven by differences in land use, vegetation height, etc.

- Artificial structures tend to absorb and reflect heat more than the natural landscape. Temperatures in urban areas can be significantly higher than in their rural surroundings (Knight et al., 2021).

Adaptation



Source: https://climate.nasa.gov/news/3176/nasas-ecostress-detects-heat-islands-in-extreme-indian-heat-wave/

Adaptation

- Differences in time-varying **vegetation** and **canopy height** may systematically impact the microclimate in the surrounding of schools.
- Vegetation may **limit solar radiation** through **tree-canopy shading** and **evapotranspiration** → **cooling effect** (Knight et al., 2021).
- Does proximity to trees and vegetation mitigate the effect of heat on schooling outcome, by increasing thermal comfort? Do 'green' surroundings reduce exposure to the 'Urban Heat Islands' effect?

Mitigation Effects: Forests

- We use **data on deforestation** from Hansen et al. (2013) to track **variations in forests' density** over time. ▸ Image

- 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).

- We measure the **extent of tree cover**, in hectares, **within a radius of 1 km (≈314 ha) and 0.5 km (≈79 ha) from each school**.

Mitigation Effects: Forests

Two concerns:

1. Within a radius of 0.5 km (1 km) from each school, only about 2.5% (5%) of students in the sample are treated.
2. Tree/forest loss is not as exogenous as weather: it may correlate with local economic development, or natural disasters.

We therefore prune our sample with **coarsened exact matching (CEM)** (Iacus, King, and Porro, 2012), based on spatial proximity and school characteristics.

▸ Standardized mean differences

Mitigation Effects: Forests

	Pass		Pass with Distinction	
	(1)	(2)	(3)	(4)
Temperature	-0.441***	-0.437***	-1.155***	-1.156***
	(0.0246)	(0.0246)	(0.0052)	(0.0052)
	[-0.0036]	[-0.0036]	[-0.1766]	[-0.1766]
Forest ha 1 km	0.0009		0.023***	
	(0.0061)		(0.0018)	
	[0.00001]		[0.0035]	
Forest ha 0.5 km		0.167***		0.121***
		(0.0239)		(0.0064)
		[0.0014]		[0.0185]
CEM	**N**	**N**	**N**	**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	2,637,871	2,640,460	18,884,151	18,884,151

APEs in square brackets.

Mitigation Effects: Forests

Based on the full sample results, on average:

- 2.6 hectares of forest (i.e. 3.5 football pitches) over an area of 79 hectares offset the impact of an increase in temperature by 1°C, on the probability of **passing** the exam.
- 9.5 hectares of forest offset the impact of an increase in temperature by 1°C, on the probability of **passing with distinction**.

Mitigation Effects: Forests

	Pass		Pass with Distinction	
	(1)	(2)	(3)	(4)
Temperature	-4.871***	-6.386***	-0.269***	-0.192**
	(0.3212)	(0.5869)	(0.0454)	(0.0645)
	[-0.2542]	[-0.3360]	[-0.0490]	[-0.0356]
Forest ha 1 km	0.210***		0.026***	
	(0.0368)		(0.0060)	
	[0.0110]		[0.0046]	
Forest ha 0.5 km		0.778***		0.073***
		(0.1188)		(0.0177)
		[0.0410]		[0.0136]
CEM	**Y**	**Y**	**Y**	**Y**
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	26,734	608,936	298,957

APEs in square brackets.

Summary & Conclusions

- **Heat negatively affect students' schooling outcomes in India.**
- **Loss in human capital** may be massive given the large numbers of affected students.
- Relying on **natural cooling** in the form of increasing **vegetation can mitigate some of the heat effects**: 2.6 hectares of forest within a radius of 0.5 km from the school offset an increase in temperature of 1°C.

References I

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

Currie, Janet, Eric A. Hanushek, E. Megan Kahn, Matthew Neidell, and Steven G. Rivkin. 2009. "Does Pollution Increase School Absences?" *The Review of Economics and Statistics* 91 (4):682–694.

Deschênes, Olivier and Michael Greenstone. 2011. "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US." *American Economic Journal: Applied Economics* 3 (4):152–85.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth. 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.

References II

Graff Zivin, Joshua, Yingquan Song, Qu Tang, and Peng Zhang. 2020. "Temperature and high-stakes cognitive performance: Evidence from the national college entrance examination in China." *Journal of Environmental Economics and Management* 104 (102365).

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." *Science* 342 (6160):850–853.

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2012. "Causal Inference without Balance Checking: Coarsened Exact Matching." *Political Analysis* 20 (1):1–24.

References III

Knight, Teri, Sian Price, Diana Bowler, Sian King, Ko Konno, and Raja Lorena Richter. 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.

McCormack, Kristen. 2023. "Education Under Extremes: Temperature, Student Absenteeism, and Disciplinary Infractions." *Mimeo* .

Muñoz Sabater, J., E. Dutra, A. Agustí-Panareda, C. Albergel, G. Arduini, G. Balsamo, S. Boussetta, M. Choulga, S. Harrigan, H. Hersbach, B. Martens, D. G. Miralles, M. Piles, N. J. Rodríguez-Fernández, E. Zsoter, C. Buontempo, and J.-N. Thépaut. 2021. "ERA5-Land: a state-of-the-art global reanalysis dataset for land applications." *Earth System Science Data* 13 (9):4349–4383.

Park, R Jisung, Joshua Goodman, Michael Hurwitz, and Jonathan Smith. 2020. "Heat and learning." *American Economic Journal: Economic Policy* 12 (2):306–39.

References IV

Rohini, P., M. Rajeevan, and A. K. Srivastava. 2016. "On the Variability and Increasing Trends of Heat Waves over India." *Scientific Reports* 6, 26153.

Appendix: Summary statistics

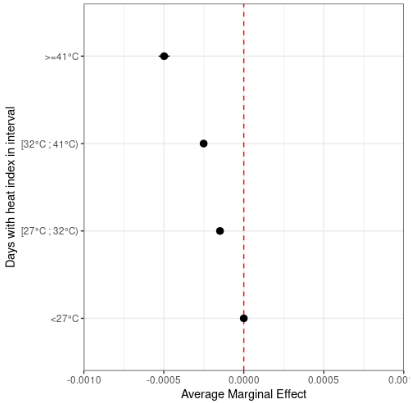
	All schools		Only schools with variation in outcome			
			Pass		Pass with Distinction	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Pass	0.988	0.110	0.909	0.288	0.988	0.111
Pass with distinction	0.710	0.454	0.610	0.488	0.691	0.462
Boy	0.518	0.500	0.520	0.500	0.517	0.500
Temperature	30.657	1.513	30.480	1.534	30.650	1.519
Wet Bulb Temperature	21.860	1.611	21.610	1.575	21.856	1.611
Heat Index	31.996	1.958	31.710	1.954	31.992	1.953
Relative humidity	57.964	7.443	57.481	6.937	57.969	7.476
Wind speed	2.634	0.649	2.593	0.603	2.627	0.650
Precipitation	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 ha 1km	32.465	68.197	30.450	65.243	32.868	68.858
Forest ha 500m	8.057	17.463	7.502	16.659	8.156	17.625
Forest ha 250m	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	

Back

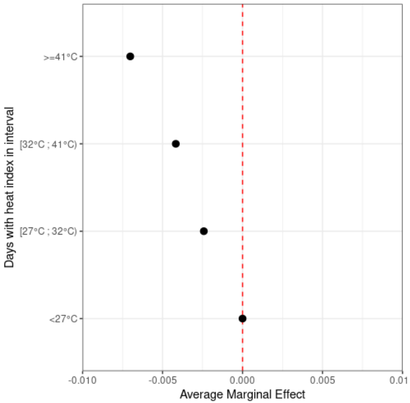
Heat & Humidity

	Pass			Pass with Distinction		
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature	-0.472***			-1.444***		
	(0.0503)			(0.0101)		
Humidity	-0.088***			-0.309***		
	(0.0243)			(0.0050)		
Temperature * Humidity	0.001			0.004***		
	(0.0007)			(0.0002)		
Wet Bulb Temperature		-0.494***			-0.455***	
		(0.0219)			(0.0047)	
Heat Index			-0.258***			-0.332***
			(0.0139)			(0.0029)
Wind	Y	Y	Y	Y	Y	Y
Precipitation	Y	Y	Y	Y	Y	Y
Air pollution	Y	Y	Y	Y	Y	Y
School FE	Y	Y	Y	Y	Y	Y
School year dummy vv.	Y	Y	Y	Y	Y	Y
Observations	2,780,912	2,554,345	2,525,282	19,915,599	18,493,322	18,251,535
BIC	1,384,972	1,265,650	1,249,897	21,864,784	20,390,798	20,100,041

Heat & Humidity



a) Pass



b) Pass with distinction

Back

Heterogeneity Analysis

- Previous evidence shows that men are more likely to suffer from severe heat than women (Deschênes and Greenstone, 2011), so we investigate differences between **boys and girls**.
- Given that our data does not capture local variations in temperature, we believe that **rural and urban schools** might experience systematic differences due to the 'Urban Heat Islands' (UHI) effect.

Back

Heterogeneity Analysis

	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

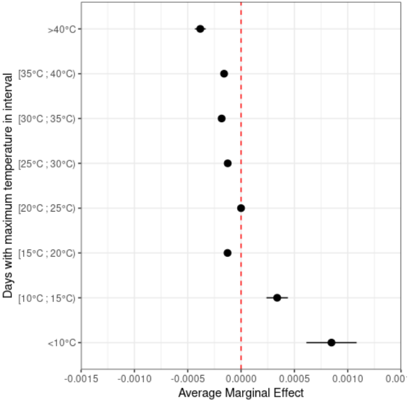
Heterogeneity Analysis

	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

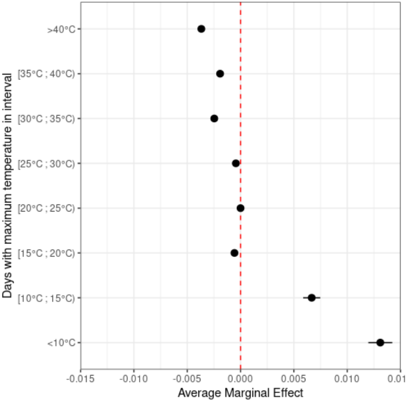
Robustness

	Pass		Pass with Distinction	
	(1)	(2)	(3)	(4)
	OLS	*Logit*	*OLS*	*Logit*
Temperature	-0.032*		-0.190***	
	(0.0136)		(0.0055)	
Days below 10°C		0.105***		0.085***
		(0.0119)		(0.0028)
Days 10-15°C		0.042***		0.043***
		(0.0052)		(0.0019)
Days 15-20°C		-0.016***		-0.004***
		(0.0013)		(0.0002)
Days 25-30°C		-0.016***		-0.003***
		0.0008		(0.0002)
Days 30-35°C		-0.023***		-0.016***
		(0.0010)		(0.0002)
Days 35-40°C		-0.020***		-0.012***
		(0.0014)		(0.0003)
Days above 40°C		-0.047***		-0.024***
		(0.0024)		(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

Robustness



a) Pass



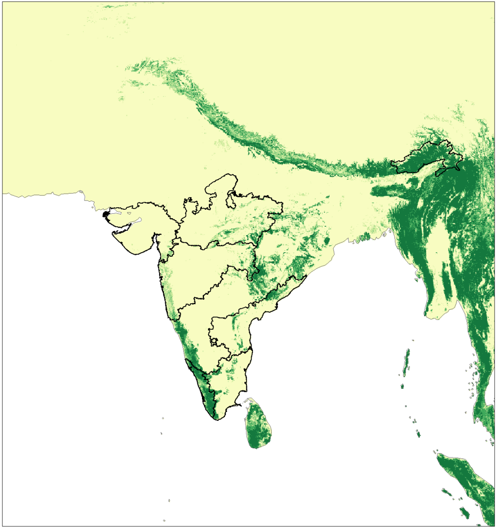
b) Pass with distinction

March Temperatures

	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

Back

Mitigation Effects: Forests



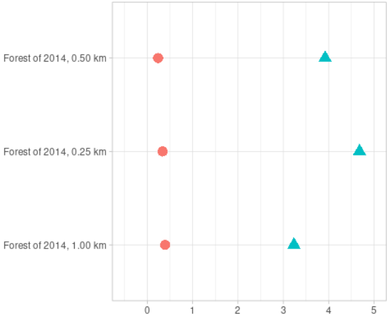
back

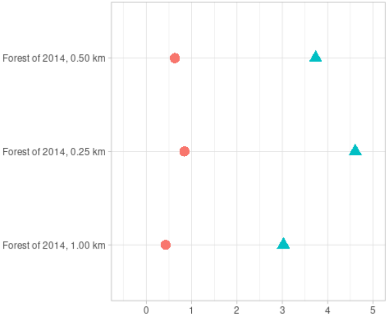
Standardized mean differences, Matching

Matching

- **Treated group:** schools that experience forest loss in any of the school-years in our panel, within a 1 km (0.5 km) radius.
- **Control group:**
  - Schools that never experience any tree cover loss within a radius of 2 km, in any year t or $t-1$.
  - Matched on:
    - district;
    - amount of forest cover as of 2014;
    - public or private school;
    - years in which the school is observed.
  - One-to-one match between students.

Standardized mean differences, Matching





Notes: Standardized mean differences of the matching variables before (triangle) and after (circle) matching. The plot on the left shows balance in the samples used to predict the probability of passing, the plot on the right shows balance in the samples used to predict the probability of earning a distinction.

Back