

Heat and Learning*

Joshua Goodman
Harvard University
Joshua_Goodman@hks.harvard.edu

Michael Hurwitz
College Board
mhurwitz@collegeboard.org

R. Jisung Park
University of California, Los Angeles
jisungpark@luskin.ucla.edu

Jonathan Smith
Georgia State University
jsmith500@gsu.edu

April 11, 2019

Abstract

We demonstrate that heat inhibits learning and that school air-conditioning may mitigate this effect. Student fixed effects models using 10 million PSAT-retakers show hotter school days in years before the test reduce scores, with extreme heat being particularly damaging. Weekend and summer temperature has little impact, suggesting heat directly disrupts learning time. New nationwide, school-level measures of air-conditioning penetration suggest patterns consistent with such infrastructure largely offsetting heat's effects. Without air-conditioning, a 1°F hotter school year reduces that year's learning by one percent. Hot school days disproportionately impact minority students, accounting for roughly five percent of the racial achievement gap.

*Thanks to Lawrence Katz, Joseph Aldy, Joshua Graff-Zivin, Matt Neidell, Alan Barreca, Lucas Davis, Olivier Deschenes, Matt Kahn, Sarah Reber, Wes Yin, Manisha Shah, Greg Duncan and Patrick Behrer for valuable comments and feedback. We also thank seminar participants at Harvard, UCLA, UCSB, UCSD, IZA, CESifo, and West Point, as well as participants at the annual meetings of the AEA, SEA and SREE. Melanie Rucinski provided excellent research assistance.

1 Introduction

Hotter countries tend to be poorer, with each 1°F increase in average temperature associated with 4.5 percent lower GDP per capita (Dell et al., 2009). Students in hotter places also tend to exhibit lower levels of standardized achievement for any given age or grade. Across countries, a 1°F increase in average annual temperature is associated with 0.02 standard deviation lower math scores for 15-year-olds taking the 2012 Programme for International Student Assessment (PISA), as shown in the top panel of Figure 1. Within the United States, students in the hottest decile of the county-level climate distribution score on average 0.12 standard deviations worse on 3rd through 8th grade math tests, as seen in the bottom panel of that figure. Our primary contributions are to show that part of the cross-sectional relationship between temperature and academic achievement is causal, that heat’s cumulative learning impacts may be mitigated by school air-conditioning, and that differential heat exposure and learning impacts of heat account for a non-trivial portion of racial achievement gaps in the US.

Why might cumulative heat exposure reduce human capital accumulation? The contemporary US context of this study makes less relevant channels often studied in less developed settings, such as: health and disease burden (Bleakley, 2010; Cho, 2017); agricultural income and the opportunity costs of schooling (Shah and Steinberg, 2017); or institutional norms and political stability more broadly (Acemoglu et al., 2001; Dell et al., 2012; Hsiang et al., 2013). We provide evidence consistent with the possibility that, in the US, heat most likely affects learning directly by altering human physiology and cognition. Even moderately elevated temperatures can impair decision-making and cause substantial discomfort, and short-term impacts of heat on cognition have been extensively documented in laboratory settings (Mackworth, 1946; Seppanen et al., 2006). Hot classrooms may thus reduce the effectiveness of instructional time through physiological impacts on both students and teachers, making it harder for both to focus and accomplish a given set of learning tasks. In cases of extreme heat, schools may close or dismiss students early, directly reducing the amount of instructional time.

To estimate the causal impact of cumulative heat exposure on human capital accumulation, we link local daily weather data to test scores of 10 million American students from the high school classes of 2001-14 who took at least twice the PSAT, a nationally standardized exam designed

to assess students' cumulative learning in high school.¹ We also construct the first nationwide measures of school level air-conditioning penetration in the United States by surveying students and guidance counselors across the country about heat-related conditions in approximately 12,000 high schools. Student fixed effects regressions identify the impact of heat exposure during the prior school year by leveraging within-student variation in temperature over multiple test takes. Our identification strategy relies on the premise that variation in temperature over successive school years for a given student is uncorrelated with unobserved determinants of learning. We provide evidence consistent with that assumption, showing that selection into test-taking and retaking is not endogenous to temperature, even when controlling for regional trends in warming and secular changes in school quality or student composition.

We then generate three primary findings about the impact of heat on human capital accumulation. First, cumulative heat exposure reduces the rate of learning. A 1°F hotter school year in the year prior to the test lowers scores by approximately 0.2 percent of a standard deviation, or slightly less than one percent of an average student's learning gain over a school year.² Extreme heat is particularly damaging. Relative to school days with temperatures in the 60s (°F), each additional school day with temperature in the 90s (°F) reduces achievement by one-sixth of a percent of year's worth of learning. A day above 100 (°F) has an effect that is up to 50 percent larger. These effects are precisely estimated, are robust to controlling for test-day weather, and are not predicted by heat exposure in the year following the test. Only school-day exposure to higher temperatures affects test scores. Hot summers and weekends have little impact on achievement and controlling for such exposure does not shrink the magnitude of impact of hot school days. This suggests that heat's disruption of instructional time is responsible for the observed drop in test scores. That our effects are robust to controlling for heat-driven labor market shocks and pollution levels suggests that economic and health-related channels observed in other settings are likely not of first-order importance in this context (Cho, 2017; Garg et al., 2017).

Importantly, these learning effects appear to be cumulative and persistent beyond just the year prior to the test. Hot school days two, three and four years prior to the test also lower scores, so that the cumulative effect of elevated temperature over multiple school years is substantially

¹The PSAT is designed to be a test of knowledge and not of general intelligence or IQ.

²On average, students score 0.3 standard deviations higher on their second PSAT take than their first PSAT taken one year prior.

larger than that of a single school year. This suggests that any compensatory investments made by students, parents, or teachers in response to such heat shocks do not fully offset their impacts. Heat-related disruptions thus appear to reduce the rate of human capital accumulation over time. The implied magnitudes are non-trivial, particularly when considering predicted effects of climate change. For the average student, a sustained increase in temperature of 3.6°F (2°C) lowers achievement gains by two percent of a standard deviation, or approximately seven percent of an average year's worth of learning. This is despite relatively high average levels of income and air-conditioning in the United States compared to most other countries.

Our study shows that cumulative heat exposure can reduce the rate of human capital accumulation, and thus speaks to a longstanding debate on the relationship between geography and economic development (Acemoglu et al., 2001; Rodrik et al., 2004; Dell et al., 2012). A growing literature shows consistent evidence that the short-term impact of heat on exam days reduces cognitive performance but has not determined whether such effects are transitory or lead to permanent reductions in the stock of human capital (Graff Zivin et al., 2017; Park, 2017). Studies that provide evidence of medium-term impacts have occurred in contexts where heat's direct physiological effects are hard to distinguish from other channels such as agricultural output or health status (Cho, 2017; Garg et al., 2017). The only other paper that precisely identifies the impact of cumulative heat exposure on human capital accumulation is Isen et al. (2017), which focuses on *in utero* exposure and thus identifies a very different channel from the learning channel we study here. Our results are consistent with a longstanding lab-based literature documenting adverse cognitive impacts of hot temperature, the long-run implications of which have not previously been tested in real-world learning environments.

Our second major finding is evidence consistent with the possibility that school air-conditioning mitigates the impacts of heat on learning. We generate the first nationwide, school-specific air conditioning penetration rates, based on survey questions sent to students and high school counselors. We then show that, in the cross section, school air-conditioning penetration reported in 2016 is strongly related to heterogeneity in heat's effects, such that hot school days reduce learning by one-fifth as much in fully air-conditioned schools as in schools with no air conditioning. To better identify the mitigating effect of school air-conditioning, we use a triple-difference strategy that combines our within-student comparisons with within-school changes over time in re-

ported air-conditioning status. This approach reveals that later cohorts in schools that increased air-conditioning penetration experienced reduced heat-related learning impacts relative to earlier cohorts in those same schools and relative to cohort differences across schools that did not increase air-conditioning penetration. Both variation in the cross section and over time are thus consistent with the hypothesis that school air-conditioning reduces the negative impacts of hot school days. Our estimates, if taken as causal and combined with estimates of the achievement-earnings relationship from Chetty et al. (2011), suggest that in hot areas such as Houston, TX, the present value of air-conditioning is approximately \$2.1 million per year for each 1,000 student high school. Global climate change means however that the private return to air-conditioning will increase by approximately \$500,000 per school by 2040-2050 for the median US school, according to our back-of-the-envelope calculations.³

That school air-conditioning may improve learning speaks directly to a recent literature showing that school resources can have positive impacts on educational achievement (Jackson et al., 2015; Lafortune et al., 2018). Existing studies that focus on school infrastructure funding generally estimate the impact of broad funding packages and not targeted upgrades to specific school facilities (Cellini et al., 2010).⁴ Two recent studies finding positive achievement impacts of school infrastructure investment generally are consistent with the positive impacts of school air-conditioning we document. A comprehensive school construction project in New Haven, CT that included installation of air-conditioning in schools that lacked it raised reading scores by 15 percent of a standard deviation (Neilson and Zimmerman, 2014). Major school building and renovation efforts in the Los Angeles Unified School District raised reading and math scores between five and 10 percent of a standard deviation, most likely due to “reduced distractions from inadequate heating, cooling, or other aspects of the physical environment” (Lafortune and Schonholzer, 2017). The lack of achievement impact from recent school infrastructure investment in Texas is consistent with our data suggesting most schools in that state were likely air-conditioned to begin with (Martorell et al., 2016). Our paper is the first to attempt to isolate the impact of school air-

³Episodes of acute heat exposure are becoming more frequent in many parts of the world, and are predicted to increase at an accelerating rate (Stocker, 2014). Importantly, much of this warming will occur in places and during times of year that do not currently feature such temperature extremes, meaning that many local institutional arrangements – whether in terms of the timing of examinations or policies regarding protective built infrastructure (e.g. air-conditioning) – might not be efficiently adapted to new expected climate distributions.

⁴A notable exception is (Stafford, 2015), who provides evidence that mold remediation and ventilation improvements had positive impacts on performance in 66 Texas elementary schools.

conditioning from other aspects of school infrastructure. These results also highlight the fact that small reduced form impacts of heat in developed countries may mask large biological and welfare impacts, underscoring the importance of taking adaptations such as air-conditioning into account when studying the relationship between climatic variables and economic outcomes (Graff Zivin and Neidell, 2014; Deschenes et al., 2017).

Our third and final result is that the temperature environment in which learning occurs appears to be a meaningful contributor to racial and geographic achievement gaps. Black and Hispanic students' learning is roughly three times as inhibited by the prior school year's heat compared to the learning of White students. We estimate that between three and seven percent of the gap in PSAT scores between White students and Black and Hispanic students can be explained by differences in the temperature environment experienced by students in each group. These disparities appear to be driven both by school-level investments, such as differential air-conditioning penetration, as well as the geographic distribution of racial minorities whereby Black and Hispanic students overwhelmingly reside in hotter locations than White students. This is consistent with the theory of residential sorting, which suggests that lower income individuals are more likely to sort into areas with lower levels of environmental amenities and local public goods (Tiebout, 1956; Roback, 1982; Banzhaf and Walsh, 2008). We also find that historically cool places experience much larger impacts per unit of warming. This suggests that, without high average levels of adaptation in the South, the North-South achievement gap would be even larger than it is currently. Minority students appear, however, less likely to have school air-conditioning even in a given historical climate.

That heat has larger impacts on Black and Hispanic students and accounts for a non-trivial proportion of the racial achievement gap are facts not documented previously in the vast literature on racial disparities in educational outcomes (Jencks and Phillips, 2011; Duncan and Murnane, 2011; Fryer Jr, 2011). Most empirical analyses in this literature have focused on social factors such as teacher quality, racial bias, or neighborhood culture, and have more rarely explored disparities in the physical environment. "Environmental factors" in this literature often explicitly or implicitly refer to the social environment or use the term to denote non-genetic factors. The few studies that assess the impact of physical environmental factors on achievement are either non-identified cross-sectional analyses (Durán-Narucki, 2008) or assess short-run outcomes such as cognitive

performance during a hot period (Mackworth, 1946; Seppanen et al., 2006; Park, 2017). A substantial proportion of the variation in achievement remains unexplained by traditional socioeconomic variables (Reardon et al., 2017). Our findings suggest that the physical factors such as temperature and the built environment may play a larger role in explaining disparities in achievement than previously realized.

Our empirical strategy relies on temperature fluctuations across multiple years as opposed to multiple decades. Our central estimates of the impact of heat and of air conditioning as a mitigating factor are thus driven by these short to medium term changes in the efficacy of instructional time. Longer-term exposure to heat may change students' study and other behaviors in ways we do not capture. School air conditioning may, for example, be a workplace amenity capable of attracting higher quality teachers and thus affecting student outcomes over time through an improved teacher workforce. Our estimates will not capture such longer-term impacts of a school's physical environment on student outcomes and thus may underestimate the total impacts of such investments.

2 Data and Empirical Strategy

2.1 Data

We combine three primary data sources on test scores, temperature and school air-conditioning. Test score data come from the College Board, which administers the PSAT exam to millions of American high school students annually. The PSAT consists primarily of a reading and a math section.⁵ The test is offered once a year roughly during the third week of October and most students first take it in 10th or 11th grade, though some start as early as 9th grade.⁶ Students take and can retake the PSAT for a variety of reasons, including preparation for the SAT college entrance exam, qualification for National Merit Scholarships, and provision of information about their college readiness to them and their schools. The PSAT has multiple advantages over other tests used

⁵A writing section has been added in more recent years. Basic scientific concepts and history are assessed as part of the reading comprehension section.

⁶We use the PSAT rather than the SAT for two reasons. First, the SAT is offered at seven different times of the year, making it harder to assign easily comparable measures of long-term heat exposure to a given exam take. Second, the PSAT is taken by roughly twice as many students as the SAT because the latter is taken in later grades by a more college-oriented and thus selected set of takers.

to study the impact of heat on cognitive skills, including: the test is given once a year on a fixed date with advanced registration required, limiting the scope of endogenous taking or timing decisions; proctoring is nationally harmonized and the test is centrally graded, limiting potential for endogenous score manipulation of any sort (Dee et al., 2016; Park, 2017); and the test is designed to assess cumulative progress on skills learned during secondary school, rather than generalized intelligence, making it arguably better-suited for assessing the effects of formal schooling.

We have test scores and dates from all takes for the universe of PSAT-takers from the intended high school classes of 2001-2014. Our primary outcome is a student's combined math and reading score, standardized by test administration. The student-level data also contains basic demographic information such as gender, race, parental education, and residential ZIP code, which we use to assign students ZIP code level mean incomes as reported by the Internal Revenue Service. High school identifiers allow us to assign students to testing locations and other school-level characteristics. We observe only the most recent high school attended and thus permanently assign that location to each student across all PSAT takes. Doing so reduces concerns about potentially endogenous migration but may generate measurement error in the heat exposure treatment variable. The magnitude of that error is, however, likely quite small. Data from the 2005-12 American Community Survey show that, among 16-17 year-old high school students, only 2.8 percent lived in a different Public Use Microdata Area in the year prior and fewer than one percent lived in a state not contiguous with their current one. Very few high school students substantially change temperature zones in the course of a year.

Daily temperature data come from the National Oceanic and Atmospheric Administration's Daily Global Historical Climatology Network, which includes station-level data for thousands of weather stations across the United States. We focus on the subset of nearly 3,000 weather stations with daily temperature data available for at least 95 percent of the days from 1996 to 2014, the time period covering potential test-taking dates of our sample. Doing so allows us to assign each high school to a single, stable weather station over the entire time period, which avoids endogeneity concerns driven by the possibility that stations coming online or going offline are somehow correlated with local population growth, economic conditions or temperatures conditions in ways that might contaminate our estimates (Auffhammer and Mansur, 2014). We impute the small proportion of missing daily observations with those from the nearest stations with non-missing data. We

assign each high school to the nearest weather station, resulting in an average distance of 9.7 miles between a student’s test site and weather station being used to measure temperature at that site.

We construct two primary measures of cumulative heat exposure experienced by a student: the average daily maximum temperature and the number of days that temperature exceeded a given multiple of 10°F in the 365 days prior to the test.⁷ We use daily maximum temperature because schooling occurs during the daytime when such temperatures usually occur.⁸ We focus particularly on temperature experienced on school days, treating summer and weekend temperatures as separate sources of variation.⁹ We also use the weather stations to construct test date temperature, rain and snowfall, as well as cumulative rain and snowfall exposure over the year prior to the test, which help account for potential independent effects of such precipitation (Goodman, 2014).

School-level air-conditioning data do not exist at the national level and very rarely exist at the state and local level. We generate measures of school-level air-conditioning penetration through a survey the College Board regularly administers by e-mail to all SAT takers and to high school guidance counselors registered to administer PSAT or other exams. In 2016, we added to the survey the statement “On hot days, classrooms get too hot.” Respondents could choose “Never”, “Some of the time”, “Most of the time”, or “All the time”.¹⁰ We received valid responses from students in nearly 12,000 schools enrolling 87 percent of our sample’s PSAT-takers. To construct school-level measures from individual responses, we assign the four possible responses a value of 0, $\frac{1}{3}$, $\frac{2}{3}$, and 1, then average across all students within a school. We interpret this measure as the fraction of a school’s classrooms with adequately functioning or sufficiently frequently activated air-conditioning. This measure of school air-conditioning penetration has the advantage of being reported by the students themselves and of being based on the largest set of responses to our ques-

⁷Focusing on the prior year implies that the measured heat exposure occurred prior to a given PSAT administration but after the most recent one before that. Because of slight annual variation in the timing of the PSAT, we exclude the third week of October from these measures to guarantee no overlap between PSAT administrations and the constructed measure of heat exposure. The year prior to the test take thus runs from late October of the preceding year to mid-October of the current year.

⁸We note that it is possible for nighttime heat to affect learning through disrupted sleep as well. To the extent that daytime high and nighttime low temperatures are correlated, it is possible that our estimates may include some impacts due to disrupted sleep, though the results on weekend days versus weekdays is suggestive of effects driven primarily by instructional time, as we discuss below.

⁹No comprehensive national data set of school calendars covering this time period exists, so we assign to each student a likely school start and end date based on the calendar of the largest school district in that student’s state, as seen in Figure A.1. We then divide the year into three periods: school days, the summer, and weekends or national holidays.

¹⁰Respondents were also allowed to choose “I don’t know.” We coded such answers as missing.

tions about air-conditioning. It has the disadvantage of measuring air-conditioning penetration at a single point in time and thus may correlate with other unobserved school-level factors.

A different measure of school air-conditioning penetration comes from our addition to the survey of two questions posed to high school guidance counselors: “How many of your school’s classrooms have air-conditioning?” and “Ten years ago, how many of your school’s classrooms had air-conditioning?” Guidance counselors could respond with “None”, “Fewer than half”, “About half”, “More than half”, and “All”. To construct school-level measures, we assign these possible responses a value of 0, $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$, and 1, then average across all guidance counselors within a school. This measure has the disadvantage that we received responses from guidance counselors in only about 2,000 high schools. It has the advantage, however, of directly asking about air-conditioning and of allowing us to measure air-conditioning penetration at two points in time instead of one. Variation in air-conditioning penetration over time is more plausibly exogenous than cross-sectional measures, helping us construct stronger causal claims about the impact of school air-conditioning in offsetting heat’s impacts. Air conditioning upgrades may nonetheless be correlated with other school improvements that protect against high temperatures.

Finally, we supplement our three primary data sources with additional data on residential air-conditioning, local economic conditions and pollution levels. We construct county-by-year data on residential air-conditioning penetration by combining county-level residential air-conditioning penetration estimates from the 1980 decennial census with data on changes in such penetration over time by census region from the Energy Information Agencies Residential Energy Consumption surveys.¹¹ We estimate county-level economic conditions by constructing the logarithm of annual payroll per capita from the U.S. Census Bureau’s County Business Patterns, focusing on sectors that the National Institute of Occupational Safety and Health categorizes as being highly exposed to weather (namely: construction, mining, transportation, manufacturing, agriculture and utilities). Based on evidence of the adverse impact of hot days on highly exposed sector payroll, we use this measure to control for local economic shocks driven by annual fluctuations in heat (Deryugina and Hsiang, 2014; Behrer and Park, 2017). Similarly to our temperature measures, we

¹¹We use the reported penetration rates in 1980 as a basis and then extrapolate based on the region-level growth rate of total air-conditioning penetration recorded by RECS, which provide penetration rates by region from 1980 to 2009 with a two or three-year frequency. We linearly interpolate growth rates for the missing years and assign counties their corresponding regional growth rate. Using this growth rate and the observed penetration rate in 1980 we create a measure of penetration in every county in each year from 1980 to 2011. We top-code penetration at 100 percent.

construct both cumulative and test date measures of exposure to major air pollutants (ozone, sulfur dioxide, nitrogen dioxide, carbon monoxide, and PM10 particulate matter) using data from the Environmental Protection Agency’s Ambient Air Monitoring network. Previous research suggests that such pollutants can affect student absenteeism and performance, at least in the short run (Currie et al., 2009; Ebenstein et al., 2016), though there is relatively less evidence regarding the impact of longer-term exposures.

2.2 Summary Statistics

As shown in Table 1, the starting sample comprises over 38 million test scores from 27 million students in the high school classes of 2001 to 2014, who took the PSAT as early as October of 1998 and as late as October of 2012.¹² Because our primary identification strategy relies on within-student variation in heat exposure for identification, we focus on the subset of 21 million scores from the nearly 10 million students who took the PSAT at least twice.¹³ Those retakers are slightly more advantaged than the general pool of PSAT takers, living in higher income zip codes, more likely to have mothers with college degrees, and with first PSAT scores 0.14 standard deviations about the national average. Importantly, retakers seem geographically similar to the general pool of SAT takers. Both groups experience school days with mean temperature in the mid 60s°F in the year prior to the test, with an average of 12 school days above 90°F. Maps of takers and retakers locations also suggest that both groups have similar geographic distributions in line with the population distribution across the United States.¹⁴ On average, PSAT takers and retakers report that nearly 60 percent of their classrooms are adequately air-conditioned, with closer to 80 percent likely to have air-conditioning at home. Dividing the retaking sample by race and by income reveals that black and Hispanic students and lower income students score lower on the PSAT than their white and higher income counterparts.

The raw data suggest a strong negative relationship between cumulative heat exposure and academic achievement. Figure 2 maps average PSAT performance by county. On average, Southern counties have substantially lower test scores than do Northern counties. This tracks closely

¹²We exclude a very small number of observations of PSATs taken during October of 12th grade.

¹³In our sample, 85 percent of students take the PSAT twice and 15 percent take it three times, the maximum given testing opportunities in 9th, 10th and 11th grades.

¹⁴See Figure A.2.

with differences in heat exposure by geography, as seen in Figure 3. In Florida and southern Texas, the average school day experienced by a student is above 80°F, compared to an average school day temperature in the 50s°F in much of the Northern U.S. Southern students in many counties experience 30 or more school days above 90°F, compared to fewer than 5 such days in nearly all Northern counties. The existence of such a strong North-South temperature gradient and test score gradient could be causal or could be driven by other important regional differences. Our goal is disentangle the impact of heat from such other factors.

2.3 Empirical Approach

We exploit the fact that for students who take the PSAT multiple times, differences across takes in prior year heat exposure are likely uncorrelated with differences in other factors that might affect academic achievement. We thus ask whether students score lower immediately following a hotter year relative to their own score immediately following a cooler school year and, if so, we argue that heat is the only factor that can explain such a difference in outcomes. We implement this identification strategy with student fixed effects regressions of the following form:

$$Score_{iscyn} = \beta Heat_{sy} + \eta_i + \gamma_{cyn} + \epsilon_{iscyn} \quad (1)$$

Here, $Score$ denotes the standardized PSAT math and reading score for student i in high school s , high school class c , taking the test in October of year y for the n th time. Inclusion of student fixed effects η implies that identification comes from within-student comparisons of heat exposure and test score differences over multiple takes.¹⁵ We define cumulative heat exposure $Heat$ as the average maximum temperature experienced during school days in the year prior to the test, for all students in high school s taking the test in year y . In that case, the coefficient of interest β can be interpreted as the standard deviation impact on a student’s test score of experiencing a one degree F hotter school year on average. High school class by test year by take number fixed effects γ flexibly control for a variety of potential confounds, including differential selection into test-taking

¹⁵We use student fixed effects rather than high school fixed effects because the latter approach depends on the assumption that selection into PSAT-taking at the school level does not vary over time in ways correlated with heat exposure. This assumption fails empirically because, over the time period in question, PSAT taking expands to a wider set of students and moreso in regions of the country that are differentially affected by longer-term warming trends. The student fixed effects approach avoids this selection margin entirely.

across high school classes, differential difficulty of the test across test dates, and differential test performance based on past number of test takes. We cluster standard errors by weather station, the level of variation in our treatment variable.

We also use a specification that replaces this single heat exposure measure with a vector of counts of the number of school days falling into various temperature bins:

$$\begin{aligned} Score_{iscyn} = & \beta_1 DaysAbove100_{sy} + \beta_2 DaysIn90_{s_{sy}} + \beta_3 DaysIn80_{s_{sy}} + \beta_4 DaysIn70_{s_{sy}} \\ & + \beta_5 DaysBelow60_{sy} + \eta_i + \gamma_{cyn} + \epsilon_{iscyn} \end{aligned} \quad (2)$$

In this specification, the coefficient on days in the 90s can be interpreted as the impact of experiencing one more very hot school day, relative to a school day with temperature in the 60s. This specification allows us to study non-linearities in the effect of school year temperature.¹⁶ Identification therefore relies on both spatial variation in heat exposure, as previously shown in Figure 3, as well as temporal variation in heat exposure.¹⁷ To understand the magnitude of such identifying variation, we compute the residual from regressing heat exposure on the aforementioned student and class-year-take fixed effects. The distribution of such residuals suggests that a one standard deviation increase in mean school day temperature is about 1°F, while a one standard deviation increase in the number of schools day above 90°F is about three days.

One potential threat to identification comes from the possibility that cumulative heat exposure drives selection into taking the PSAT for the first time or choosing to retake it. To test for selection into taking the PSAT the first time, we collapse the data by high school and regress the number of first-time test-takers (as well as its logarithm) on high school fixed effects and cohort by test date fixed effects. We see no evidence that within-school fluctuations in annual heat exposure affects the number of test takers and can rule out economically meaningful effect sizes. We then run similar student-level regressions using the demographic characteristics of first-time test-takers as outcomes. The results rule out meaningful impacts of heat exposure on the observable composition of the test-taking population, particularly when controlling for differential heat trends by state.¹⁸ Finally, we ask whether heat in the year prior to the first take or in the year following

¹⁶We do not find strong evidence that cold weather affects learning in our sample, hence our focus on the upper end of the temperature distribution.

¹⁷See Figure A.3 for annual variation in average school day temperatures.

¹⁸See Table A.1 for detailed estimates.

the first take affects the probability that a student retakes the PSAT. We again find no evidence of such selection, with point estimates suggesting a 1°F hotter school year increases the probability of retaking by 0.05 percentage points and confidence intervals that rule out effects larger than 0.15 percentage points.¹⁹ In total, these results suggest little evidence of endogenous selection into test-taking or retaking as a result of cumulative heat exposure.

3 The Impact of Cumulative Heat Exposure

3.1 Prior Year Impacts

On average in the U.S., experiencing a 1°F hotter school year lowers academic achievement by 0.002 standard deviations, a result that is very precisely estimated and robust to a variety of controls for potential confounding factors. Table 2 shows these results. The first column of panel A contains the baseline specification described in equation 3, where the test score outcome is measured in hundredths of a standard deviation. The coefficient is highly statistically significant and precise enough to rule out effects smaller than 0.001 standard deviations. The magnitude of these impacts is small enough to have been missed by previous studies with less precision but large enough to imply non-trivial cumulative effects of temperature on learning. For example, the average gain in PSAT score performance between 10th and 11th grade is about 0.3 standard deviations, implying that a 1°F hotter school year reduces learning by close to one percent of the expected gains over that year. As we will see below, learning impacts are even larger for certain demographic subgroups and in the absence of air-conditioning.

That students score lower following hotter years relative to their own scores following cooler years does not appear to be driven by other channels potentially correlated with heat in the school year leading up the exam, as seen in the next four columns of Table 2. Controlling for total snowfall and rainfall in the prior year and for temperature and precipitation on the day of the exam has nearly no effect on the point estimate. This suggests that we are not mistakenly attributing to cumulative heat exposure effects that are actually driven by cumulative precipitation exposure or by contemporaneous heat exposure. Similarly, controlling for both cumulative and contemporaneous pollution exposure leaves our estimate nearly unchanged, implying that we are measuring

¹⁹See Table A.2 for detailed estimates.

the direct impact of heat and not of pollutants such as ozone that may be more common on hot days.

Controlling for county-level payroll in industries highly exposed to weather conditions also does little to our point estimate, suggesting that cumulative heat exposure is not operating through the channel of family income or local economic conditions.²⁰ The robustness of our estimate to controls for state-specific time trends suggests we are not picking up spurious correlations driven by subtle geographic differences in warming trends that may be correlated with other local changes in selection into or preparation for PSAT-taking or retaking.²¹ Regardless of which of the aforementioned controls are included, the estimated impact of a 1°F hotter school year, which represents a roughly one standard deviation change in cumulative heat exposure, is never substantially different from 0.002 standard deviations.

Given that the mean distance between weather stations and high schools in our data is just under 10 miles, the cumulative heat exposure we assign to each student may be mismeasured, particularly for students farthest away from weather stations. The final column of Table 2 limits the sample to high schools within five miles of a weather station, for which we assume that measurement error is less of an issue. For students whose high schools are particularly close to weather stations, the impact of cumulative heat exposure on academic achievement is about 25 percent larger than for the sample as a whole. Interacting heat effects with distance from the nearest sensor also suggests students farther from sensors see less measured impact from heat, with the model predicting that those in the immediate vicinity of a sensor would experience impacts similar to the estimate in that final column of Table 2.²² This is consistent with the possibility that measurement error in our treatment variable is generating downward bias in our estimates, though it is also consistent with the possibility of heterogeneous treatment effects correlated with factors that make weather stations more likely to be online near schools.

In addition to estimating the impact of generally hotter school years, we also show that very hot days are particularly damaging to student achievement. Panel B of Table 2 shows the specification in which we measure cumulative heat exposure by counts of school days falling into various

²⁰This is perhaps unsurprising given the developed country context but contrasts with research in developing countries that shows agricultural yield shocks driving schooling outcomes (Garg et al., 2017; Shah and Steinberg, 2017).

²¹The estimates are also robust to using quadratic or cubic trends instead of linear trends.

²²See column 1 of Table A.3.

temperature bins. Replacing a school day in the 60s with a hotter school day lowers achievement, with the extent of that damage increasing roughly linearly with temperature above 70°F. Consistent with our baseline specification, these estimated magnitudes also imply that a one standard deviation increase in heat exposure, or over three additional days above 90°F, lower achievement by 0.002 standard deviations. Cold days, those below 60°F, appear to have little impact on achievement. Figure 4 shows these point estimates with further disaggregation of colder school days and, consistent with laboratory studies on cognition and recent studies on labor supply and mortality, shows damage that increases with temperature starting around 70°F. Estimated impact of hot days are, like those for mean heat, robust to controls for cumulative precipitation and pollution, test day weather and pollution, local economic conditions and state-specific trends.

Heat's cumulative impact on academic achievement is not driven by one particular subject, in contrast to findings focusing on short-run cognitive impacts (Graff Zivin et al., 2017; Garg et al., 2017). Both math and reading scores drop by similar magnitudes for a given level of additional heat exposure.²³ The effect is also not driven by one particular test take. Heat prior to a first test take has relatively similar negative effects on achievement as heat prior to second take, although there is some evidence that heat's learning impact, if anything, increases with take number.²⁴ This eliminates the possibility that our results are driven by differential selection into retaking based on correlations between heat exposure, first take performance and students' belief about whether their first scores reflect their true abilities. It also makes unlikely the possibility that our results are driven by subtle warming trends over time. Using future temperature shocks as a placebo test also yields results consistent with our interpretation of these impacts as causal. Controlling for mean school day temperatures in the years one, two and three years after the exam does little to change our estimated impact of cumulative heat exposure and the coefficients on future temperature are much closer to zero than our main effect and never statistically significant.²⁵ This makes it less likely our results are a statistical artifact driven by spurious correlations between temperature and test scores.

The statistical significance of our estimates is not particularly sensitive to how we account for potential spatial correlation in temperature shocks. Our default specification clusters standard

²³See Table A.4 for details.

²⁴See Table A.5 for details.

²⁵See Table A.6 for details.

errors by weather sensor, following the suggestion by Abadie et al. (2017) to cluster at the level of treatment variation. Given potential spatial correlation between heat shocks across distinct but nearby weather sensors, we cluster by state to allow for correlations within much broader geographic areas. This very conservative approach increases our standard errors by at most 30-50 percent, so that our central estimates of the impact of average heat or days above 90°F remain highly statistically significant.²⁶ Following Conley (1999), non-parametrically adjusting our standard errors to account for spatial correlation does not inflate our standard errors significantly, again leaving our main inferences unchanged.

3.2 Productivity of Instructional Time

One indication that cumulative heat exposure affects achievement by directly lowering the productivity of instructional time comes from examining the impact of heat on three mutually exclusive sets of days in the year prior to the test: school days, weekends and national holidays during the school year, and summer days. The first two columns of Table 3 contrast our baseline specification to one in which we control for heat on summer days. Two important facts emerge. First, controlling for summer heat has little effect on our estimated impacts of school day heat. Second, the impact of summer heat on academic achievement is very small and statistically indistinguishable from zero. That summer heat has no impact on academic achievement seems to rule out potential channels such as student health or local economic conditions given that such channels should be affected by summer heat as well as school day heat.

Controlling for school year weekend and holiday heat somewhat increases the estimated effects of school day heat. We see no evidence that weekend and holiday heat lowers student achievement. This is again consistent with the possibility that time in school is critical to understanding the relationship between heat exposure and human capital accumulation. One challenge in interpreting the coefficients in column 3 of Table 3 is that residual variation in weekend heat is highly collinear with variation in school day heat. Summer heat variation, by contrast, is substantially less collinear with school day heat variation. The large, positive coefficients on weekend heat could represent a real impact of substitution away from leisure time and toward study time. They may also simply be an artifact of that high degree of collinearity. Regardless, controlling for

²⁶See Table A.7.

weekend heat does not affect the conclusion that school day heat affects academic achievement.

One further test yields results consistent with our interpretation that heat exposure interferes with actual learning. In column 4 of Table 3, we break heat exposure into three time periods corresponding to distance in time to the test: post-summer school days (the roughly two months just prior to the PSAT), summer days (roughly three to five months prior to the PSAT), and pre-summer days (roughly six to twelve months prior to the PSAT). Summer heat again has little clear impact on achievement while both pre- and post-summer heat have large negative impacts. That the damage from pre-summer heat is as large as or larger than the impact of post-summer heat suggests that heat operates not just through periods where students might be “cramming” for a test. Instead, heat appears to affect cumulative learning over a longer time frame.

3.3 Persistent and Cumulative Impacts

To provide evidence about the impacts of heat exposure further back in time than the prior school year, we augment our regression specifications 3 and 2 with temperature measures from school days in the years starting two, three, four and five years prior to each exam take. Equation 3 becomes:

$$\begin{aligned}
 Score_{iscyn} = & \beta_1 Heat1YearAgo_{sy} + \beta_2 Heat2YearsAgo_{sy} + \beta_3 Heat3YearsAgo_{sy} \\
 & + \beta_4 Heat4YearsAgo_{sy} + \beta_5 Heat5YearsAgo_{sy} + \eta_i + \gamma_{cyn} + \epsilon_{iscyn}
 \end{aligned} \tag{3}$$

This accomplishes two things. First, our estimates of prior year heat exposure from Equations 3 and 2 are downward biased if such effects persist beyond one year, and controlling for further lagged measures of temperature should help eliminate that downward bias. Failing to control for dynamic effects can cause our student fixed effects approach to understate the impact of heat because that approach implicitly assumes complete decay of effects after one year, so that heat before the first take affects only the first take and not the second take itself. Second, the estimated impacts of earlier years’ heat exposure allows us to more accurately estimate longer run impacts of exposure to different temperature environments. For example, the sum of the coefficients β_1 through β_5 estimates the test score impact of a one degree increase in the average temperature over a student’s five prior school years.

Table 4 shows these estimates, with coefficients for individual years shown in each column and the sum of those coefficients shown at the bottom of each column. Two important facts emerge. First, controlling for earlier heat exposure increases our estimates of the impact of prior year heat by 20 to 40 percent, suggesting that such estimates may have been downward biased by our baseline specification's identifying assumption that prior year heat's impact fades out entirely after that year. Second, we see fairly clear evidence that heat up to four years prior affects test scores. This evidence is clearer with respect to the impact of hot days than the impact of average temperature but is present in the point estimates of both. We see no evidence that heat five years prior has impacts on test scores and tests of further lags (not shown here) also show no clear impact. These estimates suggest that heat effects are somewhat more persistent than teacher effects, which the literature estimates fade out substantially within one to two years (Kane and Staiger, 2008; Jacob et al., 2010; Rothstein, 2010).

We take the cumulative impact over the past four years as our favored estimate of cumulative impacts for two reasons. First, because five year and further lags have little apparent impact on test scores, cumulative impacts tend to level off after four years. Second, we are estimating lagged heat exposure based on the high school a student is enrolled in at the time of taking the PSAT. Because we do not know where a student attended primary school, going back much further in time likely increases the measurement error of our treatment variable. We therefore take the four-year cumulative impact as a lower bound on the total impact but one that our best estimates suggest is likely not far from the truth. The cumulative impact of increased heat exposure over multiple years is thus three to five times larger than the impact of just the prior year's heat. Our estimates suggest that a 1°F increase in the average temperature of the past four school years leads to about a 0.006 standard deviation decrease in test scores, or 2 percent of the typical increase in PSAT scores over a single school year. Experiencing one additional day above 90°F in each of those four years would reduce scores by between 0.002 and 0.003 standard deviations, or 1 percent of a typical school year's PSAT increase.

4 School Air-Conditioning as a Defensive Investment

4.1 Descriptive Analysis

Adaptive responses and defensive investments are important for understanding the welfare implications of environmental shocks, particularly in the long run (Graff Zivin and Neidell, 2013; Kahn, 2016; Deschenes et al., 2017). School air-conditioning represents one potential investment.²⁷ To explore its role in mitigating the effects of heat exposure, we first provide descriptive analysis of the prevalence of school air-conditioning across the U.S.

Figure 5 shows county-level averages of school air-conditioning penetration as measured by the extent to which students (panel A) and guidance counselors (panel B) report that “On hot days, classrooms get too hot.” The resulting map is roughly the inverse of an average temperature map. Students and counselors are much less likely to report hot classrooms in the hotter regions of the country and much more likely to report hot classrooms in cooler regions. Students in the Northeast, for example, report that heat interferes with learning on the majority of hot days. Students in the South report heat interfering with learning on only about one-fourth of hot days.

Because this reporting may partly reflect the extent to which students and counselors are accustomed to heat, rather than actual air-conditioning status, we ask guidance counselors directly about the fraction of classrooms with air-conditioning. That map, in panel A of Figure 6, looks quite similar to the map of hot classrooms. According to guidance counselors, nearly all classrooms in the South have air-conditioning while the majority of classrooms in the Northeast lack it. We therefore interpret student reports of hot classrooms as a measure of air-conditioning penetration. That our student-generated measure based on hot classrooms has a nearly 0.7 correlation with the counselor-generated responses to the direct question about air conditioning suggests the former is largely measuring actual school air conditioning.²⁸

These various measures are the first we know of to provide national school-level estimates of

²⁷Teachers and parents seem to value school air-conditioning, judging by recent labor disputes and community petitions. During a major teacher strike in Chicago in 2012, “Timetable for air-conditioning” was one of four major contract demands, with an agreement to provide universal air-conditioning eventually reached in 2016. Parents and teachers in a number of major school districts such as New York City, Los Angeles, and Denver have have recently signed petitions asking districts to upgrade air-conditioning equipment. See: <http://www.denverpost.com/2011/09/08/heat-related-illnesses-spur-petition-for-sept-school-start-in-denver/>.

²⁸See Figure A.5, which shows a tight relationship between counselor reports of air conditioned classrooms and student reports of classrooms too hot for learning.

air-conditioning status. We also note, as seen in panel B of Figure 6, that both home and school AC seem to vary substantially by region, with lower penetration rates particularly in more mountainous regions of the country.

4.2 School Air-Conditioning as Potential Mitigator

We use two approaches to assess whether school air-conditioning can offset heat’s adverse impact on learning. First, we interact the cross-sectional measure of air-conditioning penetration reported by students with our heat exposure measure and add it to our baseline specification. This regression has the form:

$$Score_{iscyn} = \alpha Heat_{sy} + \beta Heat_{sy} * SchoolAC_s + \eta_i + \gamma_{cyn} + \epsilon_{iscyn} \quad (4)$$

The coefficient α now has the interpretation of the impact of heat on a school with no air-conditioning, while β represents the predicted difference between that impact and the impact on a fully air-conditioned school.

This analysis, if taken as causal, suggests that school air-conditioning almost fully offsets the impacts of cumulative heat exposure. In column 1 of Table 5, the main coefficient implies that, for students in schools with no air-conditioning, a 1°F hotter school year lowers achievement by 0.0032 standard deviations. The interaction coefficient suggests that, in fully air-conditioned schools, this effect is 0.0025 standard deviations smaller. For the average student, school air-conditioning thus appears to offset 73 percent of the learning impact of hot school days. This interaction coefficient may represent the causal impact of school air-conditioning but it may also be picking up effects of other factors correlated in the cross-section with a school’s air-conditioning status.

In particular, school air conditioning may be correlated with other adaptations students have made to deal with hot environments or with wealth or other resources available to students that independently improve their achievement. To account for the possibility that school air conditioning is proxying for other adaptive measures, we add to equation 4 additional terms in which heat exposure is interacted with county-level home air-conditioning rates and mean school year temperature. To account for school air conditioning as a proxy for other academic resources available

to students, we also control for such interactions with ZIP code level income, the racial composition of a school’s PSAT takers, and per-pupil district-level school capital expenditures, the last of which helps account for the possibility that schools with air conditioning have made other capital investments that boost achievement.²⁹ The results of this augmented specification are shown in column 2 of Table 5. The topmost coefficient now suggests that, for a student with neither school nor home air-conditioning, a 1°F hotter school year lowers achievement by 0.0057 standard deviations. School air-conditioning and home air-conditioning respectively offset 41 and 57 percent of this effect, implying that a student with access to both would see little negative impact of heat exposure. That the magnitude of the school air-conditioning coefficient does not change substantially with the addition of these controls implies that omitted variable bias from such sources is unlikely to explain the observed effect.

We make one further attempt to isolate the impact of school air-conditioning. To do so, we use the change over time in penetration rates implied by differences in counselors’ reports about their schools’ air-conditioning status in 2016 versus 10 years before that. We assign to students a variable *SchoolACchange*, which represents their high school cohort’s change in air-conditioning penetration rate relative to 2006 implied by the counselors’ answers. Cohorts from 2006 and earlier are assigned a value of zero and more recent cohorts are assigned a change linearly interpolated from the counselors two responses. We then run the following specification:

$$Score_{iscyn} = \beta Heat_{sy} * SchoolACchange_{sc} * HSClass_i + \delta Heat_{sy} * SchoolACchange_{sc} + \mu Heat_{sy} * HSClass_i + \nu Heat_{sy} + \eta_i + \gamma_{cyn} + \epsilon_{iscyn} \quad (5)$$

Here, *HSClass* is a continuous measure of a student’s cohort. In effect, the coefficient β from this triple-difference approach estimates whether schools that have installed additional air-conditioning over time see the impact of heat shrink across cohorts, relative to schools that have not added air-conditioning.

The results in column 3 of Table 5 suggest that this is the case. The coefficient on the triple interaction term is positive and highly statistically significant, implying that later cohorts do see

²⁹We use sample-demeaned versions of these control variables so that coefficients can be interpreted as impacts on students with average school temperatures, income, racial composition and capital expenditures.

smaller impacts of heat in schools that improved air-conditioning penetration, relative to schools that did not. This conclusion is unchanged when we control for additional interactions with home air-conditioning and the other controls described previously. This provides additional evidence consistent with the possibility that school air-conditioning mitigates the impacts of heat exposure. The main threat to validity here is that a school's adoption of air-conditioning correlates with other unobserved changes over time in that school that might also mitigate the impacts of heat (e.g. more tree cover, improved ventilation). Though possible, it seems likely that changes in school air-conditioning penetration are more exogenous than cross-sectional variation in such penetration. That both approaches yield consistent results suggests school air-conditioning may mitigate a substantial portion of the learning impacts of heat exposure.

To provide a back-of-the-envelope estimate of the monetized value of school air-conditioning, we apply previous estimates of the relationship between test scores and later life earnings from Chetty et al. (2011). That paper finds that having a teacher who raises test scores by 0.1 standard deviations results in a net present value of \$8,500 in future increased earnings for current 16 year olds.³⁰ Our estimates in Table 5 suggest that school air-conditioning offsets about 0.0025 standard deviations in learning damage for each 1°F increase in temperature. This translates into a net present value of \$212 in recovered future earnings per student, per 1°F increase in temperature during a given school year. In a city like Houston, where the average school day is approximately 80°F and thus 10°F above the point where heat's impact on learning first appears, the present value of school air-conditioning is approximately \$2,120 per year for each student, \$53,000 per year for each 25 student classroom, or \$2.1 million per year for each 1,000 student high school. Put differently, the extent to which school air-conditioning would offset the earnings loss driven by the 5°F increase predicted by climate change models is \$1,060 per student, \$26,500 per classroom, or just over \$1 million per high school.

Although these are rough estimates, benefit values of this order of magnitude imply that school infrastructure may more than justify their costs. Based on available data, we conservatively esti-

³⁰Chetty et al. (2011) compute a \$7,000 net present value in increased earnings for the typical 12 year old student in their sample. We apply their five percent discount rate to generate the \$8,500 figure for the typical 16 year old student taking the PSAT. An important assumption we make by using these estimates is that the achievement gains due to a better teacher result in later life impacts that are equivalent in magnitude to those associated with having a climate-controlled learning environment. For instance, if teachers impart valuable non-cognitive skills but air-conditioning does not, these estimates would be overstated.

mate the amortized cost of school air conditioning to be approximately \$125,000 per school per year, or \$125 per student per year for a 1,000 person school.³¹ Given variation in the average amount of school year heat, we estimate that the cost of improving scores by 1 percent of a standard deviation for the average US school is between \$25 and \$125 per student per year.³² For comparison, the average cost associated with the Tennessee STAR experiment – which reduced class sizes by roughly one-third at a cost of \$3,602 per student per year (\$2,151 in 1995 dollars) and resulted in test score gains of approximately 0.22 standard deviations Krueger (1999) – is approximately \$163 per 1 percent of a standard deviation improvement per student.

5 Heat and Achievement Gaps

5.1 Inequality in School Air-Conditioning

We next document racial and income gaps in school air-conditioning penetration rates. To do so, we regress the student-generated measure of school air-conditioning penetration on an indicator for being black or Hispanic (relative to being white) and separately on an indicator for being in the lower half of the ZIP code income distribution (within cooler and hotter halves of the country to account for temperature differences). We also control for a quartic in school-level mean school year temperature experienced over the sample time period.³³ As panel A of Table 6 shows, black and Hispanic students report 1.7 percentage point (9 percent of a standard deviation) lower rates of school air-conditioning penetration relative to white students, while lower income students report 2.8 percentage point (15 percent of a standard deviation) lower rates than higher income

³¹We note that the average school AC budget allocations from two recent bond initiatives in New York City and Denver public schools were reported as being \$100,000 and \$886,000 per school respectively. We therefore conservatively assume that these allocations only represent up-front fixed costs, and that variable costs associated with operating school air conditioners are an additional 10% of fixed costs each year. Conservatively assuming \$886,000 in up-front fixed costs, a 5% discount rate and an average lifetime of 20 years, we estimate total amortized costs of \$125,000 per school per year. (For Denver, see: <https://www.chalkbeat.org/posts/co/2016/06/09/how-denver-public-schools-could-spend-70-million-cooling-its-hottest-schools/>. For NYC, see: <https://www1.nyc.gov/office-of-the-mayor/news/261-17/mayor-de-blasio-chancellor-fari-a-city-council-every-classroom-will-have-air>.)

³²The benefits of installing school air conditioning will likely be proportional to the average school year heat that is offset. In a place like Houston, which experiences approximately 35 school days above 90 and 48 days in the 80's, our estimates (using coefficients from column 1 of table 2) suggest that air conditioning offsets around 5.7 percent of a standard deviation in learning losses that would occur due to hot temperature. In a cooler place such as New York, which experiences 2 days above 90°F and 14 days in the 80's, we estimate the present value of air conditioning to be approximately 1 percent of a standard deviation.

³³In these regressions, we limit the sample to one observation per student and use all PSAT takers to get a more representative picture of national patterns.

students. These disparities appear somewhat stronger in cooler areas of the country than in hotter areas, defined by dividing schools into those below and above the national median school year temperature.

Panel B of Table 6 also measures gaps in school air-conditioning penetration, instead using as an outcome an indicator having “inadequate” school air-conditioning. We define schools with inadequate air-conditioning as those where the average student responded that classrooms get too hot on hot days “Most of the time” or “All the time”. Black and Hispanic students are 1.6 percentage points more likely than white students to be in schools with inadequate air-conditioning. Lower income students are 6.2 percentage points more likely to be in schools with inadequate air-conditioning than their higher income counterparts. These disparities are again substantially larger in cooler areas of the country than in hotter areas. Figure 7 further shows that racial gaps in school air-conditioning access are not explained purely by ZIP code level income differences.

5.2 Heterogeneous Impacts of Heat

Heat exposure has heterogeneous impacts by race, income, and geography, which we show in Table 7.³⁴ As panel A shows, the impact of prior year heat on black and Hispanic students is three times larger than the impact on white students. The impact of prior year heat on students in lower income ZIP codes is twice as large as on those from higher income ZIP codes.³⁵ The last two columns split the sample into high schools with below and above national median school year temperatures. Even though cooler areas of the country are less likely to experience extreme heat, students in these areas seem to experience more learning disruption per unit of extreme heat, consistent with lower average levels of defensive investments. Each school day above 90°F has more than three times the negative impact on test scores in cooler areas than in hotter areas.³⁶ We also see suggestive evidence that heat is more damaging to achievement in more humid areas of

³⁴We find no evidence of heterogeneity by student gender.

³⁵Measurement error may imply that these racial and income gaps in heat’s effects are, if anything, somewhat understated. The last four columns of Table A.3 suggest that measurement error may generate more underestimation in heat’s effects for Black and Hispanic students and low income students than their White and high income counterparts. If weather sensors were immediately adjacent to high schools, these estimates suggest that heat’s impacts would be about 25 percent larger for minority and low income students but largely unchanged for White and high income students.

³⁶In cooler areas, the observed impact of heat could be driven partly by higher temperatures generating an increased number of otherwise scarce pleasant days for students to skip school. That days above 90°F show negative impacts and are too hot to be considered pleasant likely implies that channel is not the only explanation of heat’s effects.

the country.³⁷

Panel B shows cumulative impacts measured by adding coefficients from the previous four years of heat exposure. Experiencing 1°F hotter school years over the past four years has a nearly 80 percent larger impact on black and Hispanic students than on white students. The impact of one additional day above 90°F in each of the preceding four school years has a 40 percent larger impact on black and Hispanic students than on white students. Interestingly, four-year cumulative impacts of heat exposure do not appear to vary by ZIP code income as much as one-year impacts. The cumulative effects of a 1°F increase in school year temperatures is nearly twice as high in cooler areas, while the cumulative impact of each year having an additional day above 90°F is five times larger in cooler areas.

That the achievement of minority and lower income students suffers more from heat exposure is consistent with minority households and high-minority schools having fewer resources to make defensive investments, either *ex ante* or *ex post* (Graff Zivin and Neidell, 2013; Kahn, 2016). Wealthier students may be able to compensate for lost learning time by getting additional instruction from their parents or private tutors. Such students may also be more likely to attend schools where teachers have better capacity to compensate for lost learning time by adjusting lesson plans or adding more instructional time to the day.³⁸ A simpler explanation for this heterogeneity is that minority and lower income students have less access to school conditioning, as we have shown, and home air-conditioning, both of which help offset the negative impacts of heat, though it appears that minorities and low income students also tend to live in historically hotter climates which appear to feature other effective adaptations.³⁹

The estimates in Table 7 allow us to compute the fraction of the racial achievement gap that is driven by a combination of the heterogeneous impacts of heat by race and racial differences in heat exposure. To do this, we ask how much larger racial achievement gaps are than they would be if all students always experienced school day temperatures that appear optimal for learning (i.e. those in the 60s). Our estimates suggest that total exposure to days above 90°F lowers black

³⁷Though heat indices accounting for humidity are not available from our weather sensor data, Table A.8 estimates heat effects separately by typical humidity and finds that heat appears more damaging in humid areas than arid areas.

³⁸Both research and media reports suggest teachers are aware of the adverse impacts of heat on student performance and make efforts to offset some of those impacts *ex post*. Park (2017) finds that New York City teachers selectively boosted grades of students who experienced hot exam sittings and scored just below pass/fail cutoffs.

³⁹While our residential AC data is not student-specific, we note using RECS microdata that Black and Hispanic households are 7% and 6% less likely respectively to have residential air conditioning for a given climate.

and Hispanic students' achievement by 0.043 standard deviations (-0.00277 standard deviations per additional day above 90°F * 15.7 days above 90°F). For white students, the comparable effect is 0.021 standard deviations (-0.00197 standard deviations per additional day above 90°F * 10.06 days above 90°F). This means that extremely hot days widen the racial achievement gap by approximately 0.022 standard deviations, or about three percent of the 0.8 standard deviation gap in PSAT performance between black and Hispanic students and white students. If we also account for the cumulative impacts of school days between 70 and 90°F, then excess heat accounts for about seven percent of the racial achievement gap.⁴⁰ The physical environment students are exposed to, as measured by school day temperatures, thus accounts for a small but non-trivial portion of the racial achievement gap.

Similarly, our analyses suggest that the disruptive influence of heat may play a causal role in generating cross-sectional differences in standardized achievement. The North-South gradient in PSAT performance is 0.023 standard deviations per °F hotter average climate. This presumably captures many correlated factors, including differences in teacher quality, parental incomes, and rates of air conditioning. Our estimate of the causal (cumulative) impact of a 1°F hotter climate on learning is approximately -0.0056 standard deviations at current average levels of air conditioning (column 4 of Table 4). This suggests that up to 24 percent (-0.0056/0.023) of the North-South gradient in PSAT scores may be attributable to the causal impact of a hotter climate. This most likely overstates heat's contribution to that gradient, given the positive correlation between hotter climate and levels of air conditioning, but suggests that a non-trivial portion of existing cross-sectional differences in learning may be caused by differences in the thermal conditions under which learning occurs.

6 Conclusion

We provide the first evidence that cumulative heat exposure inhibits cognitive skill development and that defensive investments such as school air-conditioning may mitigate this effect. Our findings imply that the physiological impacts of heat directly interfere with learning. This work high-

⁴⁰Though the estimated impacts of school days between 70 and 90°F are smaller and noisier than those from days above 90°F, the cumulative coefficient on such days for black and Hispanic students is -0.0004 standard deviations and they experience nearly 84 such days a year. The estimated impact of such days on white students is zero (out to four decimal places).

lights the understudied role that students' and teachers' physical environments play in generating educational outcomes. It also highlights the extent to which disparities in such physical environments contribute to inequality in educational outcomes such as the racial achievement gap. The results allow us to estimate the value of one potential public investment, school air-conditioning, that may help reduce such gaps.

Evidence that heat exposure affects human capital accumulation points to a potential understudied channel through which heat may affect macroeconomic growth. A variety of recent papers have documented clear connections between country-level temperature fluctuations and growth but the mechanisms explaining that connection have remained a matter of speculation (Dell et al., 2012; Burke et al., 2015). Hypothesized channels include the negative impacts of heat on physical health and on labor productivity, particularly for physically intensive occupations (Hsiang, 2010; Heal and Park, 2016). Our evidence suggests that heat not only interferes with the physical capabilities of a nation's workforce but also with its cognitive capacities, and most crucially the rate at which valuable skills are accrued by the workforce over time.

Understanding the causal relationship between cumulative heat exposure and learning is of heightened policy relevance given accelerating warming in most parts of the world, and given that the overwhelming majority of the world's population does not yet have access to air-conditioning (Davis and Gertler, 2015). Based on current estimates of projected warming in the U.S., we engage in the following thought experiment: By 2050, how much heat-related learning disruption can we expect for the average high school student, relative to a student attending high school in 2010? Median climate change scenarios for the contiguous United States predict average warming of roughly 5°F by 2050. To generate a summary measure of the impact of climate change on future learning, we take the average treatment effect in degree F terms from above (0.006 standard deviations) and multiply by the extent of predicted mid-century warming (5°F), yielding an estimate of 0.03 standard deviation lower achievement, or approximately 10 percent of the average learning achieved in a given school year. Without further investments in school infrastructure, climate change would likely also result in a further widening of current racial achievement gaps.⁴¹

⁴¹These calculations do not account for potential non-linearities in the damage function for temperatures outside the range of historical experience, such as days above 110°F. It is important to note that these back-of-the-envelope estimates do not account for heterogeneity in warming or air-conditioning penetration patterns across regions, or for the costs of installing air-conditioning, both of which are likely to be important determinants of net welfare impacts. They also do not account for potential migration responses.

Further questions about the impact of heat on learning remain. The average American student experiences approximately 12 school days above 90°F per year, whereas the average Indian student experiences over 100 such days annually. What portion of the achievement gap between hot and cool countries is explained by the direct impact of heat exposure on learning documented here? Do these impacts on cognitive skill and learning have longer-term impacts on students' economic outcomes, affecting rates of economic mobility and convergence? Other than school air-conditioning, what alternative investments or actions can be taken to mitigate the impacts of heat on learning? While it is possible that individuals in hotter climates are better acclimated and thus experience reduced sensitivity per unit of heat exposure, given much lower levels of air-conditioning and the potential for other correlated health or nutritional impacts, we speculate that our estimates represent a conservative appraisal of the inhibiting influence of a hotter climate on human capital development. We hope future work addresses such questions.

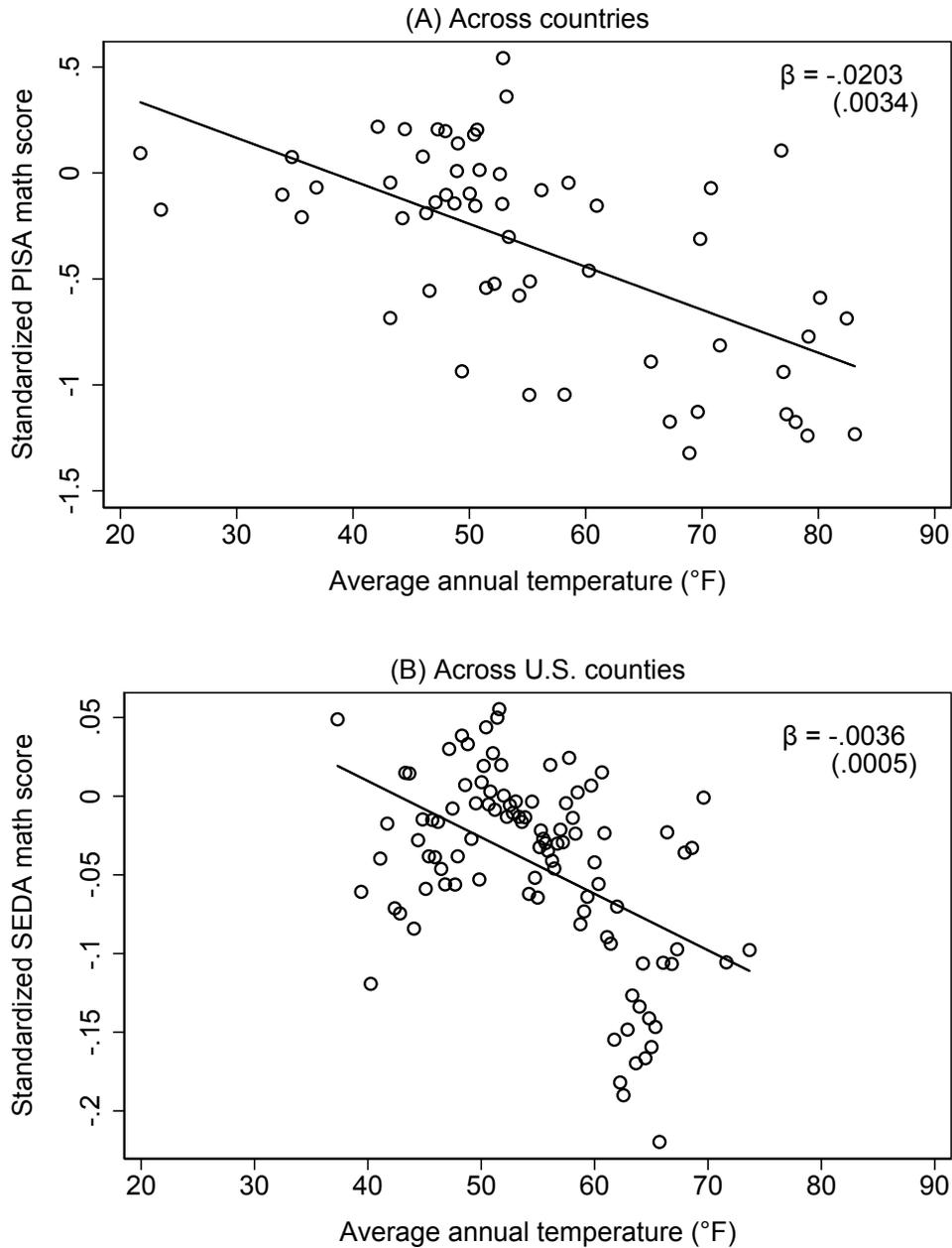
References

- Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge (2017), "When should you adjust standard errors for clustering?" Technical report, National Bureau of Economic Research.
- Acemoglu, Daron, Simon Johnson, and James A Robinson (2001), "The colonial origins of comparative development: An empirical investigation." *The American Economic Review*, 91, 1369–1369.
- Auffhammer, Maximilian and Erin T Mansur (2014), "Measuring climatic impacts on energy consumption: A review of the empirical literature." *Energy Economics*, 46, 522–530.
- Banzhaf, H Spencer and Randall P Walsh (2008), "Do people vote with their feet? an empirical test of tiebout." *American Economic Review*, 98, 843–63.
- Behrer, A Patrick and Jisung Park (2017), "Will we adapt? temperature, labor and adaptation to climate change." *Harvard Kennedy School Working Paper*.
- Bleakley, Hoyt (2010), "Malaria eradication in the americas: A retrospective analysis of childhood exposure." *American Economic Journal: Applied Economics*, 2, 1–45.
- Burke, Marshall, Solomon M Hsiang, and Edward Miguel (2015), "Global non-linear effect of temperature on economic production." *Nature*, 527, 235–239.
- Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein (2010), "The value of school facility investments: Evidence from a dynamic regression discontinuity design." *The Quarterly Journal of Economics*, 125, 215–261.
- Chetty, Raj, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan (2011), "How does your kindergarten classroom affect your earnings? evidence from project star." *The Quarterly Journal of Economics*, 126, 1593–1660.
- Cho, H (2017), "Effect of summer heat on test scores: A cohort analysis." *Journal of Environmental Economics and Management*, 83, 185–196.
- Conley, Timothy G (1999), "Gmm estimation with cross sectional dependence." *Journal of econometrics*, 92, 1–45.
- 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, 682–694.
- Davis, Lucas W and Paul J Gertler (2015), "Contribution of air conditioning adoption to future energy use under global warming." *Proceedings of the National Academy of Sciences*, 112, 5962–5967.
- Dee, Thomas S, Will Dobbie, Brian A Jacob, and Jonah Rockoff (2016), "The causes and consequences of test score manipulation: Evidence from the new york regents examinations." *NBER Working Paper*.
- Dell, Melissa, Benjamin F Jones, and Benjamin A Olken (2009), "Temperature and income: reconciling new cross-sectional and panel estimates." *American Economic Review*, 99, 198–204.
- Dell, Melissa, Benjamin F Jones, and Benjamin A Olken (2012), "Temperature shocks and economic growth: Evidence from the last half century." *American Economic Journal: Macroeconomics*, 4, 66–95.
- Deryugina, Tatyana and Solomon M Hsiang (2014), "Does the environment still matter? daily temperature and income in the united states." *NBER Working Paper*.
- Deschenes, Olivier, Michael Greenstone, and Joseph S Shapiro (2017), "Defensive investments and the demand for air quality: Evidence from the nox budget program." *American Economic Review*, 107, 2958–2989.

- Duncan, Greg J and Richard J Murnane (2011), *Whither opportunity?: Rising inequality, schools, and children's life chances*. Russell Sage Foundation.
- Durán-Narucki, Valkiria (2008), "School building condition, school attendance, and academic achievement in new york city public schools: A mediation model." *Journal of environmental psychology*, 28, 278–286.
- 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, 36–65.
- Fahle, Erin M, Benjamin R Shear, Demetra Kalogrides, Sean F Reardon, Richard DiSalvo, and Andrew D Ho (2017), "Stanford education data archive."
- Fryer Jr, Roland G (2011), "The importance of segregation, discrimination, peer dynamics, and identity in explaining trends in the racial achievement gap." In *Handbook of Social Economics*, volume 1, 1165–1191, Elsevier.
- Garg, Teevrat, Maulik Jagnani, and Vis Taraz (2017), "Human capital costs of climate change: Evidence from test scores in india."
- Goodman, Joshua (2014), "Flaking out: Student absences and snow days as disruptions of instructional time." *NBER Working Paper*.
- Graff Zivin, Joshua, Solomon M Hsiang, and Matthew Neidell (2017), "Temperature and human capital in the short-and long-run." *Journal of the Association of Environmental and Resource Economists*.
- Graff Zivin, Joshua and Matthew Neidell (2013), "Environment, health, and human capital." *Journal of Economic Literature*, 51, 689–730.
- Graff Zivin, Joshua and Matthew Neidell (2014), "Temperature and the allocation of time: Implications for climate change." *Journal of Labor Economics*, 32, 1–26.
- Heal, Geoffrey and Jisung Park (2016), "Temperature stress and the direct impact of climate change: A review of an emerging literature." *Review of Environmental Economics and Policy*, 10, 347–362.
- Hsiang, Solomon M (2010), "Temperatures and cyclones strongly associated with economic production in the caribbean and central america." *Proceedings of the National Academy of sciences*, 107, 15367–15372.
- Hsiang, Solomon M, Marshall Burke, and Edward Miguel (2013), "Quantifying the influence of climate on human conflict." *Science*, 341, 1235367.
- Isen, Adam, Maya Rossin-Slater, and Reed Walker (2017), "Relationship between season of birth, temperature exposure, and later life wellbeing." *Proceedings of the National Academy of Sciences*, 201702436.
- Jackson, C Kirabo, Rucker C Johnson, and Claudia Persico (2015), "The effects of school spending on educational and economic outcomes: Evidence from school finance reforms." *The Quarterly Journal of Economics*, 131, 157–218.
- Jacob, Brian A, Lars Lefgren, and David P Sims (2010), "The persistence of teacher-induced learning." *Journal of Human resources*, 45, 915–943.
- Jencks, Christopher and Meredith Phillips (2011), *The Black-White test score gap*. Brookings Institution Press.
- Kahn, Matthew E (2016), "The climate change adaptation literature." *Review of Environmental Economics and Policy*, 10, 166–178.
- Kane, Thomas J and Douglas O Staiger (2008), "Estimating teacher impacts on student achievement: An experimental evaluation." Technical report, National Bureau of Economic Research.

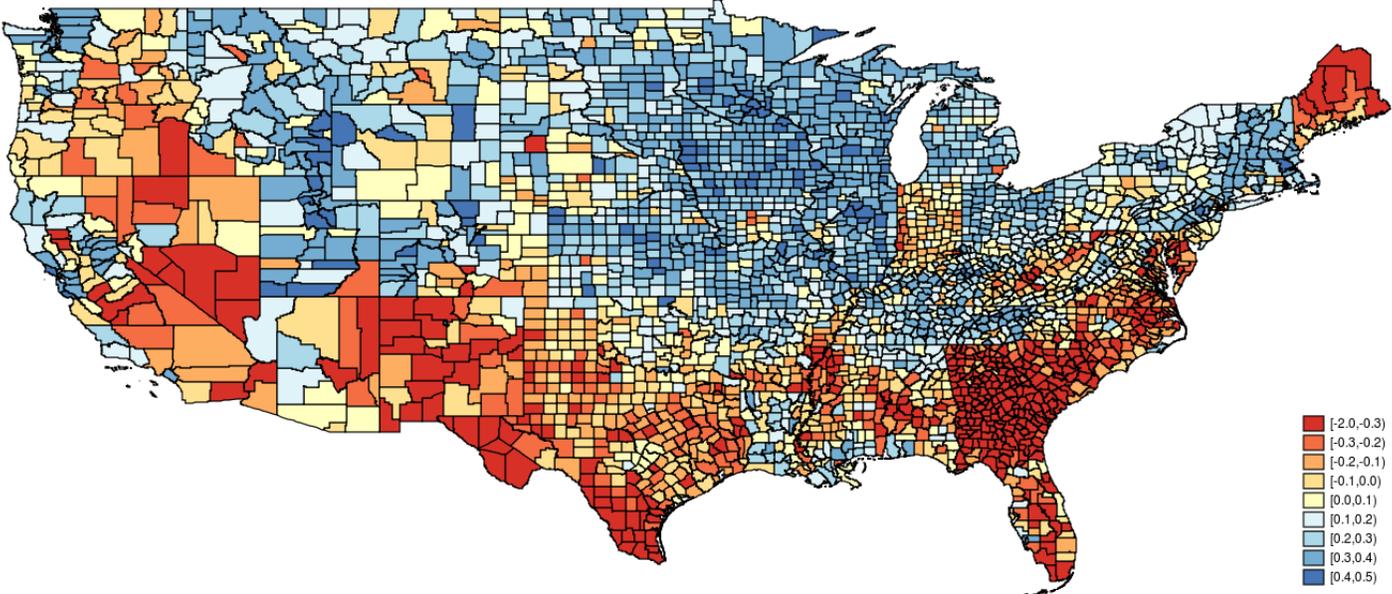
- Krueger, Alan B (1999), "Experimental estimates of education production functions." *The quarterly journal of economics*, 114, 497–532.
- Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach (2018), "School finance reform and the distribution of student achievement." *American Economic Journal: Applied Economics*, 10, 1–26.
- Lafortune, Julien and David Schonholzer (2017), "Do school facilities matter? measuring the effects of capital expenditures on student and neighborhood outcomes." *Working Paper*.
- Mackworth, Norman H (1946), "Effects of heat on wireless operators." *British journal of industrial medicine*, 3, 143.
- Martorell, Paco, Kevin Stange, and Isaac McFarlin (2016), "Investing in schools: capital spending, facility conditions, and student achievement." *Journal of Public Economics*, 140, 13–29.
- Neilson, Christopher A and Seth D Zimmerman (2014), "The effect of school construction on test scores, school enrollment, and home prices." *Journal of Public Economics*, 120, 18–31.
- Park, Jisung (2017), "Hot temperature and high stakes exams: Evidence from new york city public schools." *Harvard University Working Paper*.
- Reardon, Sean F, Demetra Kalogrides, and Ken Shores (2017), "The geography of racial/ethnic test score gaps. cepa working paper no. 16-10." *Stanford Center for Education Policy Analysis*.
- Roback, Jennifer (1982), "Wages, rents, and the quality of life." *Journal of political Economy*, 90, 1257–1278.
- Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi (2004), "Institutions rule: the primacy of institutions over geography and integration in economic development." *Journal of economic growth*, 9, 131–165.
- Rothstein, Jesse (2010), "Teacher quality in educational production: Tracking, decay, and student achievement." *The Quarterly Journal of Economics*, 125, 175–214.
- Seppanen, Olli, William J Fisk, and QH Lei (2006), "Effect of temperature on task performance in office environment." *Lawrence Berkeley National Laboratory*.
- Shah, Manisha and Bryce Millett Steinberg (2017), "Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital." *Journal of Political Economy*, 125, 527–561.
- Stafford, Tess M (2015), "Indoor air quality and academic performance." *Journal of Environmental Economics and Management*, 70, 34–50.
- Stocker, Thomas (2014), *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
- Tiebout, Charles M (1956), "A pure theory of local expenditures." *Journal of political economy*, 64, 416–424.

Figure 1: Climate and Achievement Across Geographies



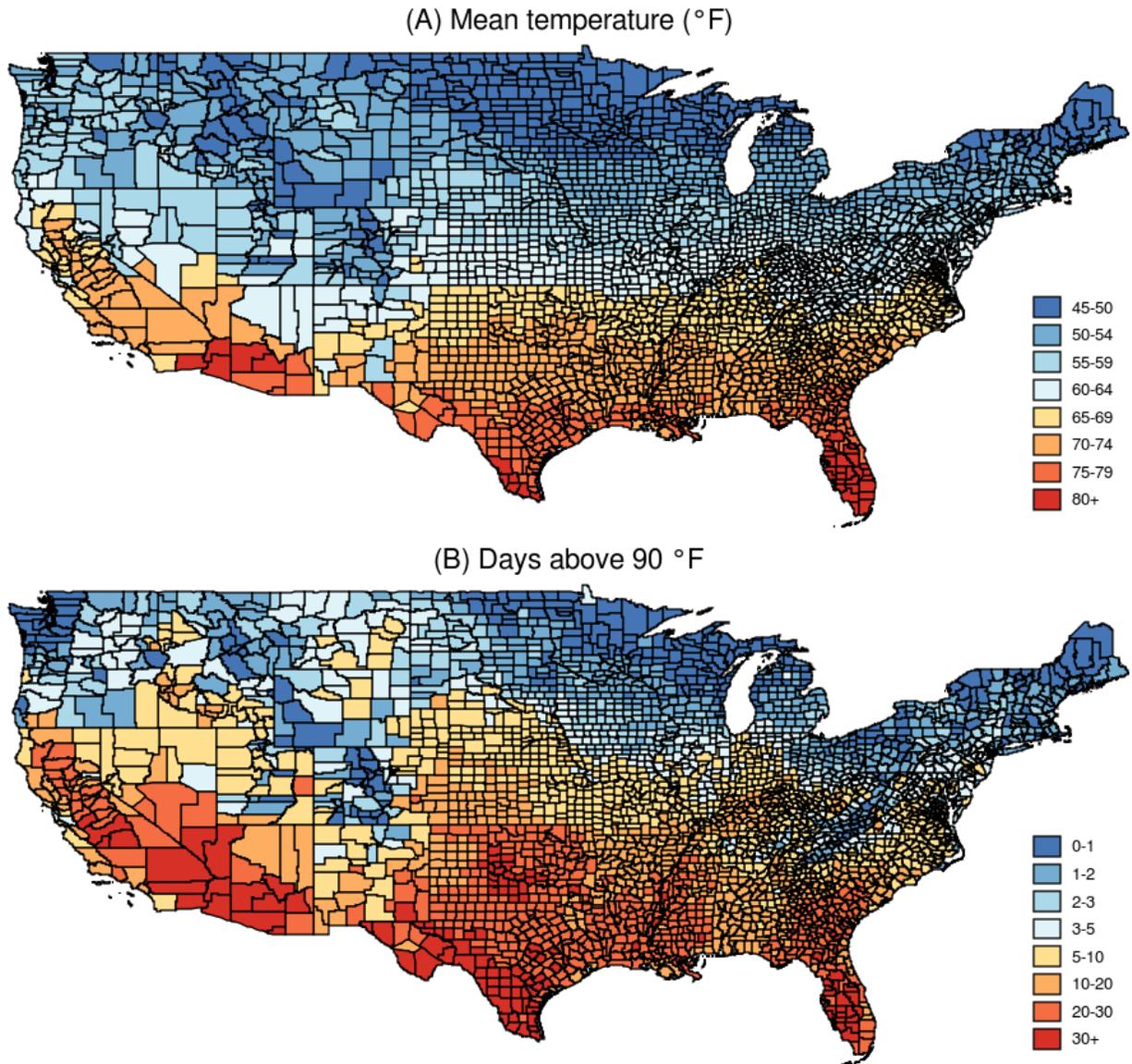
Notes: The above figure shows a scatterplot of mean 2012 PISA (panel A) or SEDA (panel B) math scores and average annual temperature by country or U.S. county. Average annual temperatures are measured over the period 1980-2011. Panel B shows a binned percentile plot of standardized 3rd-8th grade math scores (2009-2013) by percentile of the county-level average temperature distribution, with scores standardized by subject, grade and year as in Fahle et al. (2017). Also shown is a fitted line and slope coefficient from a bivariate regression of scores on temperatures, using heteroskedasticity robust standard errors.

Figure 2: Spatial Distribution of PSAT Z-Scores



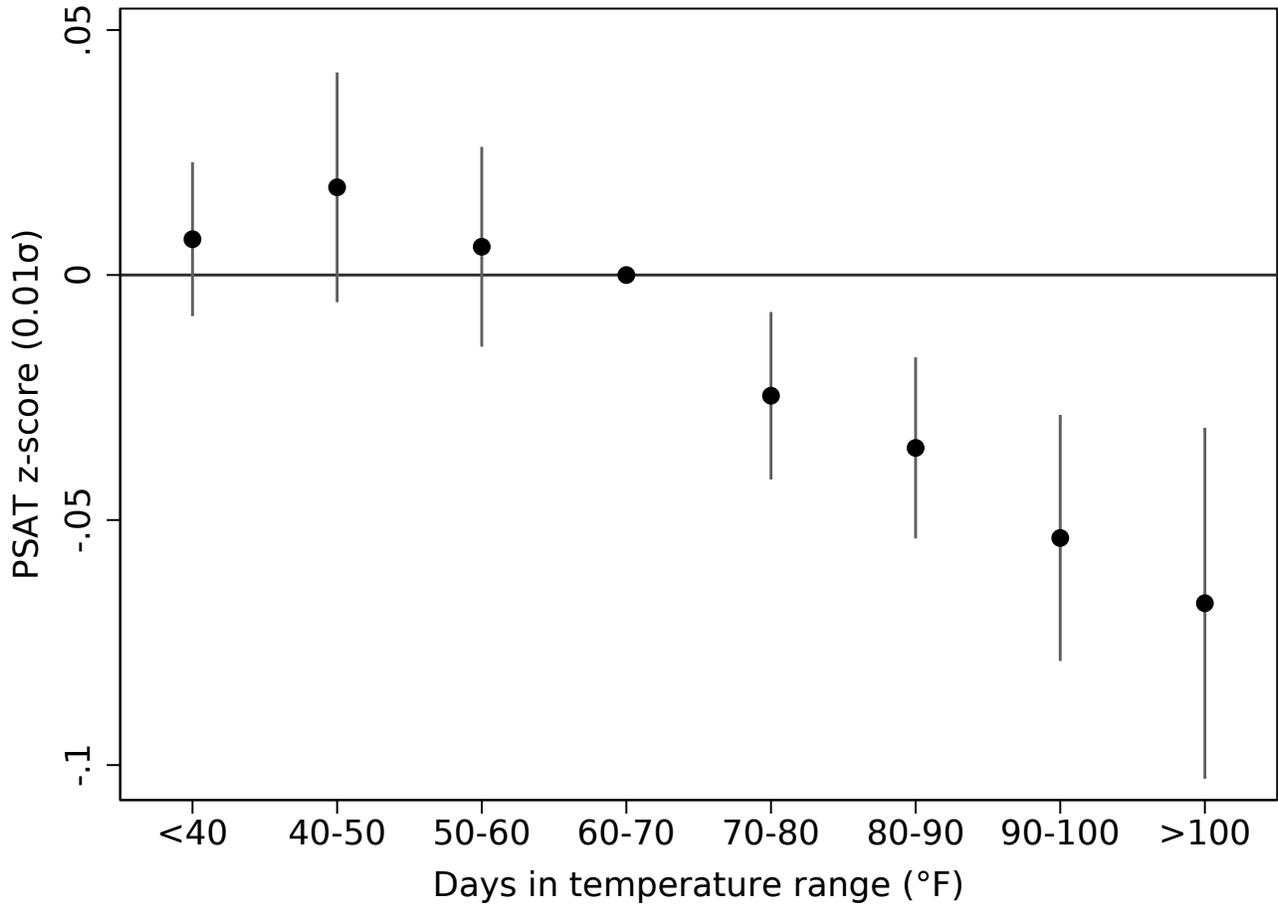
Notes: The above figure shows county-level average standardized PSAT scores from the high school classes of 2001-14.

Figure 3: Spatial Variation in Prior Year Temperature



Notes: The above figure shows the mean daily maximum temperature (panel A) and number of days above 90°F (panel B) experienced by students on school days in the 365 days prior to taking the PSAT, by county. The sample consists of all PSAT-takers from the high school classes of 2001-14, whose PSATs were taken between 1997 and 2012.

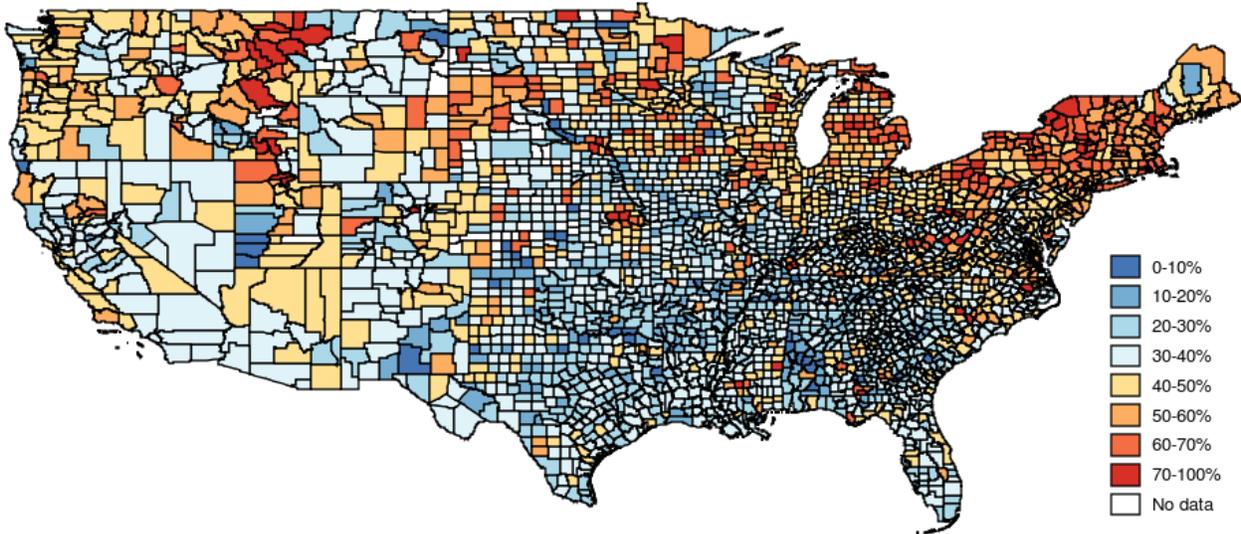
Figure 4: Cumulative Hot Days and Test Performance



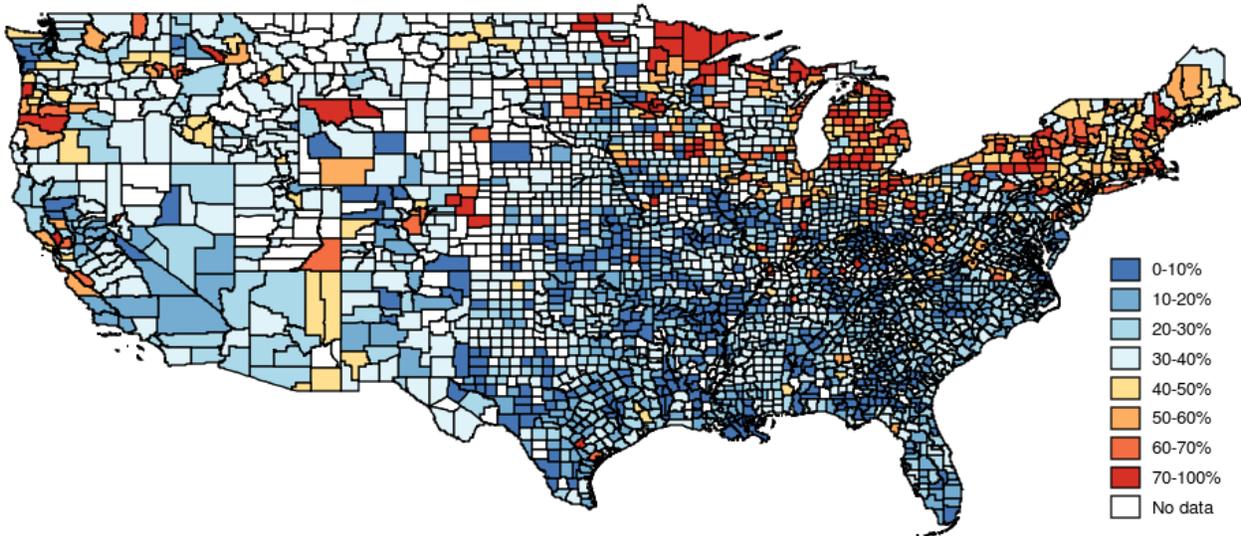
Notes: Shown above are coefficients from a regression of hundredths of a standard deviation in PSAT total (math plus reading) scores on the number of school days within a given temperature range during the 365 days preceding the PSAT take. The regression includes student fixed effects and fixed effects for each combination of cohort, test date and take number. Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). The sample comprises all students from the high school classes of 2001-14 who took the PSAT more than once.

Figure 5: Hot Classrooms

(A) Fraction of hot days when classrooms get too hot (students)

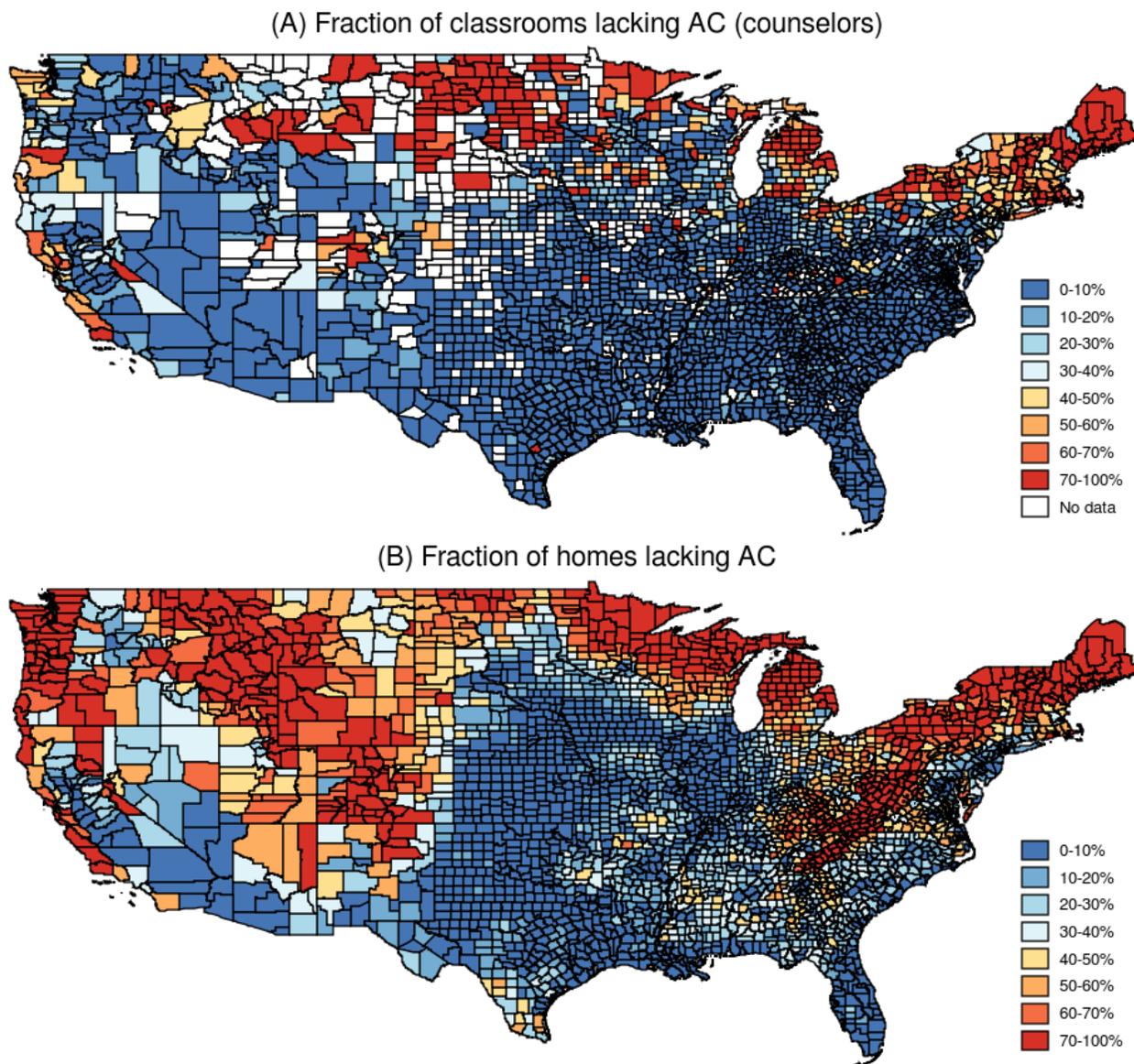


(B) Fraction of hot days when classrooms get too hot (counselors)



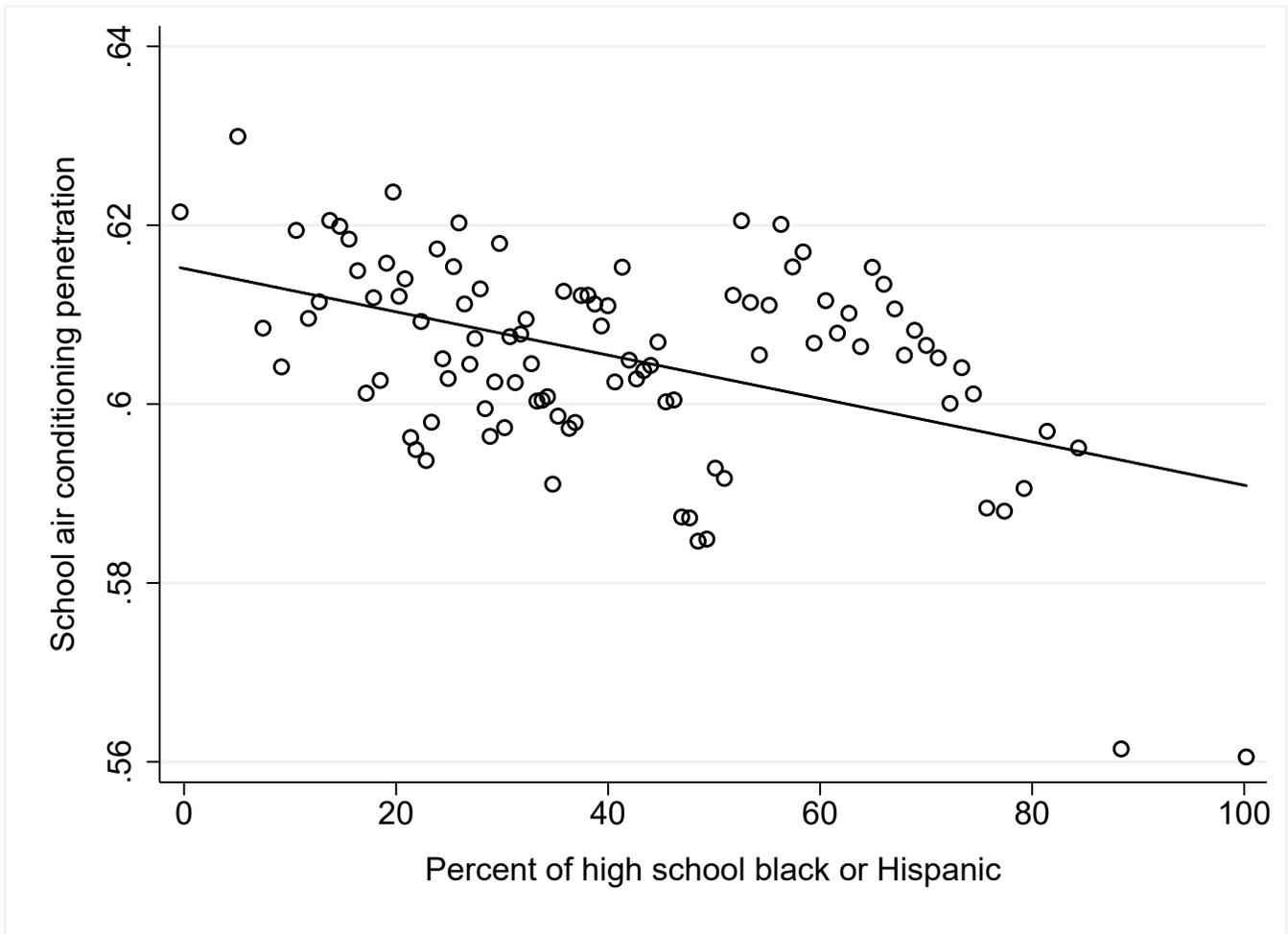
Notes: The above figure shows by county the mean fraction of classrooms reported as “too hot” on hot days by students (panel A) and guidance counselors (panel B). Both measures are derived from student or counselor responses to a survey administered by the College Board. The sample consists of all PSAT-takers from the high school classes of 2001-14, whose PSATs were taken between 1997 and 2012.

Figure 6: School and Home Air-conditioning



Notes: The above figure shows by county the mean fraction of classrooms (panel A) and homes (panel B) lacking air-conditioning. Classroom measures are derived from guidance counselor responses to a survey administered by the College Board. Home measures are derived from the 1980 Census and the 1993-2015 quadrennial Residential Energy Consumption Surveys. The sample consists of all PSAT-takers from the high school classes of 2001-14, whose PSATs were taken between 1997 and 2012.

Figure 7: School Air-conditioning by Percent Black or Hispanic



Notes: The above figure is a binned percentile plot of high school air-conditioning penetration rates as implied by student survey responses, by percentile of the school-level percent black or Hispanic distribution. It plots residual variation after controlling for average daily maximum temperature by school and average income by zip code between 1997 and 2012.

Table 1: Summary Statistics

	PSAT Retakers					
	All takers (1)	All retakers (2)	Black or Hispanic (3)	White (4)	Lower income (5)	Higher income (6)
(A) Demographics						
Female	0.53	0.55	0.56	0.54	0.56	0.53
White	0.58	0.58	0.00	1.00	0.45	0.69
Black or Hispanic	0.29	0.28	1.00	0.00	0.43	0.16
Mother has B.A.	0.22	0.33	0.18	0.40	0.20	0.44
ZIP code mean income	63.2	69.6	49.4	78.3	38.9	95.6
(B) PSAT scores						
Retook PSAT	0.36	1.00	1.00	1.00	1.00	1.00
Total takes	1.42	2.15	2.17	2.14	2.14	2.15
First PSAT z-score	-0.00	0.14	-0.49	0.40	-0.15	0.40
(C) Temperature						
Mean temperature (°F)	65.1	65.8	68.8	64.2	65.9	65.6
Days above 90 °F	11.9	12.2	15.7	10.6	12.7	11.7
(D) Air conditioning						
Classrooms with AC	0.58	0.59	0.60	0.58	0.58	0.60
Homes with AC	0.77	0.80	0.85	0.79	0.77	0.82
N (scores)	38,303,474	21,076,009	6,023,145	12,161,058	9,570,444	11,322,404
N (students)	27,023,119	9,795,654	2,775,607	5,689,371	4,462,169	5,259,910

Notes: Notes: Mean values of key variables are shown. Column 1 includes comprises all students from the high school classes of 2001-14 who took the PSAT at least once. Column 2 includes only those who took the PSAT more than once. Columns 3-6 include subgroups of retakers, with columns 5 and 6 respectively including below and above median ZIP code-level income within below and above median temperature areas.

Table 2: Prior Year Temperature and PSAT Scores

	(1)	(2)	(3)	(4)	(5)	(6)
(A) Average heat						
Mean temperature (°F)	-0.181*** (0.028)	-0.211*** (0.036)	-0.185*** (0.029)	-0.182*** (0.028)	-0.176*** (0.027)	-0.230*** (0.042)
(B) Hot days						
Days above 100 °F	-0.067*** (0.018)	-0.077*** (0.020)	-0.075*** (0.017)	-0.065*** (0.018)	-0.064*** (0.016)	-0.098*** (0.033)
Days in 90s (°F)	-0.053*** (0.013)	-0.061*** (0.014)	-0.059*** (0.013)	-0.053*** (0.013)	-0.053*** (0.012)	-0.064*** (0.018)
Days in 80s (°F)	-0.035*** (0.009)	-0.037*** (0.010)	-0.039*** (0.009)	-0.035*** (0.009)	-0.029*** (0.009)	-0.046*** (0.013)
Days in 70s (°F)	-0.024*** (0.008)	-0.024*** (0.008)	-0.026*** (0.008)	-0.024*** (0.008)	-0.023*** (0.008)	-0.020* (0.012)
Days below 60 °F	0.010 (0.009)	0.013 (0.009)	0.008 (0.008)	0.010 (0.009)	0.010 (0.007)	0.007 (0.011)
N	21,046,448	21,046,448	21,046,448	21,046,448	21,046,448	5,378,273
Prior year snow, rain	No	Yes	No	No	No	No
Test day weather	No	Yes	No	No	No	No
Pollution	No	No	Yes	No	No	No
Economic conditions	No	No	No	Yes	No	No
State-specific trends	No	No	No	No	Yes	No
Sensor within 5 miles	No	No	No	No	No	Yes

Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). Coefficients in each column and panel come from a regression of hundredths of a standard deviation in PSAT total (math plus reading) scores on the weather measure(s) shown. Temperature is measured with the daily maximum temperature from school days in the 365 days preceding the PSAT take. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. Column 2 adds controls for prior year rainfall and snowfall, as well as test day temperature, rainfall and snowfall. Column 3 controls for prior year and test day pollution levels (carbon monoxide, ozone, sulfur dioxide, nitrogen dioxide and PM10). Column 4 controls for the logarithm of per capita county-level payroll in industries highly exposed to weather. Column 5 adds state-specific linear time trends. Column 6 limits the sample to high schools within 5 miles of the nearest weather sensor. The sample comprises all students from the high school classes of 2001-14 who took the PSAT more than once.

Table 3: Timing of Temperature Shocks

	(1)	(2)	(3)	(4)
<hr/> (A) Mean temperature (°F) <hr/>				
School days, 1 year prior	-0.181*** (0.028)	-0.205*** (0.030)	-0.270*** (0.043)	
Summer days, 1 year prior		0.039 (0.026)		0.047* (0.026)
Weekend days, 1 year prior			0.114*** (0.038)	
School days, post-summer				-0.061** (0.025)
School days, pre-summer				-0.160*** (0.029)
<hr/> (B) Days above 90 °F <hr/>				
School days, 1 year prior	-0.056*** (0.012)	-0.061*** (0.011)	-0.073*** (0.016)	
Summer days, 1 year prior		0.016 (0.011)		0.018 (0.011)
Weekend days, 1 year prior			0.043 (0.028)	
School days, post-summer				-0.074*** (0.019)
School days, pre-summer				-0.074*** (0.016)
N	21,046,448	21,046,448	21,046,448	21,046,448

Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). Coefficients in each column and panel come from a regression of hundredths of a standard deviation in PSAT total (math plus reading) scores on the weather measure(s) shown. School day temperature is measured with the daily maximum temperature from school days in the listed 365 day period relative to the PSAT take. Summer temperature is measured across all days in the summer break preceding the PSAT take. Weekend temperature is measured across all weekends and national holidays in the 365 days preceding the PSAT take. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. The sample comprises all students from the high school classes of 2001-14 who took the PSAT more than once.

Table 4: Lagged and Cumulative Impacts of Heat Exposure

	(1)	(2)	(3)	(4)	(5)
<u>(A) Mean temperature (°F)</u>					
1 year prior	-0.181*** (0.028)	-0.189*** (0.036)	-0.185*** (0.041)	-0.213*** (0.044)	-0.215*** (0.048)
2 years prior		-0.017 (0.036)	-0.010 (0.052)	-0.096 (0.061)	-0.099 (0.071)
3 years prior			0.012 (0.043)	-0.088 (0.057)	-0.093 (0.080)
4 years prior				-0.168*** (0.049)	-0.173** (0.076)
5 years prior					-0.009 (0.058)
Cumulative impact	-0.181*** (0.028)	-0.206*** (0.066)	-0.182 (0.121)	-0.565*** (0.178)	-0.589** (0.296)
<u>(B) Days above 90 °F</u>					
1 year prior	-0.056*** (0.012)	-0.066*** (0.013)	-0.078*** (0.014)	-0.083*** (0.015)	-0.080*** (0.016)
2 years prior		-0.031*** (0.010)	-0.051*** (0.013)	-0.069*** (0.018)	-0.061*** (0.021)
3 years prior			-0.048*** (0.014)	-0.072*** (0.021)	-0.054* (0.028)
4 years prior				-0.043*** (0.017)	-0.024 (0.025)
5 years prior					0.031 (0.019)
Cumulative impact	-0.056*** (0.012)	-0.097*** (0.021)	-0.178*** (0.034)	-0.268*** (0.060)	-0.188* (0.096)
N	21,046,448	21,046,448	21,046,448	21,046,448	21,046,448

Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* p<.10 ** p<.05 *** p<.01). Each coefficient comes from a separate regression of hundredths of a standard deviation in PSAT total (math plus reading) scores on the weather measure(s) shown. Panel A measures mean temperature using the daily maximum temperature from school days in the 365 day period starting the given number of years prior to the PSAT take. Panel B measures the number of such school days above 90 °F and controls for the number of days in other temperature ranges, so that days in the 60s are the reference category. Beneath each column is the cumulative impact of heat exposure, generated by adding the listed coefficients. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. The sample comprises all students from the high school classes of 2001-14 who took the PSAT more than once.

Table 5: Adaptation through School Air Conditioning

	(1)	(2)	(3)	(4)
Mean temp.	-0.322*** (0.067)	-0.569*** (0.104)	-0.229*** (0.060)	-0.456*** (0.169)
Mean temp. * School AC penetration	0.253** (0.099)	0.235*** (0.070)		
Mean temp. * HS class * School AC penetration change			0.117*** (0.039)	0.116*** (0.038)
Mean temp. * HS class			-0.011 (0.015)	0.003 (0.016)
Mean temp. * School AC penetration change			-0.218 (0.168)	-0.275* (0.153)
Mean temp. * Home AC penetration		0.324*** (0.110)		0.211 (0.185)
N	18,665,967	18,665,967	2,935,907	2,935,907
Interactions with ZIP code income, school racial composition, typical school temperature, and district capital expenditures	N	Y	N	Y

Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* p<.10 ** p<.05 *** p<.01). Coefficients in each column come from a regression of hundredths of a standard deviation in PSAT total (math plus reading) scores on the heat measure and interaction shown. Mean temperature is measured by the daily maximum temperature from school days in the 365 days preceding the PSAT take. Columns 1 and 2 interact heat with school air conditioning penetration rates as reported by students in 2016. Columns 3 and 4 interact heat with high school class, the change in school air conditioning penetration rates between 2006 and 2016 as reported by guidance counselors, and the interaction of those two variables. Columns 2 and 4 also control for interactions between heat and county-level home air conditioning penetration rates, as well as (not shown) sensor-level mean school year temperature, ZIP code-level income, the school-level fraction of PSAT-takers who are black or Hispanic, and the school district's average per-pupil capital expenditures from 1996-2000. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. The sample comprises all students from the high school classes of 2001-14 who took the PSAT more than once and whose school air conditioning penetration rate (columns 1 and 2) or change in that rate (columns 3 and 4) are non-missing.

Table 6: Heterogeneity in School Air Conditioning Access

	All schools (1)	Cooler areas (2)	Hotter areas (3)
<hr/> (A) School AC <hr/>			
Black or Hispanic	-0.017*** (0.004)	-0.020*** (0.006)	-0.016*** (0.005)
Lower income	-0.028*** (0.005)	-0.034*** (0.008)	-0.020*** (0.005)
<hr/> (B) Inadequate school AC <hr/>			
Black or Hispanic	0.016*** (0.006)	0.022* (0.012)	0.015** (0.006)
Lower income	0.062*** (0.008)	0.082*** (0.016)	0.039*** (0.006)
N	22,347,878	11,176,342	11,171,536

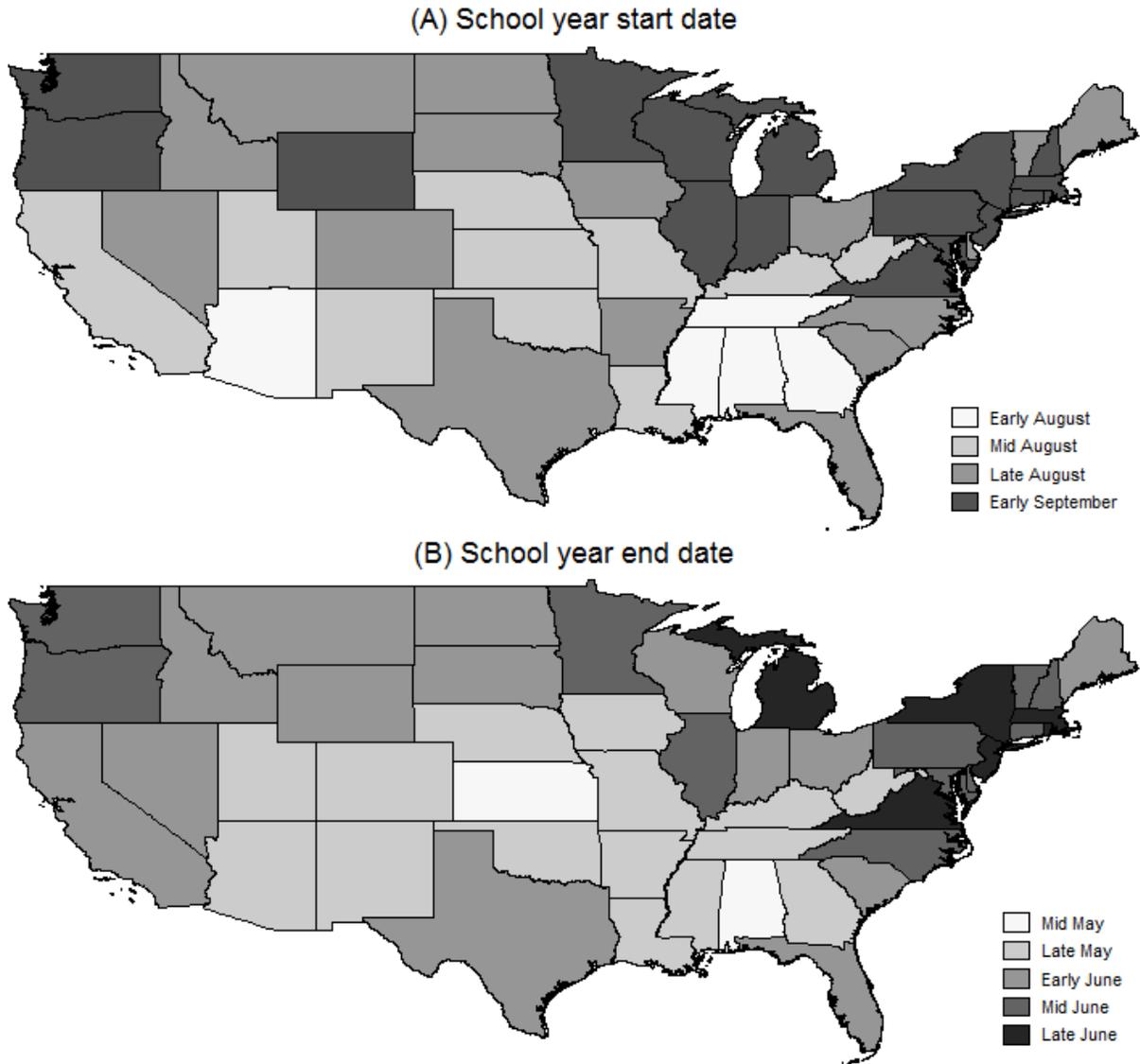
Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). Each coefficient comes from a separate regression of school air conditioning penetration rate (panel A) or an indicator for inadequate air conditioning (panel B) on the listed subgroup indicators. Inadequate air conditioning is defined as the average student's response that, on hot days, classrooms are too hot to learn "most of the time" or "all of the time." The reference groups in each panel are white students (top row) and higher income ZIP codes (bottom row). Lower income refers to students living in below median income ZIP codes within hotter and cooler areas. Each regression controls for a quartic in school-level mean temperatures over the entire time period. Cooler and hotter areas identify schools whose long-term mean temperatures are below or above the median. The sample comprises one observation from each PSAT-taker from the high school classes of 2001-14.

Table 7: Heterogeneity by Race, Income and Geography

	Black or Hispanic (1)	White (2)	Lower income (3)	Higher income (4)	Cooler areas (5)	Hottest areas (6)
<hr/> (A) 1 year prior <hr/>						
Mean temperature (°F)	-0.320*** (0.043)	-0.093*** (0.019)	-0.217*** (0.034)	-0.115*** (0.022)	-0.215*** (0.047)	-0.170*** (0.047)
Days above 90 °F	-0.072*** (0.015)	-0.027*** (0.008)	-0.067*** (0.013)	-0.030*** (0.010)	-0.082*** (0.029)	-0.025* (0.013)
<hr/> (B) Cumulative impact <hr/>						
Mean temperature (°F)	-0.516*** (0.217)	-0.292* (0.151)	-0.348** (0.159)	-0.446** (0.180)	-1.312*** (0.358)	-0.712*** (0.236)
Days above 90 °F	-0.277*** (0.070)	-0.197*** (0.054)	-0.222*** (0.050)	-0.241*** (0.063)	-0.621*** (0.168)	-0.122* (0.067)
<hr/> (C) Average heat <hr/>						
Mean temperature (°F)	68.8	64.2	65.8	65.8	58.1	73.5
Days above 90 °F	15.7	10.6	12.5	11.8	3.9	20.5
N	6,023,145	12,161,058	10,658,547	10,234,301	10,535,013	10,511,435

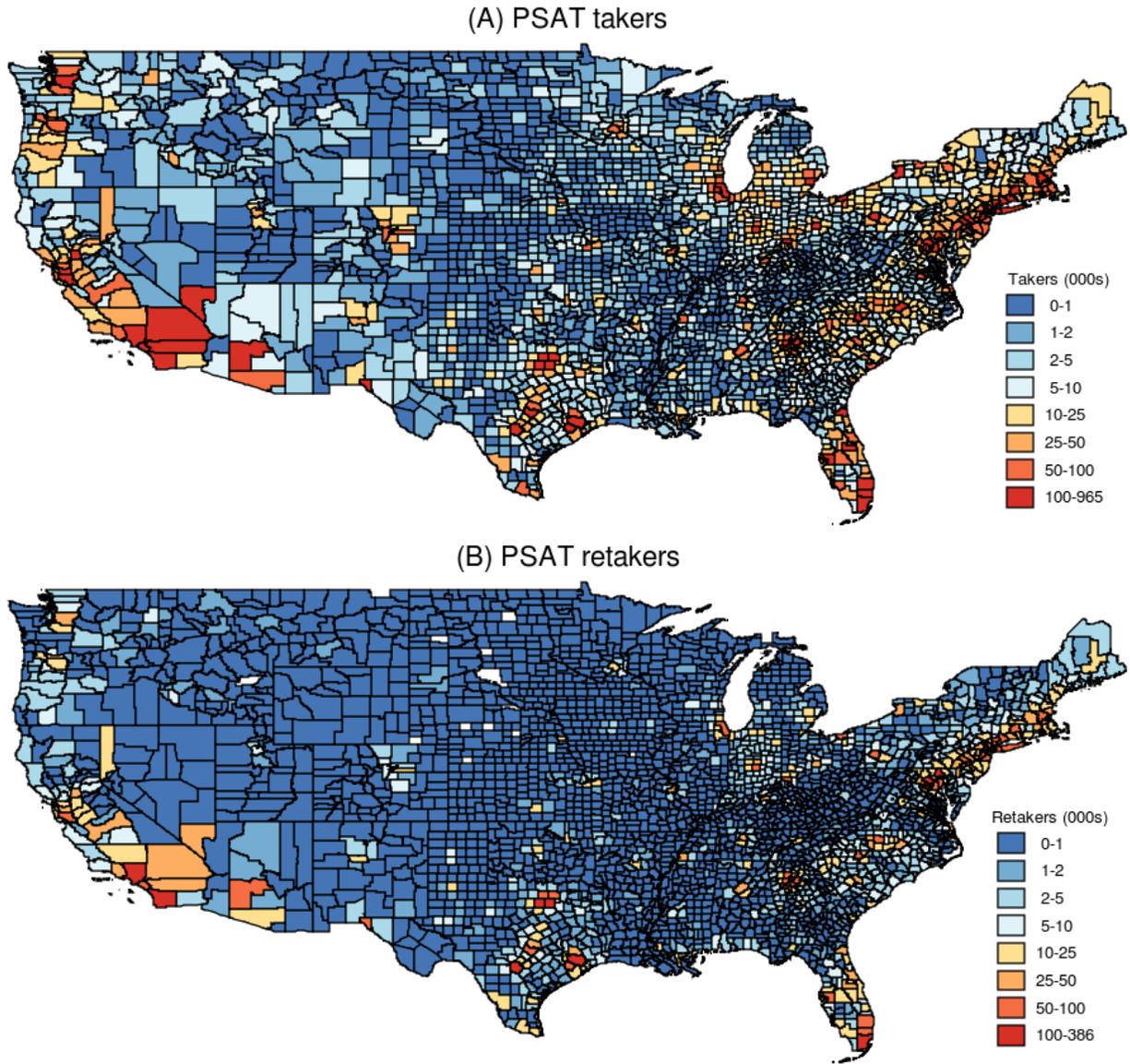
Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* p<.10 ** p<.05 *** p<.01). Each coefficient comes from a separate regression of hundredths of a standard deviation in PSAT total (math plus reading) scores on the weather measure(s) shown. The first row in panel A measures mean temperature using the daily maximum temperature from school days in the 365 days preceding the PSAT take. The second row in panel A measures the number of such school days above 90 °F and controls for the number of days in other temperature ranges, so that days in the 60s are the reference category. The first and second rows in panel B measure the cumulative impact of heat by summing lagged coefficients from each of the four years preceding the PSAT take. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. The sample comprises all students from the high school classes of 2001-14 who took the PSAT more than once. Columns 3 and 4 contain students living in below and above median ZIP code-level income within below and above median temperature areas.

Figure A.1: School Year Calendars by State



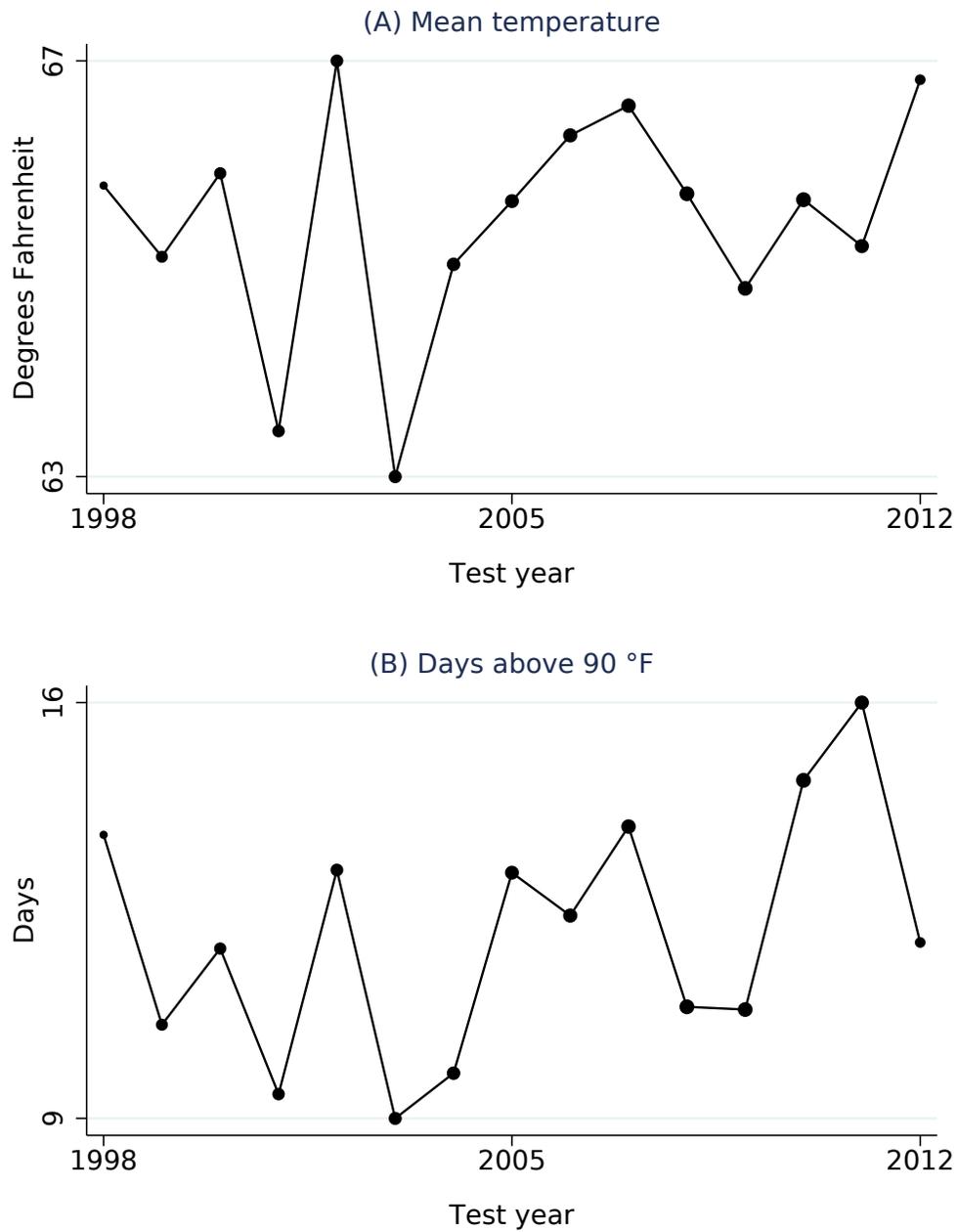
Notes: The above figure shows state's approximate school year start and end dates based on the largest school district in each state and as of 2016.

Figure A.2: Spatial Distribution of PSAT Takers



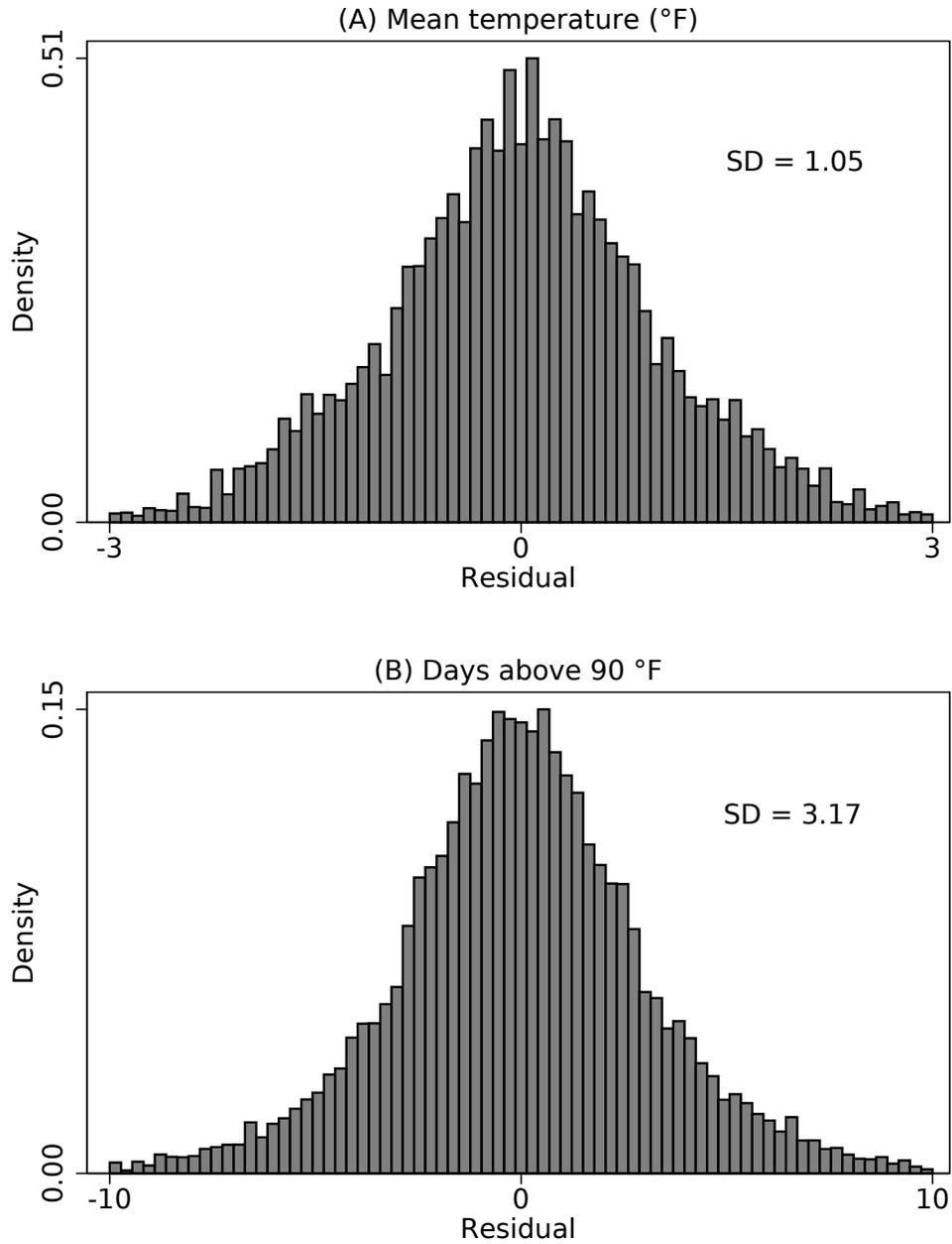
Notes: The above figure shows by county the total number of PSAT takers (panel A) and retakers (panel B) from the high school classes of 2001-14.

Figure A.3: Temporal Variation in Prior Year Temperature



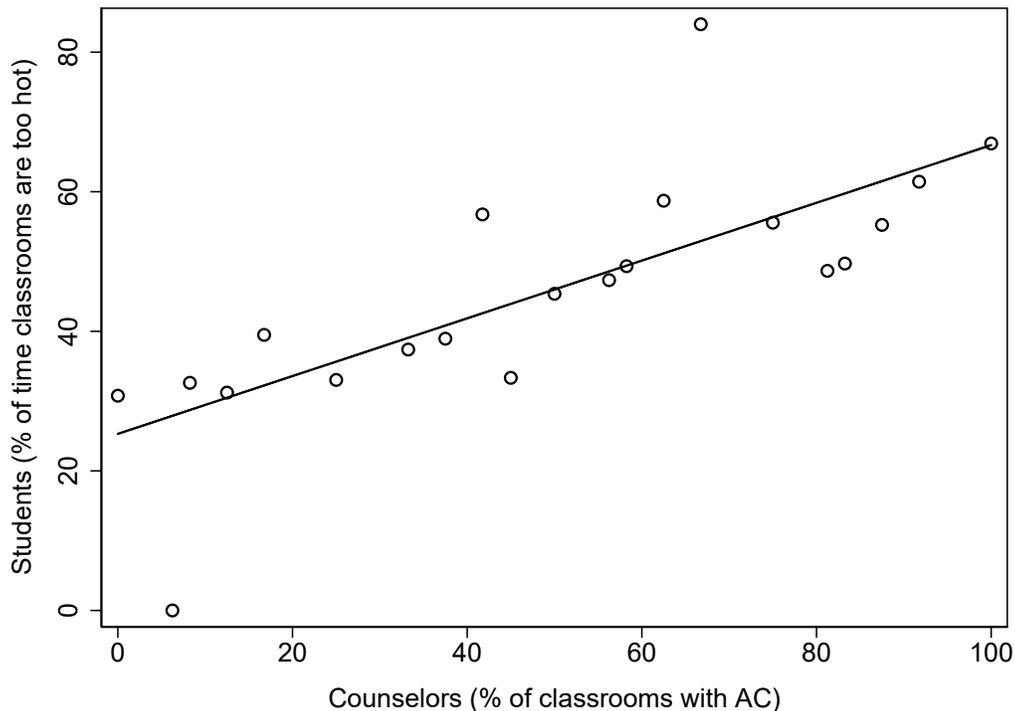
Notes: The above figure shows the mean daily maximum temperature (panel A) and number of days above 90°F (panel B) experienced by students on school days in the 365 days prior to taking the PSAT, by test year. The sample consists of all PSAT-takers from the high school classes of 2001-14, whose PSATs were taken between 1997 and 2012. Dot size is proportional to the number of students in each test year. Test year 1997 is excluded due to the small number of observations.

Figure A.4: Residuals of Prior Year Temperature



Notes: The above figure shows the distribution of residuals resulting from regressions on student fixed effects of the mean daily maximum temperature (panel A) and number of days above 90°F (panel B) experienced by students on school days in the 365 days prior to taking the PSAT. All regressions include fixed effects for each combination of cohort, test date and take number. The figure excludes residuals with magnitude above three (panel A) and 10 (panel B). The standard deviation of the full set of residuals is shown in each panel.

Figure A.5: Students and Counselors Reports of School Air Conditioning



Notes: The above figure shows a bin scatter comparing two measures of school air conditioning penetration rates. The vertical axis shows our preferred measure, constructed from student reports of the fraction of hot days on which classrooms are too hot for learning. The horizontal axis shows counselor reports of the fraction of classrooms with air conditioning. The sample includes all PSAT retakers for whom these two measures are non-missing. The correlation between the measures in this sample is 0.67.

Table A.1: Temperature and PSAT-Taking

	Takers (1)	Ln(takers) (2)	Female (3)	Black or Hispanic (4)	Mother has B.A. (5)	Father has B.A. (6)
<hr/> (A) Baseline <hr/>						
Mean temperature (°F)	-0.0107 (0.0902)	-0.0021 (0.0017)	-0.0001 (0.0001)	0.0007*** (0.0002)	0.0003 (0.0003)	0.0003 (0.0003)
<hr/> (B) State trends <hr/>						
Mean temperature (°F)	-0.0924 (0.0812)	-0.0013 (0.0016)	0.0001 (0.0001)	0.0001 (0.0002)	-0.0001 (0.0002)	-0.0001 (0.0002)
N	686,977	686,977	27,021,552	27,021,552	27,021,552	27,021,552

Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* p<.10 ** p<.05 *** p<.01). Each coefficient comes from a separate regression of the listed characteristic of PSAT-takers on the average daily maximum temperature from school days in the 365 days preceding a student's first PSAT take. All regressions include fixed effects for each high school and for each combination of cohort and test date. Panel B also includes state-specific linear time trends. The sample comprises all students from the high school classes of 2001-14 who took the PSAT at least once.

Table A.2: Temperature and Retaking

	(1)	(2)	(3)	(4)	(5)
Prior year temperature (°F)	0.0005 (0.0005)		0.0005 (0.0005)	0.0003 (0.0005)	-0.0001 (0.0004)
Following year temperature (°F)		0.0003 (0.0004)	0.0003 (0.0005)	0.0003 (0.0005)	-0.0004 (0.0004)
N	27,021,551	27,021,551	27,021,551	27,021,551	27,021,551
Test day temperature	No	No	No	Yes	Yes
State-specific time trends	No	No	No	No	Yes

Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). Each coefficient comes from a separate regression of the probability of retaking the PSAT on the weather measure(s) shown. Yearly temperatures are measured with the daily maximum temperature from school days in the 365 days preceding and following a student's first PSAT take. All regressions include fixed effects for each high school, for each combination of gender, race, income and parental education, and for each combination of cohort and test date. Columns 4 and 5 control for temperature on the day of the first PSAT take. Column 5 includes state-specific linear time trends. The sample comprises all students from the high school classes of 2001-14 who took the PSAT at least once.

Table A.3: Measurement Error from Sensor Distance

	All students (1)	Black or Hispanic (2)	White (3)	Low income (4)	High income (5)
Mean temp. (°F)	-0.237*** (0.045)	-0.389*** (0.061)	-0.085*** (0.025)	-0.279*** (0.050)	-0.091*** (0.032)
Sensor distance * Mean temp. (°F)	0.006** (0.003)	0.008* (0.004)	-0.001 (0.002)	0.006** (0.003)	-0.002 (0.002)
N	21,046,448	6,023,145	12,161,058	10,658,547	10,234,301

Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). Each coefficient comes from a separate regression of hundredths of a standard deviation in PSAT total (math plus reading) scores on the weather measure(s) shown. Each regression measures mean temperature using the daily maximum temperature from school days in the 365 days preceding the PSAT take. Also included is the interaction between temperature and distance from the nearest weather sensor (in miles). All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. The sample in column 1 comprises all students from the high school classes of 2001-14 who took the PSAT more than once. Columns 4 and 5 contain students living in below and above median ZIP code-level income within below and above median temperature areas.

Table A.4: Temperature Effects by Test Subject

	Math (1)	Verbal (2)
<hr/>		
(A) Average heat		
Mean temperature (°F)	-0.159*** (0.034)	-0.177*** (0.023)
<hr/>		
(B) Hot days		
Days above 90 °F	-0.042*** (0.014)	-0.062*** (0.010)
N	21,046,448	21,046,448

Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* p<.10 ** p<.05 *** p<.01). Coefficients in each column and panel come from a regression of hundredths of a standard deviation in PSAT math or reading scores on the weather measure shown. Panel A measures temperature with the daily maximum temperature from school days in the 365 days preceding the PSAT take. Panel B measures the number of such school days above 90 °F. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. Panel B also controls for the number of days in other temperature ranges, so that days in the 60s are the reference category. The sample comprises all students from the high school classes of 2001-14 who took the PSAT more than once.

Table A.5: Heterogeneity by Take Number

	(1)	(2)	(3)	(4)
Mean temp. (°F)		-0.152*** (0.034)	-0.200*** (0.027)	-0.269*** (0.069)
Mean temp. * 1st take	-0.152*** (0.034)		0.048** (0.021)	0.116 (0.079)
Mean temp. * 2nd take	-0.200*** (0.027)	-0.048** (0.021)		0.069 (0.065)
Mean temp. * 3rd take	-0.269*** (0.069)	-0.116 (0.079)	-0.069 (0.065)	
N	21,046,448	21,046,448	21,046,448	21,046,448

Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). Coefficients in each column and panel come from a regression of hundredths of a standard deviation in PSAT total (math plus reading) scores on interactions between take number and the average daily maximum temperature from school days in the 365 days preceding the PSAT take. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. The sample comprises all students from the high school classes of 2001-14 who took the PSAT more than once.

Table A.6: Future Temperature Shocks

	(1)	(2)	(3)	(4)	(5)
Mean temp., 1 year prior (°F)	-0.181*** (0.028)	-0.229*** (0.039)	-0.182*** (0.028)	-0.178*** (0.029)	-0.228*** (0.049)
Mean temp., 1 year after (°F)		-0.090* (0.048)			-0.092 (0.080)
Mean temp., 2 years after (°F)			0.053 (0.032)		-0.006 (0.073)
Mean temp., 3 years after (°F)				-0.037 (0.033)	-0.037 (0.051)
N	21,046,448	21,046,448	21,046,448	21,046,448	21,046,448

Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). Coefficients in each column and panel come from a regression of hundredths of a standard deviation in PSAT total (math plus reading) scores on the weather measure(s) shown. School day temperature is measured with the daily maximum temperature from school days in the listed 365 day period relative to the PSAT take. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. The sample comprises all students from the high school classes of 2001-14 who took the PSAT more than once.

Table A.7: Clustering by weather sensor versus state

	(1)	(2)
<hr/>		
(A) Average heat		
Mean temperature (°F)	-0.181*** (0.028)	-0.181*** (0.041)
<hr/>		
(B) Hot days		
Days above 100 °F	-0.067*** (0.018)	-0.067*** (0.014)
Days in 90s (°F)	-0.053*** (0.013)	-0.053*** (0.020)
Days in 80s (°F)	-0.035*** (0.009)	-0.035** (0.014)
Days in 70s (°F)	-0.024*** (0.008)	-0.024* (0.015)
Days below 60 °F	0.010 (0.009)	0.010 (0.014)
N	21,046,448	21,046,448
Clustering by	Weather sensor	State

Notes: Heteroskedasticity robust standard errors clustered by weather sensor (column 1) or state (column 2) are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). Coefficients in each column and panel come from a regression of hundredths of a standard deviation in PSAT total (math plus reading) scores on the weather measure(s) shown. Temperature is measured with the daily maximum temperature from school days in the 365 days preceding the PSAT take. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. The sample comprises all students from the high school classes of 2001-14 who took the PSAT more than once.

Table A.8: Heterogeneity in Prior Year Temperature Impacts by Humidity

	Humid areas (1)	Arid areas (2)
Mean temperature (°F)	-0.186*** (0.038)	-0.111*** (0.028)
Days above 90 °F	-0.068*** (0.014)	-0.035** (0.015)
N	14,492,113	6,435,636

Notes: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses (* $p < .10$ ** $p < .05$ *** $p < .01$). Each coefficient comes from a separate regression of hundredths of a standard deviation in PSAT total (math plus reading) scores on the weather measure(s) shown. The first row measures mean temperature using the daily maximum temperature from school days in the 365 days preceding the PSAT take. The second row measures the number of such school days above 90 °F and controls for the number of days in other temperature ranges, so that days in the 60s are the reference category. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. The sample comprises all students from the high school classes of 2001-14 who took the PSAT more than once. Columns 1 and 2 contain students living in areas with below and above the median of the US summer humidity distribution, as characterized by ZIP code level summer dewpoints.