Heat and Learning in Primary and Secondary Education

Travis Roach* and Jacob Whitney†

February 12, 2019

Abstract

Changing weather patterns and extreme events are not the only outcomes of global climatic change. We investigate the impact of changing weather conditions on human capital development by studying achievement on standardized tests in Math and English/Language Arts for students in grades 3-8. Here we show, that increasing average temperature levels and particularly hot days reduce student learning and achievement. We find that each additional day above 100F degrees decreases student achievement on these tests, and that a 6.5F degree difference in average temperatures has the same impact as living in an area with a lower median income of about $1,000. This study confirms many findings in the received literature on global climate change and human capital acquisition and productivity.

JEL Classification Codes: H23, I21, J24, Q51, R11

Keywords: Climate Change, Student Achievement, Human Capital

*Corresponding Author. Department of Economics, University of Central Oklahoma, troach2@uco.edu.

†Oklahoma Council on Economic Education, jwhitney2@uco.edu

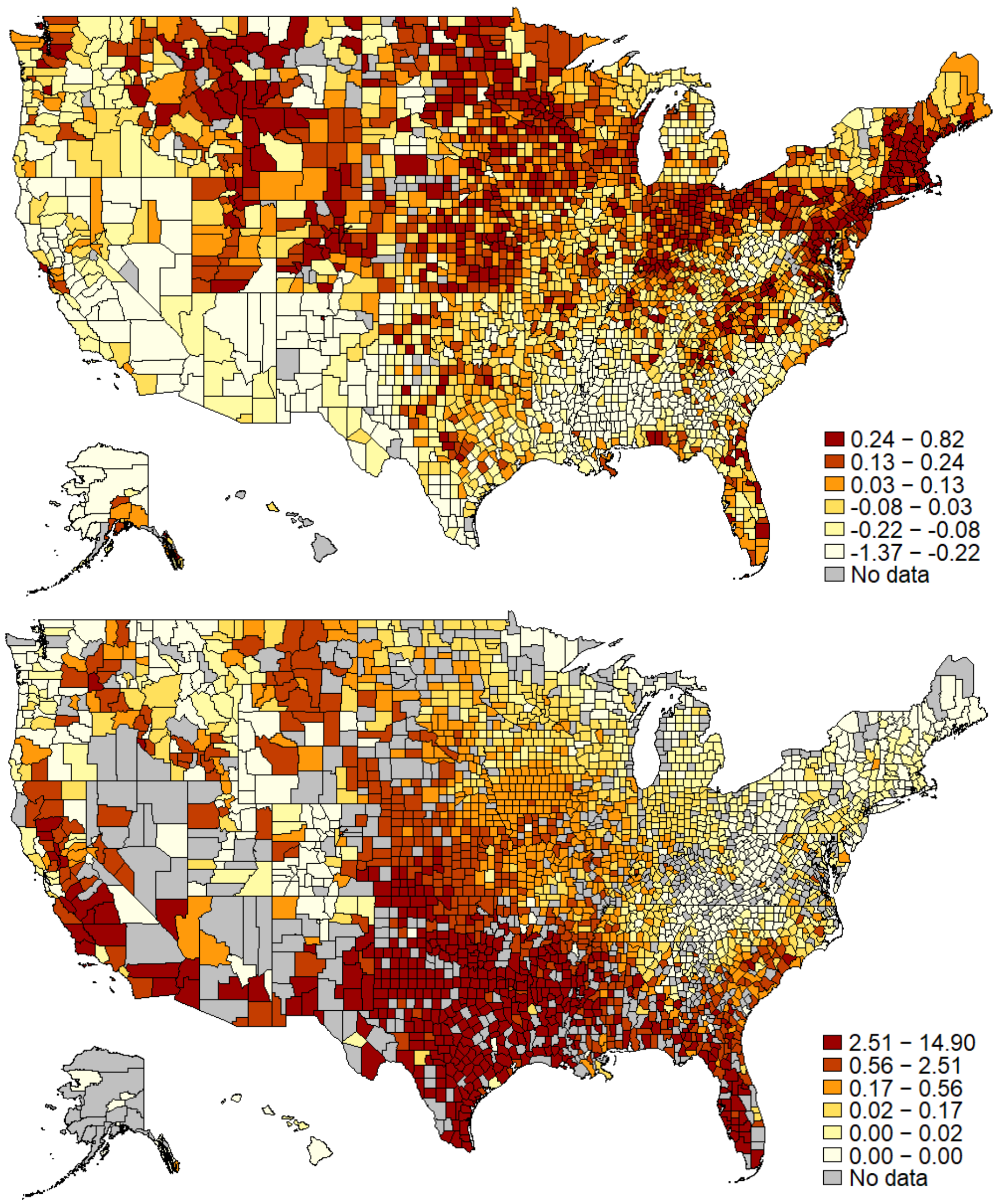
1 Introduction and Background

Despite more than a century of scientific evidence discussing the effect of accumulating greenhouse gasses on temperature (Foote, 1856), policies intended to limit these emissions have been slow to pass or are, at least in some countries, a political non-starter. Regardless of the political appetite to limit greenhouse gas accumulation, the effects of global climate change have already taken shape. These impacts can be seen in the heightened severity of droughts, hurricanes, and (of course) extreme temperatures. Beyond changing meteorological conditions, the effects of global climate change stand to impact human interactions and development in fundamental ways. These effects are sometimes extreme, for instance the lost economic value of one's home being destroyed in a fire or hurricane, and these effects are sometimes a little more banal, for instance difficulty paying attention or motivating one's self on a hot day. Regardless of the emotions drawn when discussing these differing impacts, they matter economically. And much like the accumulation of emissions from individual decisions have global consequences, the accumulation of climate-driven impacts stand to have long-lasting and non-trivial economic impacts.

There is a small but growing literature on the impacts of climatic change and "human capital" that has established a causal link between hot days or climbing temperatures and negative learning or productivity outcomes. We contribute to this literature by studying achievement on standardized tests for young students from 3rd through 8th grade for the entire United States. Our paper is the first to study this young of an age group without being limited geographically. Using nationally-comparable data on achievement in math and english/language arts at the individual school and grade level, we determine how changes in average temperatures and weather extrema during school days impact learning.

We find that rising average temperatures harm student learning. This result is robust to multiple specifications that vary in demographic controls, the subject tested, as well as the inclusion or exclusion of key parameters. We further find that each additional hot day harms student learning.

Figure 1: Average Test Scores (top) and Days over 90F (bottom) in 2013



1.1 Climate Change and Human Capital

Brain imaging and psychometric testing have shown that heat and thermal stress cause heightened brain stress to fulfill cognitive processes like attention, memory, verbal learning, information processing and concentration (Hocking et al. (2001)). Without technology to directly detect how thermal stress inhibits brain function, economists have studied how thermal stress manifests in our daily lives by linking temperature anomalies to changes in performance and human capital development. Zivin et al. (2015) study the impact of short-run weather and long-run achievement using data from the National Longitudinal Survey of Youth (NLSY) and meteorological conditions on the day of the assessment. They find that math scores decline at temperatures above 21C. Park (n.d.) also finds a negative impact of hot days on learning (and achievement) using data on students in NYC public schools that take exams that are needed to graduate in June. Specifically, Park (n.d.) finds that hot temperatures during an exam result in reduced performance. They show that taking an exam on a 90-degree day reduces performance by 14 percent of a standard deviation, which also impacts a students chance to graduate by 10.9%. Garg et al. (n.d.) show that 10 or more days in a year with an average daily temperature above 29C lowered both math and reading scores. In a study closely related to the methodology used here, Goodman et al. (2018) provides causal evidence of both the impact of heat on student learning, and the mitigating effects of air conditioning using a nationally representative sample of high school PSAT takers. They find that 1-degree higher temperature in the previous school year reduces learning by 1% of a years worth of learning. They deduce that heat reduces academic achievement by reducing the productivity of instructional time. To show this, the authors use a unique data set on the quality of air conditioning in schools and show that air conditioning appears to offset nearly all of the damages of heat exposure.

Rising average temperatures have also been linked to lower productivity and income growth. Heal and Park (2013) show that years with warmer than average temperatures are associated with lower output per-capita in hot climates, and higher output per-capita for countries in cold ones. The authors find an effect size of approximately 3%-4% per degree. Dell et al. (2009) show that a negative relationship exists between temperature and income, even within regions and states of a country, and that a 1-degree Celsius rise in temperature is associated with a 1.2-1.9% decline in municipal per-capita income. Deryugina and Hsiang (2014) go beyond linking hot temperatures and income and find temperature ranges that maximize income. They find that total personal income per capita is highest when the temperature is between 9-15C (48.2-59F). They further find that a day with an average daily temperature of 29C decreases productivity by .065% relative to a day with an average daily temperature of 15C, which means that a 29C day is 23.6% less productive than an average day. The same relative change lowers annual earnings by 0.11%.

2 Modeling Strategy

Prior authors have established a causal link between heat, heat exposure, and human capital and productivity. We aim to examine this link for young students in the United States, which would be a new population (at the national level) to study this effect of global climatic change. We are able to identify the causal impact of heat on learning because there is exogenous variation in weather events across space and time, and this variation even impacts cohorts of students within the same school at different periods in their academic careers. Using panel data methods we are able to account for fixed factors that impact student learning over time that are related to place (access to parks, air conditioning availability, proximity to quality grocery stores, or other neighborhood-level effects) and idiosyncratic differences in state governance that are uncorrelated with warm or cool temperatures (e.g. state funding formulas that dictate funding for education). The two data sets we draw from are described separately below. Figure 1 visually shows the connection we investigate, and summary statistics for all model parameters are included in Table 1.

2.1 Weather Data

We use information on daily temperature averages from the National Oceanic and Atmospheric Administration's (NOAA) global monitoring system. This data is collected by 58,578 individual stations in the United States which yield over 149 million daily observations for the time sample studied here.[1] To construct the geographic school district-level estimates, we take the median temperature reading from all stations within a 20 km radial distance from each county centroid,

[1]This data includes information on the daily temperature average, minimum, and maximum, and any precipitation.

and link each school district to the county it is located within. We also restrict temperature anomalies to those that would impact in-class learning. To do this we only use observations on Monday-Friday during school months as in Goodman et al. (2018). Finally, for each academic school year (and each school) we build variables that will capture the effect of heat on student learning at both the extensive and intensive margins. First, we calculate the average temperature observed which captures changes along the extensive margin, and then we calculate the amount of days that the temperature fell within a temperature range, above 100F, above 90F (but below 100F), above 80F (but below 90F), and so on. These variables capture changes along the intensive margin which is similar to the method used in Goodman et al. (2018).

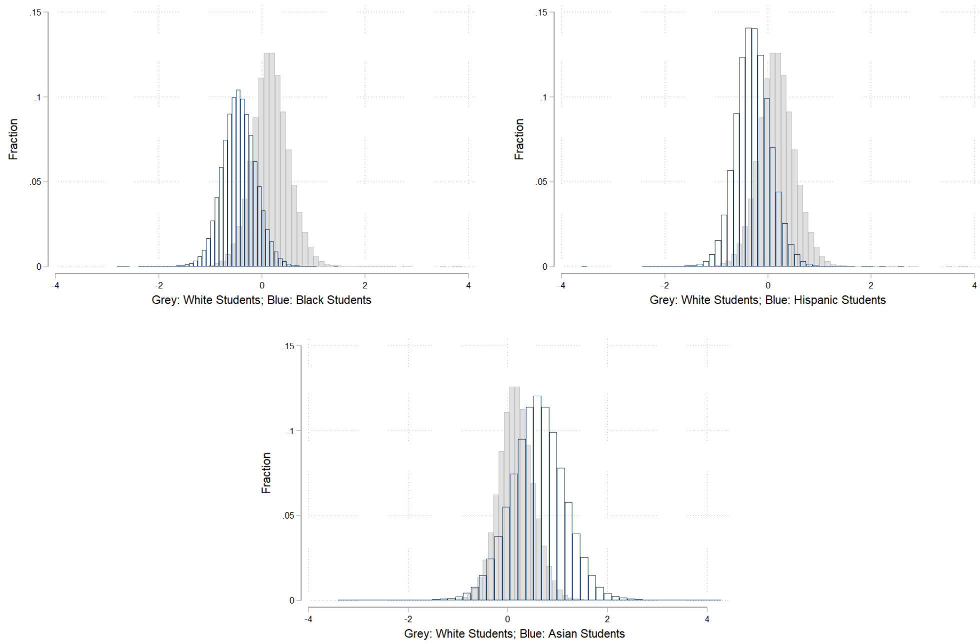
2.2 Education Data

To gauge student learning and achievement we use data from the Stanford Educational Data Archive (SEDA) that has been constructed to compare proficiency outcomes nationally. These data include assessment outcomes from the 2008-09 school year to the 2014-15 school year for students in 3rd grade through 8th grade for both English/Language Arts and Math tests.

Under federal legislation, each state is required to test every student in grades 3-8 and in one high school grade in Math and English/Language Arts (ELA) each year.[2] Each state is allowed to administer the test of their choice that measures student achievement relative to the state's standards for each subject. States also determine standards regarding the level of performance considered "proficient" for each grade and subject. The SEDA (2.1) achievement data is constructed using state-recorded counts of students deemed proficient which are reported to the U.S. Department of Education. To nationally standardize achievement and proficiency, the SEDA data is constructed such that the estimated means and standard deviations at the school-level summarize the achievement represented by the observed counts in the raw (student-level) data. To do this, ordered probit models are fit at a 'geographic school district' or GSD-level.[3] The process is documented in Ho and Reardon (2012), Reardon, Shear, Castellano and Ho (2017), and Reardon

[2]The SEDA data excludes the high school year that is tested.

[3]Most traditional public schools have the same geographic and administrative district. There are a number of other types of schools (e.g. charter schools, virtual schools) that belong to an administrative district but do not have a geographic boundary. The SEDA data assigns each such school to a GSD.



et al. (2018). The scale used for this paper is the 'cohort scale' which is standardized by dividing by the national grade-subject-specific standard deviation for a given cohort in the data (for example, students in 4th grade in 2009 and 8th grade in 2013). Ho and Reardon (2012) note that this metric can be interpreted as an effect size which can be used to describe change over time in test scores.

A perennial issue that we must account for in our model is the 'achievement gap' among white and minority students. White students have historically performed better on standardized tests relative to black students (Vanneman et al. (2009)) and Hispanic students (Hemphill and Vanneman (2011)), and this gap is persistent over time. While there is a large literature devoted to discussing and determining the causes of these gaps, one possible explanation that is related to our study is that of instructional time. Lavy (2015) shows that each additional hour of instructional time yields a 0.15 standard deviation improvement in student achievement. Given the received literature on the mitigating effects of air conditioning (Goodman et al. (2018), Garg et al. (n.d.), Laurent et al. (2018)), part of the achievement gap may be due to climate adaptation and school funding.

Much like the findings of prior authors, we recognize an achievement gap in the SEDA data, which is discussed in greater detail in Reardon, Kalogrides and Shores (2017). Figure 2 shows the histogram of student achievement outcomes black, Hispanic and Asian students (blue bars)

Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Average Proficiency	825,814	0.04522	0.400055	-3.18686	3.80456
Median Income	799,894	62.02792	26.43517	12.4995	228.5294
Average Temp.	701,624	65.49427	8.901525	24.34941	101.1936
Days over 100	701,624	1.669123	5.21492	0	68
Days over 90	701,624	9.019023	11.14708	0	94
Days over 80	701,624	30.13792	15.00081	0	189
Days over 70	701,624	32.84046	11.36682	0	176
Days Under 60	701,624	89.89703	42.82277	0	216
% Black	825,461	0.085485	0.173608	0	1
% White	825,461	0.740867	0.275585	0	1
% Hispanic	825,461	0.127164	0.20034	0	1
% Asian	825,461	0.021662	0.049384	0	0.8

Table 1: Cross-sectional unit of observation is a grade-level in a school (3rd-8th), over time from 2009-2015.

compared to white students (grey bars).[4] We account for this achievement gap by including the percentage representation of black, white, Hispanic, and Asian students for each grade in each school as a regressor. Our dependent variable is the average score for each grade in each school in a year, so we do this to account account for the effect that race and ethnicity may play in average scores that are independent of the observed weather.

2.3 Empirical Model

The SEDA data includes information on average achievement within a school for each grade (3rd-8th) from 2009 to 2015. Thus, there are repeat observations for (mostly) the same group of students as they progress over time and are witness to differing weather and temperatures while still being impacted by school-district (neighborhood) fixed effects. Our identification is achieved by controlling for grade, state, school district, and year fixed effects (and their interactions) while exogenous temperature anomalies vary over space and time within these geographic boundaries. Our main estimating equation is presented below, where the dependent variable, $Score_{igsy}$, is the average score in school i, for grade g, on subject s, in year y. In later models this dependent variable is the average score of students of certain race backgrounds (in a school, year, grade), or

[4] The SEDA data is normalized such that a score of 0 indicates performance on par with the rest of the cohort. Lower scores indicate lower proficiency and achievement, and vice versa.

on certain tests (Math or ELA).

$$
\begin{aligned}
Score_{igsy} &= \beta_0 + \beta_1 AvgTemp_{iy} + \beta_2 Over100_{iy} + \beta_3 Over90_{iy} + \beta_4 Over80_{iy} + \beta_5 Over70_{iy} + \beta_6 Under60_{iy} \\
&\quad + \beta_7 MedianIncome_{iy} + \sum_{i=0}^{n} \pi_{0+i} RacePercent_{igy} + \mu_i + \gamma_g + \xi_{gy} + \varepsilon_{igsy}
\end{aligned}
$$

The primary variables of interest are $AvgTemp_{iy}$ which measures the average temperature observed for each school over the year, and the count variables $Over100_{iy}$ (and others) which measure the impact that days above a threshold temperature have on performance. While we do not explore the impetus for lower or higher average grades by ethnicity, we account for the percentage of white, black, Hispanic, and Asian students in each grade so that our results are not biased due to spatial sorting of these students that may coincide with hotter or cooler areas. We also account for the median income in a school area, $MedianIncome_{iy}$. The variables, μ_i, γ_g, and ξ_{gy} denote various fixed effects that we toggle for robustness. We are not able to control for specific adaptation measures like air conditioning as in Goodman et al. (2018) or Laurent et al. (2018). Thus, we include school district fixed effects in each of the model specifications presented below to proxy for these adaptation measures. All standard errors are clustered by geographic school district (GSD).

3 Results

Our primary results are presented in table 2. We begin by exploring the simple effect of temperature on scores while only controlling for school district fixed effects (specification 1). Here we see that each 1f degree change in average temperatures decreases student achievement. When we control for the median income in the school district area (specification 2) the effect of a 1F degree change becomes more muted, however it remains negative and statistically significant. Once we control for grade, year, and state fixed effects (specification 3), the effect of temperature changes on student achievement again lessens compared to when there are no other controls in the model (spec. 1), however we arrive at a consistent estimate of the impact that a 1F degree change in average school day temperatures has on student achievement (which is still negative and statistically significant at the 1% level). Specifications 4 and 5 both include controls for the racial makeup of the school. In these models we see that the impact of an additional 1F degree in average school day temperatures is negative and in the same range as when race controls are excluded. We note that the estimated effects for each race conforms with prior authors' research on achievement gaps, as well as the raw data plotted in caparison with one another (Figure 2). We note, too, that the marginal effect of an increase in median income is lower when race controls are included. We expect that when race is not included as a control variable the estimated effect of a change in median income is biased upwards because white and Asian students score better than black or Hispanic students on average, and that failing to include these measures picks up an endogenous relationship between the median income of an area and a school's racial composition. The effect of an additional day above 100F remains negative and statistically significant across model specifications.

The scaling of coefficients is difficult to interpret as they measure the change in a variable that has been standardized. We thus compare the estimated effect of an additional 1F in average temperatures with the effect of median incomes.[5] We expect a 6.53F increase in average temperatures to have the same negative effect as living in an area with a median income that is $1,000 lower. This is a considerable impact given the recent report of the Intergovernmental Panel on Climate Change (IPCC (2018)). The IPCC (2018) report states that a 1.5C average temperature change can be expected (about 2.7F). This large of a change is similar to the estimated effect of living in an area with median incomes that are lower by approximately $413. Obviously, areas that will warm by more than 1.5C can expect to witness even lower achievement due to hotter days.

Tables 3 uses specification 4 and 5 from above, but uses test scores disaggregated by test subject. Our results are strikingly similar to the main model which pools these scores. This implies that both math and English/language arts (ELA) proficiency are hurt by increasing average temperatures and days over 100F. Specifications 6 and 7 use math scores as the dependent variable, and specifications 8 and 9 use ELA scores as a dependent variable.

[5] Median income is measured in $1,000s of dollars.

Table 2: Primary Estimates

	(1)	(2)	(3)	(4)	(5)
Average Temp.	-0.00246***	-0.00243***	-0.00134***	-0.00169***	-0.00122***
	(0.00043)	(0.00033)	(0.00026)	(0.00029)	(0.00025)
Over 100	-0.00166***	-0.00135***	-0.00127***	-0.00140***	-0.00123***
	(0.00022)	(0.00019)	(0.00017)	(0.00018)	(0.00017)
Over 90	-0.00124***	-0.00057***	-0.00029***	-0.00017	-0.00012
	(0.00019)	(0.00015)	(0.00011)	(0.00013)	(0.00011)
Over 80	-0.00045***	-0.00018**	-0.00005	0.00001	-0.00002
	(0.00011)	(0.00008)	(0.00006)	(0.00008)	(0.00006)
Over 70	-0.00067***	-0.00057***	-0.00044***	-0.00042***	-0.00044***
	(0.00010)	(0.00008)	(0.00007)	(0.00007)	(0.00007)
Under 60	-0.00009	-0.00018***	-0.00010**	-0.00016***	-0.00010**
	(0.00006)	(0.00005)	(0.00004)	(0.00004)	(0.00004)
Median Income		0.00929***	0.00969***	0.00750***	0.00797***
		(0.00064)	(0.00061)	(0.00055)	(0.00057)
% Black				-0.09863	-0.29687***
				(0.07568)	(0.06684)
% White				0.46007***	0.25187***
				(0.06095)	(0.05429)
% Hispanic				0.37452***	0.12758**
				(0.06829)	(0.05883)
% Asian				0.94157***	0.62160***
				(0.11437)	(0.09312)
School District FE	Y	Y	Y	Y	Y
Grade FE	N	N	Y	N	Y
Year FE	N	N	Y	N	Y
State FE	N	N	Y	N	Y
R2	0.692	0.704	0.714	0.708	0.716
N	695320	673495	673495	673495	673495

Table 2: Robust standard errors clustered at the school-district level. ***, ** indicates statistical significance at the 1% and 5% level, respectively.

Table 3: Math/ELA Proficiency

	(6)	(7)	(8)	(9)
Average Temp.	-0.00162***	-0.00098***	-0.00160***	-0.00131***
	(0.00035)	(0.00031)	(0.00029)	(0.00026)
Over 100	-0.00085***	-0.00060***	-0.00227***	-0.00218***
	(0.00022)	(0.00021)	(0.00017)	(0.00016)
Over 90	-0.00002	0.00008	-0.00038***	-0.00038***
	(0.00017)	(0.00014)	(0.00012)	(0.00010)
Over 80	-0.00008	-0.00003	0.00006	-0.00003
	(0.00010)	(0.00008)	(0.00007)	(0.00006)
Over 70	-0.00019**	-0.00019**	-0.00068***	-0.00071***
	(0.00009)	(0.00008)	(0.00007)	(0.00007)
Under 60	0.00010*	0.00016***	-0.00038***	-0.00033***
	(0.00005)	(0.00005)	(0.00004)	(0.00004)
Median Income	0.00739***	0.00815***	0.00762***	0.00781***
	(0.00062)	(0.00061)	(0.00051)	(0.00057)
% Black	-0.06167	-0.28296***	-0.13807*	-0.31623***
	(0.09653)	(0.07844)	(0.07258)	(0.06649)
% White	0.52819***	0.28052***	0.38771***	0.21728***
	(0.07844)	(0.06485)	(0.06200)	(0.05470)
% Hispanic	0.44125***	0.17361**	0.31071***	0.08867
	(0.09037)	(0.07200)	(0.06579)	(0.05836)
% Asian	1.04985***	0.67901***	0.82571***	0.55560***
	(0.14235)	(0.11043)	(0.10495)	(0.09016)
School District FE	Y	Y	Y	Y
Grade FE	N	Y	Y	Y
Year FE	N	Y	Y	Y
State FE	N	Y	Y	Y
R2	0.719	0.727	0.756	0.765
N	327581	327581	345914	345914

Table 3: Robust standard errors clustered at the school-district level. Specification 6-7 use math scores, and 8-9 use ELA scores. ***, ** indicates statistical significance at the 1% and 5% level, respectively.

4 Policy Implications and Conclusion

The impacts of global climate change and increasing temperatures are already being felt globally. The present research shows that failing to limit further greenhouse gas accumulation will lead to negative outcomes for student learning. Using school and grade specific data on student achievement on standardized tests for the entire United States, we find that each 1F increase in average temperatures results in lower student achievement in both mathematics and ELA tests. This result is robust across many model permutations. We further find that each additional day above 100F results in lower student achievement.

We couple the core finding of this paper with the recent IPCC report (IPCC (2018)) which states that a 1.5C average temperature increase is nearly eminent (barring substantial changes to current practices). This change in average temperatures is equivalent to the impact of living in an area with a median income that is $413 lower. For an area that might be in the upper tail of average temperature increases (4.5C or 8.1F), this is tantamount to living in an area with median incomes that are $1,239 dollars lower.

References

Dell, M., Jones, B. F. and Olken, B. A. (2009). Temperature and Income: Reconciling New Cross-Sectional and Panel Estimates, *American Economic Review* **99**(2): 198–204.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.99.2.198*

Deryugina, T. and Hsiang, S. M. (2014). Does the Environment Still Matter? Daily Temperature and Income in the United States, *Working Paper 20750*, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w20750*

Garg, T., Jagnani, M. and Taraz, V. (n.d.). Human Capital Costs of Climate Change: Evidence from Test Scores in India.
URL: *https://www.isid.ac.in/ epu/acegd2016/papers/TeevratGarg.pdf*

Goodman, J., Hurwitz, M., Park, J. and Smith, J. (2018). Heat and Learning, *Working Paper 24639*, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w24639*

Heal, G. and Park, J. (2013). Feeling the Heat: Temperature, Physiology & the Wealth of Nations, *Working Paper 19725*, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w19725*

Hemphill, F. C. and Vanneman, A. (2011). *Achievement Gaps: How Hispanic and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress. Statistical Analysis Report. NCES 2011-459*, National Center for Education Statistics.
URL: *https://eric.ed.gov/?id=ED520960*

Ho, A. D. and Reardon, S. F. (2012). Estimating Achievement Gaps From Test Scores Reported in Ordinal Proficiency Categories, *Journal of Educational and Behavioral Statistics* **37**(4): 489–517.
URL: *https://doi.org/10.3102/1076998611411918*

Hocking, C., Silberstein, R. B., Lau, W. M., Stough, C. and Roberts, W. (2001). Evaluation of cognitive performance in the heat by functional brain imaging and psychometric testing, *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology* **128**(4): 719–734.

IPCC (2018). IPCC - SR15, Global Warming of 1.5c.
URL: *http://www.ipcc.ch/report/sr15/*

Laurent, J. G. C., Williams, A., Oulhote, Y., Zanobetti, A., Allen, J. G. and Spengler, J. D. (2018). Reduced cognitive function during a heat wave among residents of non-air-conditioned buildings: An observational study of young adults in the summer of 2016, *PLOS Medicine* **15**(7): e1002605.
URL: *https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002605*

Lavy, V. (2015). Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries, *The Economic Journal* **125**(588): F397–F424.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12233*

Park, J. (n.d.). Hot Temperature and High Stakes Exams: Evidence from New York City Public Schools.
URL: *https://scholar.harvard.edu/files/jisungpark/files/paper$_n$yc$_a$ejep.pdf*

Reardon, S. F., Kalogrides, D. and Ho, A. D. (2018). Validation methods for aggregate-level test scale linking: A case study mapping school district test score distributions to a common scale, *CEPA Working Paper No. 16-09* .
URL: *https://cepa.stanford.edu/sites/default/files/wp16-09-v201807.pdf*

Reardon, S. F., Kalogrides, D. and Shores, K. (2017). *The Geography of Racial/Ethnic Test Score Gaps. CEPA Working Paper No. 16-10*, Stanford Center for Education Policy Analysis.
URL: *https://eric.ed.gov/?id=ED579685*

Reardon, S. F., Shear, B. R., Castellano, K. E. and Ho, A. D. (2017). Using Heteroskedastic Ordered Probit Models to Recover Moments of Continuous Test Score Distributions From Coarsened Data, *Journal of Educational and Behavioral Statistics* **42**(1): 3–45.
URL: *https://doi.org/10.3102/1076998616666279*

Vanneman, A., Hamilton, L., Anderson, J. B. and Rahman, T. (2009). *Achievement Gaps: How Black and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress. Statistical Analysis Report. NCES 2009-455*, National Center for Education Statistics.
URL: *https://eric.ed.gov/?id=ED505903*

Zivin, J. S. G., Hsiang, S. M. and Neidell, M. J. (2015). Temperature and Human Capital in the Short- and Long-Run, *Working Paper 21157*, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w21157*