The Causal Effect of Adverse Temperature Shocks on Schooling Outcomes in India

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

[a]National University of Singapore
[b]Indian Institute of Management Bangalore
[c]Vienna University of Economics and Business
[d]University of Cambridge

January 25, 2024

Introduction

As climate change makes the planet progressively warmer, extreme weather events will become more severe and frequent. Temperature anomalies between March and July have been increasingly common in India during the first two decades of the century.

The existing empirical evidence suggests that severe weather shocks, such as droughts and heat-waves, may affect schooling outcomes and the quality of education. According to recent studies, cumulative heat exposure combined with the lack of air conditioning directly affects education, as it increases absenteeism and it inhibits learning both at home and at school (Park et al., 2020; McCormack, 2023). Moreover, exposure to higher temperatures during examinations has a negative effect on students' performance (Graff Zivin et al., 2020).

In this paper, we measure the causal effect of exposure to extreme temperatures on schooling success across India. We link primary school records across India with information on local weather conditions with a special focus on extreme events. We argue that heat-waves directly affect the productivity of students, and therefore that temperature shocks have an adverse effect on schooling. Given that India relies on centralized exams that do not take such special circumstances into account, local extreme temperatures may hamper kids' future success in school and, further on, employability.

We plan to corroborate our findings by focusing on standardized secondary school exams from the Indian state of Assam.

Finally, we use canopy-height and deforestation data to provide preliminary evidence that vegetation in the proximity of the schools has a mitigating effect that increases with canopy height and tree density. These results suggest that enhancing tree-canopy shading is a possible adaptation strategy that can be taken to mitigate the impact of heat on students' performance.

Data

We make use of data on schools and examinations collected within the *District Information System for Education (DISE)*, a government database created by the Indian Ministry of Education. Our main outcome variable is the result of the *Class VIII* exam, which marks the conclusion of mandatory education. We observe a repeated cross-section of approximately 48 million students who enrolled in the exam between 2015 and 2018.

We complement these data with the *High School Leaving Certificate* (HSLC) examination records provided by the Assam Board of Secondary Education. We have access to the results from about 400,000 students who took the exam in 2023, while results for the 2024 cohort are expected to be available by June 2024. As the HSLC exam is standardized at the regional level, we can reasonably assume that the difficulty of the exam is exogenous to local variations in temperature, therefore strengthening our claim of a causal link between temperature shocks and school results.

As we know schools' coordinates, we can link them to a wide set of meteorologic data. We obtain these data from the *Copernicus Climate Change Service*. We use the ERA5-Land data set, which contains hourly values for various climate related variables, from 1950 to the present. The data has a resolution of $0.1°\times0.1°$(approximately 10 km^2). We also 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.

We measure canopy height using the cross-sectional data provided by Lang et al. (2023), which shows the height of vegetation over the entire globe in 2020, with a resolution of 10×10 meters. Finally, we use data on deforestation (Hansen et al., 2013) to track variations in forests' density over time. This dataset covers yearly gross forest cover loss between 2001 and 2022, with a spatial resolution of about 30×30 meters.

Methods

The outcome are students' examination results, which are reported in the broad categories of 'fail', 'pass' and 'pass with distinction'. We therefore estimate two separate logit models that distinguish fail versus pass and fail/pass versus pass with distinction.

To identify the effect of adverse temperature shocks on test results, we regress exam results on the temperature measured at the school's location both before and during the exam period. To ensure clean measurement, we account for any time-invariant school- and location-specific aspects by estimating a fixed-effects model. Thus, we rely on temperature variation within school-location over four academic-years (2014-15 to 2017-18).

We use the algorithm suggested by Stammann, Heiß, and McFadden (2016) to estimate fixed effects logit models with large panel data, at low computational cost.

Furthermore, we plan to use a quasi-experimental setting to causally assess the impact of exceptional heat exposure on Assamese students who attended the HSLC exam between 2023 and 2024.

As temperature alone is insufficient to measure all aspects relating to thermal (mis-)comfort, we measure relative humidity, wind speed and precipitation for each school location and add these variables as supplemental meteorological indicators. We also control for local air quality, as previous studies have shown that increased air pollution may lead to worse students' performance and absenteeism (Currie et al., 2009; Ebenstein, Lavy, and Roth, 2016).

Results

In our main results we find that with higher measured temperature the odds of students passing the exam as well as the odds of students passing the exam with distinction decrease significantly. To better understand the severity of the impact of temperature shocks, we express this effect in the relative number of students affected. We compare the predicted number of students that fail or miss the distinction for the observed temperature as well as an hypothetically increased temperature. A constant increase in temperature by merely 0.5°C (0.9°F) means a drop in the number of students passing the exam by 1.0% and a drop in distinctions of 3.8%.

Moreover, we use a specification that replaces the single average temperature measure with the number of school days falling into various temperature brackets. This approach reveals that adverse effects are mostly driven by very high temperatures: the effect on the probability of passing the exam is increasingly negative for higher temperature brackets, and the effect is largest for days with maximum temperature above 40°C (104°F).

Differences in vegetation may additionally systematically impact the local climate, as vegetation limits solar radiation through tree-canopy shading and evapotranspiration has a cooling effect. Local variations in temperature are not captured by the ERA5-Land data, which comes at a spatial resolution of about 10 km^2. We therefore expect that proximity to trees mitigates the effect of heat by increasing thermal comfort. To test this hypothesis, we incorporate measures of forest loss and forest height in the proximity of school into our main models. The (preliminary) results show that students attending schools that are located in areas where trees are scarce, or that suffered losses in forest cover, are less likely to pass the exam or pass it with distinction, for a given temperature change.

Conclusion

We provide evidence that exposure to higher temperature before and during the exam period has an adverse effect on the performance of students in India. These effects are both statistically and economically significant. We contribute to the current literature by focusing on primary education in a low-middle income country. We also provide the first evidence that vegetation in the proximity of the schools may have a mitigating effect, thus suggesting a possible adaptation strategy.

References

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.

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.

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.

Lang, Nico, Walter Jetz, Konrad Schindler, and Jan Dirk Wegner. 2023. "A high-resolution canopy height model of the Earth." *Nature Ecology & Evolution* 7:1778–1789.

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

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

Stammann, Amrei, Florian Heiß, and Daniel McFadden. 2016. "Estimating Fixed Effects Logit Models with Large Panel Data." G01-V3. ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft.

van Donkelaar, Aaron, Melanie S. Hammer, Liam Bindle, Michael Brauer, Jeffery R. Brook, Michael J. N. Garay, Christina Hsu, Olga V. Kalashnikova, Ralph A. Kahn, Colin Lee, Robert C. Levy, Alexei Lyapustin, Andrew M. Sayer, and Randall V. Martin. 2021. "Monthly Global Estimates of Fine Particulate Matter and Their Uncertainty." *Environmental Science & Technology* .