

Temperature and Human Capital in the Short and Long Run

Joshua Graff Zivin, Solomon M. Hsiang, Matthew Neidell

Abstract: We provide the first estimates of the potential impact of climate change on cognitive performance and attainment, focusing on the impacts from both short-run weather and long-run climate. Exploiting the longitudinal structure of the NLSY79 and random fluctuations in weather across interviews, we identify the effect of temperature in models with child-specific fixed effects. We find that short-run changes in temperature lead to statistically significant decreases in cognitive performance on math (but not reading) beyond 26°C (78.8°F). In contrast, our long-run analysis, which relies upon long-difference and rich cross-sectional models, reveals an imprecisely estimated effect that is significantly smaller than the short-run relationship between climate and human capital.

JEL Codes: J24, Q54

Keywords: Adaptation, Climate change, Human capital

THE THREAT OF CLIMATE CHANGE and its increasing prominence in public discourse has inspired a significant body of economic research that explores the potential consequences of such change on a variety of outcomes.¹ Inspired by the neurological

Joshua Graff Zivin (corresponding author) is at the University of California, San Diego, and the National Bureau of Economic Research (NBER) (jgraffzivin@ucsd.edu). Solomon M. Hsiang is at the University of California, Berkeley, and NBER. Matthew Neidell is at Columbia University and NBER. We thank Eli Berman, Gordon Dahl, and seminar participants at the University of California, San Diego, the University of Illinois–Urbana Champagne, Lausanne, Verona, Royal Holloway University, the Paris Environmental and Energy Economics Seminar, and The Property and Environment Research Center (PERC) for numerous useful comments. We are particularly indebted to Wolfram Schlenker for generously sharing data.

1. These include impacts on such wide-ranging outcomes as agriculture, human health, and economic growth. See, e.g., Mendelsohn et al. (1994), Schlenker et al. (2006), Deschênes and Greenstone (2007), Burke et al. (2009), Schlenker and Roberts (2009), Feng et al. (2010), Hsiang (2010), Nordhaus (2010), Deschênes and Greenstone (2011), Dell et al. (2012), Barreca et al. (2013), Graff Zivin and Neidell (2013), Sinha and Cropper (2013).

Received May 31, 2016; Accepted June 7, 2017; Published online November 6, 2017.

JAERE, volume 5, number 1. © 2018 by The Association of Environmental and Resource Economists. All rights reserved. 2333-5955/2018/0501-0003\$10.00 <http://dx.doi.org/10.1086/694177>

literature that documents the brain's sensitivity to temperature (Bowler and Tirri 1974; Schiff, and Somjen 1985; and Hocking et al. 2001), we provide the first estimates of the potential impacts of climate change on cognitive performance and attainment, focusing on the impacts from both short-run weather and long-run climate. Given the importance of human capital as a principal driver of economic growth (e.g., Nelson and Phelps 1966; Romer 1986), these relationships represent an important and unexplored channel through which climate change may impact economic well-being.

Our analysis, which focuses on the same study population over both the short and long run, is to our knowledge the first of its kind and serves an important purpose.² Comparisons across the two models provide a framework through which we can examine the potential offsetting effects from adaptive behaviors, which are expected to play a critical role in determining the ultimate impacts of a gradual changing climate in the coming century (IPCC 2007; Libecap and Steckel 2011). As such, our analysis has significant implications for the interpretation of other results in the literature, as most economic studies of climate change impacts rely on identification from short-run weather phenomena.³

We begin our analysis by focusing on the relationship between weather and cognitive performance. We use assessments of cognitive ability from the children of the National Longitudinal Survey of Youth (NLSY79) and merge these data with meteorological conditions at the county level on the day of the assessment. We take advantage of the longitudinal nature of the survey to estimate models with child fixed effects, exploiting the exogenous interview date and daily fluctuations in weather across the same children over time to identify the causal effect of temperature on cognitive performance.

Using a flexible specification for temperature, we find that math performance declines linearly above 21°C (70°F), with the effect statistically significant beyond 26°C (79°F).⁴ We do not find a statistically significant relationship with the two assessments of reading performance. The disparity across mathematics and other subjects is consistent with differences in the heat sensitivity of the regions within the brain on which they rely (Hocking et al. 2001; Kiyatkin 2007). These differential effects across cognitive tasks also generally support a neurological rather than behavioral explanation for our results, a finding further bolstered by evidence that child's time to completion of each assessment is not related to temperature.

2. A similar approach has been taken to examine adaptation in the very different context of agriculture (Burke and Emerick 2012). As we describe below, our approach differs due to the dynamic accumulation of impacts in the human capital context.

3. See Dell et al. (2014) for discussion of the key conceptual challenges in translating results from short-run analyses to the long run.

4. Note that assessments were only conducted during the spring and summer, so we cannot explore the effects of colder temperatures on performance.

While the negative impacts from idiosyncratic and short-lived weather shocks have potentially important implications for the optimal scheduling of cognitively demanding tasks, the key policy question regarding long-run human capital impacts under climate change depends on the impact of a permanent shift in the distribution of weather outcomes. As such, the second stage of our analysis exploits two approaches to capture the long-run effects: long-difference fixed effects models that examine the impacts of average weather exposure between tests and cross-sectional regressions with extremely rich controls, including parental and grandparental human capital, to examine the impacts of climate exposure from birth until test taking.

Despite large effects on cognitive performance in the short run, we fail to find evidence that climate is significantly related to human capital accumulation in the long run. While the imprecision of these estimates does not allow us to rule out economically meaningful effects in the long run, the effects are significantly smaller than the projections based on short-run estimates. Moreover, allowing for a flexible functional form for temperature reveals a flat relationship between temperature and human capital over the entire temperature range, further corroborating the lack of a relationship between temperature and human capital in the long run.

The difference between our approaches for the short and long run is important because they potentially capture two distinct adaptation channels that have generally been conflated in the literature.⁵ Ex ante avoidance behavior, such as technological adoption, mobility, and cultural changes designed to buffer against the effects of climate, may limit exposure to temperature extremes.⁶ Our short-run regressions will net out all such avoidance at least insofar as they have been adopted based on historical climate up until the time of the test. Ex post compensatory behavior occurs when individuals respond to insults on hot days through subsequent investments that partially or fully offset short-run effects, thus minimizing their enduring impact. For example, if a child learns less material during a hot day, parents or teachers may invest additional time or the child may increase her effort in the following days, potentially offsetting the effect of lost learning. This compensating behavior encompasses a wide range of potential responses, and is almost certainly costly, but persistent human capital effects may thus be minimized in the long run. Such ex post behaviors will only be captured by our

5. This distinction is conceptually similar to that made by Graff Zivin and Neidell (2013) with respect to the health effects from pollution. Individuals can engage in avoidance behavior by spending more time indoors or ameliorate the impacts of exposure through the use of medical inputs, such as asthma inhalers.

6. See Deschênes and Moretti (2009), Deschênes and Greenstone (2011), and Barreca et al. (2013) for evidence on the impacts of adaptation on the relationship between temperature extremes and mortality. An example of a cultural change that reflects adaptation is differences in school schedules throughout the country: schools in southern states typically end in May, a month before schools in northern states.

long-run analysis since responses are predicated on the feedback from earlier tests and thus only depend on weather/climate indirectly.

That said, we would be remiss if we did not offer an alternative explanation for the absence of a long-run effect. Test scores are a composite measure of knowledge and performance, and the intertemporal dependencies of one on the other are largely unknown. Thus, it is possible that short-run changes in performance simply do not add up to sizable long-run changes in learning. Given our parameter estimates and simulations, the plausibility of this explanation hinges on a near-zero relationship between the two, but absent a clear mapping between our short-run and long-run outcome measures, such an explanation remains a possibility. Further research is required to draw a more definitive conclusion regarding the precise mechanism that underlies our results. Nonetheless, our analysis highlights the caution needed when using results from short-run weather shocks to project long-run climate impacts.

This paper unfolds as follows. In section 1, we provide some relevant neurological information on temperatures and brain functioning. Section 2 provides a simple conceptual framework for our econometric models. In section 3, we describe our data in more detail. Section 4 discusses the empirical strategy for the short-run analysis and presents results on the relationship between temperature and test scores. Section 5 describes the empirical strategy for the long-run analysis and presents results on the relationship between climate and human capital. Section 6 offers concluding remarks.

1. SCIENTIFIC BACKGROUND

In order for climate to affect human capital, we need a plausible mechanism that relates brain function to ambient temperature. A particularly important and likely pathway is through the environment's effect on brain temperature. The brain's chemistry, electrical properties, and function are all temperature sensitive (Bowler and Tirri 1974; Schiff and Somjen 1985; Deboer 1998; Yablonskiy et al. 2000; Hocking et al. 2001), with theory suggesting that the brain's performance as a computational network will be influenced by these parameters (Doya et al. 2007; Moore and Cao 2008; Varshney 2011). Furthermore, both warm environmental temperatures and cognitive demands can elevate brain temperature. Despite being only 2% of its mass, approximately 20% of the heat released by a human body originates in brain tissue, of which four-fifths is a direct by-product of neuronal signaling (Raichle and Mintun 2006). Under normal conditions, most excess heat diffuses into the bloodstream and is transported to either the skin or lungs, where it is then transferred to the environment. When environmental temperatures rise, heat transfer at the skin and lungs slows, reducing the flow of cool blood to the brain, which can temporarily elevate brain temperatures up to 2.5°C (Nybo and Secher 2004; Kiyatkin 2007).

Higher temperatures could have different effects on different subject areas because they use distinct parts of the brain that are differentially affected by temperature. For example, mathematical problem solving relies on the ability to retain and manipulate

abstract numerical information, functions that are largely housed in the prefrontal cortex, which stores these data in neural circuits. This region of the brain appears particularly sensitive to heat. Recent work finds that neuronal activity—a measure of mental effort—in the prefrontal cortex increases under elevated temperatures in order to achieve the same level of performance on a series of cognitive and psychometric tests as under cooler temperatures (Hocking et al. 2001). As such, it appears that the costs of a given level of cognitive performance rises as temperature increases and that this effect is particularly acute for the set of activities that rely heavily on this region of the brain, for example, mathematical reasoning.

That high temperatures could impair cognitive function is also consistent with experimental evidence that documents impaired brain function in a wide range of domains as a result of heat stress. Military research has shown that soldiers executing complex tasks in hot environments make more errors than soldiers in cooler conditions (Fine and Kobrick 1978; Froom et al. 1993). Exposure to heat has also been shown to diminish attention, memory, information retention, and processing, and the performance of psycho-perceptual tasks (e.g., Hyde et al. 1997; Hocking et al. 2001; Vasmatazidis et al. 2002). The impacts of thermal stress on working memory performance are especially relevant as cognitively challenging tasks rely more heavily on the working memory for multi-step processing.

This heat-related impairment has potentially important implications in both the short and long run. In the short run, inattention, lack of focus, and diminished cognitive function due to high temperatures can harm cognitive performance. Students are not any less intelligent on hot days, they simply struggle to access that intelligence. In the longer run, as children are repeatedly exposed to high temperatures, this lack of focus and diminished cognitive function can inhibit learning and thus retard knowledge accumulation and cognitive attainment. Clearly, the mapping from performance to learning need not be perfect, a point we further elaborate at the end of section 4.

2. CONCEPTUAL FRAMEWORK

Given the dynamic nature of human capital production, insults from warmer temperatures may accumulate, leading to decreases in human capital attainment levels. One of the key questions in this paper is whether sustained exposure to warmer temperatures, as is expected under climate change, results in accumulated effects on cognitive ability. As noted in the introduction, it is possible that short-run test performance and long-run learning are the result of two distinct processes that have little to do with one another. On the other hand, it is also plausible that short-run performance is an indication of impaired cognitive function that leads to diminished attainment in the long run.⁷ In this

7. While there is no empirical evidence to indicate which is more likely, we find the latter explanation more plausible because the point of tests, however imperfect they may be, is to measure learning.

section, we outline a framework for conceptualizing these effects in order to facilitate the interpretation of our econometric models.

We begin with a simple two-period model of cognitive performance. In the first period, performance y is defined as follows:

$$y_1 = f(k_1, [1 - a_1(w_1)] * w_1), \quad (1)$$

where k_1 represents human capital endowments at birth, w_1 is weather exposure in period 1, and a_1 is avoidance behavior in period 1. Avoidance behavior is a transient action, such as turning on air conditioning or staying indoors, which depends upon weather. As such, the second argument in the performance production function ($[1 - a_1(w_1)] * w_1$) can be viewed as a measure of the effective exposure to ambient temperatures that results from this behavioral response to local weather conditions. Any time-invariant changes in behavior, such as the adoption of air conditioning, are excluded from this model because they will be captured empirically through the use of various fixed effects.

Performance in the second period is defined similarly to first-period performance with two key distinctions. Human capital accumulates from earlier periods and individuals now have the opportunity to respond to feedback embodied in their first-period performance through compensatory behaviors. As such, second-period performance is expressed as follows:

$$y_2 = f(k_2, [1 - a_2(w_2)] * w_2, b(y_1)). \quad (2)$$

As with the initial period, performance will depend on human capital levels and exposure to weather conditions, which depends upon ambient weather and avoidance behavior. For simplicity, we assume that $k_2 = k_1 + g(y_1)$ to reflect human capital accumulation between periods, where the function g reflects the growth in human capital, which depends on prior period learning as reflected by test performance. Compensatory behaviors b are an ex post response to performance in period 1. They could include activities such as spending additional time studying or the devotion of time and resources to a more formal tutoring relationship. As noted earlier, a key feature of this behavior is that it does not require that individuals understand that their performance depends on weather.

Our short-run analysis focuses on the impact of weather on the day of the assessment on cognitive performance. Since we do not observe avoidance behavior, our short-run estimate reflects the total derivative of y_t with respect to w_t , as follows:

$$dy_t/dw_t = (1 - a_t) * f'_2(.) - w_t * f'_2(.) * da_t/dw_t, \quad (3)$$

where $f'_2(.)$ is the partial derivative of $f(.)$ with respect to the second argument. The first term represents the direct (neurological) effect of temperature on performance. The second term represents the ex ante behavioral effect of temperature, which de-

depends upon the effectiveness of that avoidance behavior in diminishing the impacts on cognitive performance and the extent of that avoidance behavior.⁸ Our empirical estimates of the short-run impacts will capture the direct effect of temperature net of any avoidance.

Our long-run analysis is focused on the impacts of climate on test performance. In our simple framework, climate is simply a combination of weather exposure in both periods. If we define climate as the vector of weather states $\{w_1, w_2\}$, the impacts of climate on test performance in period 2 can be interpreted as the sum of impacts from contemporaneous and past period weather, plus any component of the ex post response to observed first-period performance $b(y_1)$ that is affected by first-period weather and the dynamic accumulation of human capital. Thus, our long-run estimate reflects the total derivative of y_2 with respect to both elements in c , which can be expressed as follows:

$$\begin{aligned} dy_2/dw_1 + dy_2/dw_2 = & (1 - a_2) * f'_2(.) - w_2 * f'_2(.) * da_2/dw_2 \\ & + f'_1(.) * dg/dy_1 * dy_1/dw_1 \\ & + f'_3(.) * db/dy_1 * dy_1/dw_1, \end{aligned} \quad (4)$$

where dy_1/dw_1 is as defined in equation (3).

The first two terms are identical to those in equation (3) and reflect the contemporaneous effect of weather on second-period performance—both the direct and ex ante avoidance impacts. The third term ($f'_1(.) * dg/dy_1 * dy_1/dw_1$) captures the impacts of first-period weather on learning and thus human capital accumulation by period 2. The fourth term ($f'_3(.) * db/dy_1 * dy_1/dw_1$) captures the impacts of ex post behavioral responses. It appears in this climate analysis precisely because compensatory behavior responds to prior period performance. Thus, the difference between the short-run estimates characterized in equation (3) and the long-run estimates described by equation (4) will reflect the accumulated impacts of weather extremes on learning plus the impacts of any ex post compensatory behaviors undertaken.

3. DATA

The National Longitudinal Survey of Youth (NLSY) is a nationally representative sample of over 12,000 men and women in the United States aged 14–22 in 1979, with participants surveyed annually until 1994 and biannually thereafter. The survey was designed to collect detailed social and economic information for a transitioning demographic. Beginning in 1986, all children of women in the initial sample were surveyed in their homes, with various developmental assessments conducted biannually on a pre-arranged date. We focus on examinations in mathematics, reading recognition, and reading comprehension, which are derived from the Peabody Individual Achievement

8. This expression is identical for both first- and second-period performance.

Tests (PIAT) and transformed into age-specific standardized scores.⁹ These tests are designed to measure cognitive achievement and capture gains in knowledge over time, making them a popular measure of human capital in the economics literature (e.g., Todd and Wolpin 2007). All three tests, which were administered to children age five and over, have been found to have high test-retest and concurrent validity (Rodgers et al. 1994). Each child is tested across multiple waves for as long as the child is part of the survey, with test data available as early as 1988 and as late as 2006 depending on the age of the child. In our sample of 8,003 children, 80.9% were tested more than once and 41.2% were tested at least four times, enabling us to precisely estimate within-child effects of temperature. Since these tests were predominantly given during the warmer periods of the year,¹⁰ our analysis of short-run temperature effects will only be informative for temperatures in this range.

Using eight waves of the geocoded version of the NLSY, which contains the child's county of residence at each survey wave, we match data on each child's test scores with the average temperature in their county on the day of their assessment using data from Schlenker and Roberts (2009), who linearly interpolated temperatures at each county centroid using the seven nearest stations with daily temperature data. County temperature is defined as (maximum temperature + minimum temperature)/2, computed daily at the geographic centroid of each county and matched to the county of residence for each child for each wave of the survey. We also assign precipitation, specific humidity, wind speed, and pressure in an analogous fashion. We repeat a similar procedure for assigning climate, except that we match the full history of temperature (and the other meteorological variables) between successive tests and from birth until the date of the test.

Since temperature is likely to have a nonlinear relationship with our outcomes of interest, we use various definitions in our analysis. In the short-run analysis, we use both the number of degree days above 21°C (DD > 21) and below 21°C (DD < 21), as well as a nonparametric specification with a full set of indicator variables for every 2°C. As will become clear, our choice of 21°C for the degree day model was chosen based on the nonparametric analysis that revealed a kink at that level. This degree day measure is useful in studies of temperature impacts when (1) a response to daily temperature is roughly constant across days but changes nonlinearly in temperature and (2) the response to daily temperature can be well approximated by a piece-wise linear function, with kinks

9. Despite the availability of additional assessments, we focus solely on these three assessments because they were the most frequently administered across the widest age range, thus yielding the largest sample size and greatest opportunity to explore long-run outcomes.

10. Assessments were conducted between May and October, except for 1986, which was conducted between February and April. To ensure common overlap across seasons and years, however, we exclude the 1986 wave.

at the specified cut-off temperatures.¹¹ The use of indicator variables is even more flexible, allowing for a nonparametric relationship between temperature and performance.

For the long-run analysis, we use three measures of climate for the between test and lifetime exposure models. First, we take the average of the number of degree days above 21°C over the relevant time period. Second, we take the average of the 2°C indicator bins for temperature, which amounts to the percentage of days in each bin. Third, we calculate the mean January–February and July–August temperatures over the relevant time period to provide estimates with a more intuitive interpretation. We also compute the same time-period averages for humidity, precipitation, wind speed, and pressure.

Table 1 contains summary statistics for the data used in this study. Our final sample includes 8,003 children across 951 counties in 48 states that received exams during multiple survey waves. The average child completes 3.66 exams with an average of 2.15 years between them. Children’s test scores are at roughly the national median. Since assessments were conducted in warmer months, average temperature exposure on the day of the test is relatively warm, at 22.8°C (73°F), reflecting the fact that the assessments were conducted during the warmer months. Although children were given the PIAT assessments for all three subjects, discrepancies in sample sizes largely reflect differences in the ability to convert raw scores into standardized and percentile scores (Baker and Mott 1989). Since weather is unrelated to the probability of a test score being available (shown below), we are not concerned that these differences induce a sample selection bias.

4. THE SHORT RUN: TEMPERATURE AND COGNITIVE PERFORMANCE

To explore the short-run relationship between temperature and cognitive performance, we estimate linear fixed effects regression models of the following form:

$$y_{i,t} = f(\beta^{\text{SR}}, T_{c(i),t}) + \eta_1 X_{it} + \eta_2 Z_{c(i),t} + \pi(t, s(i)) + \alpha_{c(i)} + \varepsilon_{it}. \quad (5)$$

The test score (y) of child i on date t is regressed on the temperature faced by that child in county c on the same day ($T_{c(i),t}$).¹² The term β^{SR} reflects dy_t/dw_t from equation (3). We include the child’s age (X_{it}) and other meteorological variables ($Z_{c(i),t}$) that may confound the relationship between temperature and test scores. Our regression models also control for the month and weekday of the assessment and state-

11. Degree days are defined as the number of degrees by which the average daily temperature exceeds 21°C, with values below 21°C assigned a value of 0. The degree day approach has been widely used to study the nonlinear impact of temperature on crop yields (e.g., Schlenker and Roberts 2009), electricity demand (e.g., Auffhammer and Aroonruengsawat 2011), and GDP growth (Hsiang 2010).

12. We also explore lagged temperatures as well, shown below.

Table 1. Summary Statistics

| | | A. Cognitive Outcomes | | | |
|----------------------------|------------|-------------------------|-------|-------|--------------|
| | | N | Mean | SD | Within SD |
| Math | Percentile | 24,361 | 49.70 | 27.45 | 14.37 |
| | Length | 10,389 | 9.98 | 4.91 | 3.14 |
| | Energy | 24,260 | .11 | .31 | .25 |
| Reading comprehension | Percentile | 20,439 | 51.60 | 27.74 | 14.34 |
| | Length | 8,557 | 10.19 | 5.26 | 3.15 |
| | Energy | 20,041 | .16 | .36 | .28 |
| Reading recognition | Percentile | 24,229 | 56.52 | 28.50 | 13.00 |
| | Length | 10,367 | 3.59 | 1.92 | 1.25 |
| | Energy | 22,814 | .15 | .35 | .28 |
| No. of tests per child | | 24,361 | 3.66 | 1.20 | ... |
| No. of years between tests | | 16,304 | 2.15 | .64 | ... |
| | | B. Temperature Measures | | | |
| | | N | °C | SD | Within SD °F |
| Day of test: | | | | | |
| Temperature | | 24,361 | 22.77 | 4.96 | 3.12 72.99 |
| Degree days ≥ 21 | | 24,361 | 3.05 | 3.12 | 1.85 5.49 |
| Degree days < 21 | | 24,361 | 1.28 | 2.65 | 1.92 2.30 |
| Between tests: | | | | | |
| Degree days ≥ 21 | | 16,304 | 1.21 | .99 | .22 2.18 |
| Degree days < 21 | | 16,304 | 13.07 | 3.75 | .71 23.52 |
| January–February | | 16,304 | 4.61 | 6.86 | 1.33 40.30 |
| July–August | | 16,304 | 24.57 | 3.03 | .63 76.23 |
| From birth: | | | | | |
| Degree days ≥ 21 | | 24,294 | 1.23 | .98 | ... 2.22 |
| Degree days < 21 | | 24,294 | 13.08 | 3.68 | ... 23.54 |
| January–February | | 24,294 | 4.12 | 6.85 | ... 39.42 |
| July–August | | 24,294 | 24.53 | 2.95 | ... 76.15 |

Note. Length is the number of minutes for the child to complete the evaluation. Energy is the percentage of children reported by the interviewer as acting lethargic. Temperature measures are reported for the math test only. “Day of test” reflects temperature the day the test is taken. “Between tests” reflects the mean temperature between successive tests. “From birth” reflects the mean temperature from birth until the time of the test. Degree days ≥ 21 (< 21) is the number of degrees above (below) 21°C. January–February and July–August are the mean temperatures for those months.

specific nonlinear time trends in order to capture time-varying factors that influence performance ($\pi(t,s(i))$). Importantly, these time trends will capture any changes in time-invariant adaptive behaviors during our period, such as air-conditioning penetration or other avoidance “technologies,” to the extent they are common to families in each

state.¹³ The longitudinal nature of the survey enables us to specify child fixed effects ($\alpha_{c(i)}$), which controls for all time invariant characteristics of a child. The disturbance term ($\varepsilon_{i,t}$) consists of an individual idiosyncratic component and a clustered component by state, which serves three purposes: to allow for arbitrary spatial correlation across counties within a state, to allow for autocorrelation in test scores over time, and to account for the fact that the same temperature measure can be assigned to multiple children. Since the date the child has the assessment is prearranged, it is unlikely to respond to short-run changes in temperature and thus plausibly exogenous, allowing us to identify the causal effect of temperature on performance.¹⁴

As described earlier, temperature is included in our model using two distinct approaches to explore its potentially nonlinear relationship with performance: (1) a series of indicator variables for temperatures in 2°C bins from 12°C to 32°C, with 20°C–22°C as the reference category; and (2) a linear function in heating and cooling degree days with a cutoff at 21°C, so chosen because the temperature bin at 20°C–22°C was the local maximum in the first approach.

Table 2 presents the core short-run results for our three test outcomes of interest: math, reading comprehension, and reading recognition. Given our interest in temperature extremes at the high end, we begin with a specification that only includes degree days above 21°C. We then add degree days below 21°C to capture any effects that might occur at lower temperatures. Columns 1 and 3 present results with mother fixed effects (since siblings are in the sample). Columns 2 and 4 present results with child fixed effects. The math results, shown in panel A, indicate that warmer temperatures lead to a statistically significant decrease in performance. The results are insensitive to whether we include degree days below 21°C and to the type of fixed effect used.¹⁵ The estimate of -0.219 in the first row of column 4 implies that each degree day above 21°C lowers the math score by 0.219 of a percentile.

13. According to the 2001 American Housing Survey, 79.5% of all households had some form of air conditioning, with the rate of ownership much higher in warmer regions (Graff Zivin and Neidell 2014). We directly probe the role of air conditioning in blunting the short-run impacts of temperature on cognition in the appendix, available online. While the results are statistically insignificant, our coefficients change in the expected direction, that is, air conditioning appears to be protective against cognitive harm, though small in magnitude.

14. We test our exogeneity assumption by separately regressing the probability that a child is male, Hispanic, black, the child's age in months, the child's height in inches, the mothers age at the child's birth, and the birth order of the child, on our full suite of temperature dummies as well as county and state-by-year fixed effects (results available upon request). We find no systematic or significant patterns of selection by these observables with respect to the temperature on the day of the interview and exam.

15. The insignificance of degree days below 21 should be interpreted with caution since very few exams occur on cold days. As such the coefficient on $DD < 21$ largely reflects the impacts of moderate temperatures on test performance.

Table 2. Fixed Effect Estimates of Relationship between Short-Run Temperature and Cognitive Performance

| | (1) | (2) | (3) | (4) |
|---------------------------|-------------------|-------------------|--------------------|-------------------|
| A. Math: | | | | |
| Degree days ≥ 21 | -.211* [.0903] | -.205* [.0960] | -.240** [.0925] | -.219* [.0984] |
| Degree days < 21 | | | -.151 [.0899] | -.0749 [.0934] |
| Fixed effect | mother | child | mother | child |
| Observations | 24,361 | 24,361 | 24,361 | 24,361 |
| R-squared | .551 | .737 | .551 | .737 |
| B. Reading comprehension: | | | | |
| Degree days ≥ 21 | .0607 [.102] | .0611 [.102] | .0524 [.104] | .0711 [.104] |
| Degree days < 21 | | | -.0434 [.0942] | .0509 [.0985] |
| Fixed effect | mother | child | mother | child |
| Observations | 20,439 | 20,439 | 20,439 | 20,439 |
| R-squared | .601 | .779 | .601 | .779 |
| B. Reading recognition: | | | | |
| Degree days ≥ 21 | -.027 [.0899] | .0441 [.0875] | -.0325 [.0919] | .0461 [.0896] |
| Degree days < 21 | | | -.0286 [.0856] | .0101 [.0828] |
| Fixed effect | mother | child | mother | child |
| Observations | 24,229 | 24,229 | 24,229 | 24,229 |
| R-squared | .587 | .802 | .587 | .802 |

Note. The above coefficients reflect estimates of the relationship between temperature on the day of the test and cognitive performance. Standard errors clustered on state-week in brackets. All regression models control for precipitation, pressure, wind speed, humidity, and dummy variables for day of week, month, year, and state by year. Regressions with mother fixed effects also control for child sex, birth order dummies, age of mother at birth of child, and child birth weight dummies.

* $p < .05$.

** $p < .01$.

In contrast, we find that temperature does not have a statistically significant relationship with reading recognition or reading comprehension, regardless of the specification, shown in panels B and C.¹⁶ As described earlier, one potential explanation for the discrepancy in impacts by subject is that mathematical problem solving utilizes

16. Our core results remain unchanged when we drop all non-temperature weather controls, as can be seen in appendix table 3.

functions of the brain that are distinct from the other subject areas, and different parts of the brain are differentially affected by temperature.¹⁷

Figure 1 plots estimates for each of our three outcome variables using the more flexible specification for temperature. Shown in panel A, we find that child performance in mathematics shows a monotonic decline in outdoor temperatures above 22°C (71.6°F) but is relatively flat and statistically insignificant for temperatures below this point. Furthermore, two of the estimates in the four highest bins are individually statistically significant at the 5% level, with the other two at the 10% level. This monotonic relationship at the high end reassures us that the significant estimates for math in table 2 are not simply the result of a Type I error. We interpret the magnitude of our estimates as follows: changing the temperature of the outdoor environment from 20°C–22°C (68°F–71.6°F) to 30°C–32°C (86°F–89.6°F) decreases a child's mathematics score by 1.6 percentile points, which is a sizable 0.12 of a standard deviation. The predicted effect using estimates from our degree days specification matches these results quite closely, suggesting that our math estimates are largely insensitive to how we specify our temperature variable.

The nonparametric results for both reading outcomes, shown in panels B and C of figure 1, are consistent with the results in table 2. The coefficients are small, statistically insignificant, and relatively flat across the entire temperature range, providing additional support for the conclusion that performance on these measures is unaffected by temperature.

One potential concern about the results thus far relates to migration. The inclusion of movers could bias our results if moving to a location with a very different climate were correlated with other factors that determine test scores. To address this concern, we classify any individual as a "mover" if the difference in average July–August temperatures between their origin and destination location is >2 degrees Celsius. We then use this new variable to address potential concerns about bias using two approaches. Panel A of table 3 presents results when we include mover status as a covariate in our regression by defining it as a time-varying variable that takes the value of 1 when someone moves (and 0 otherwise). Our point estimates are identical to those found in table 2. Panel B of table 3 limits our analysis to those who never move and here again the results are very similar to those obtained when we include the full sample of subjects. Together, these results suggest that migration-induced bias is not a significant concern in this setting.

Another issue relates to the interpretation of these results since weather may affect the child's value from alternative activities, which may affect the child's effort on the exam. For example, warmer weather makes playing outside more attractive, and a child

17. We note that math is always the first of the three exams, so increased fatigue cannot explain this pattern.

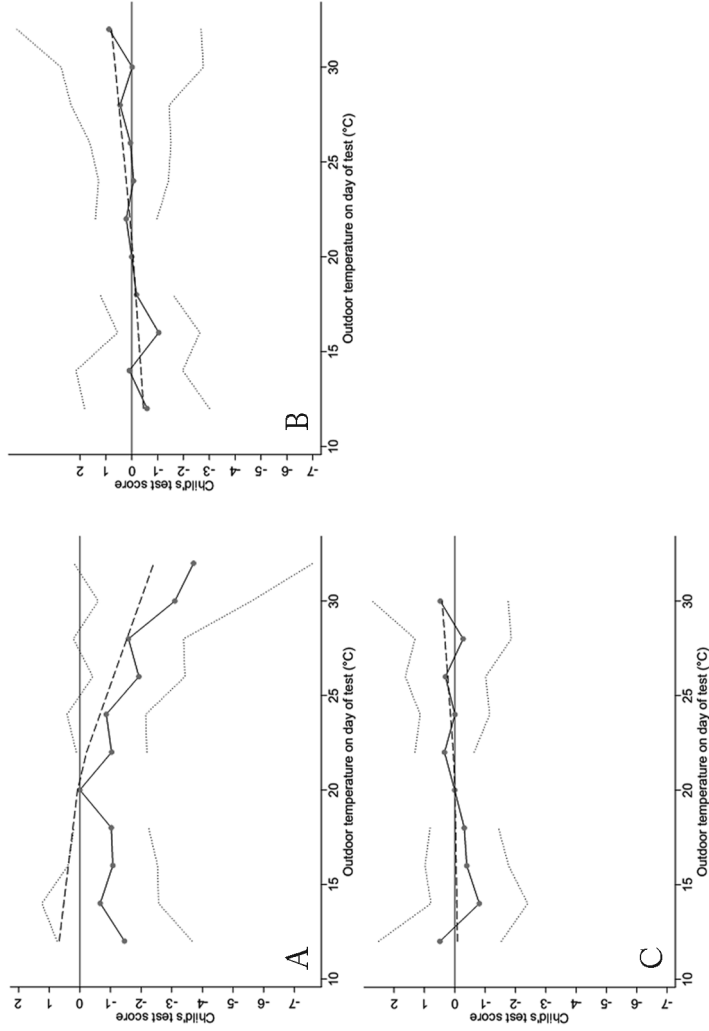


Figure 1. Relationship between short-run temperature and cognitive performance. *A*, Mathematics. *B*, Reading comprehension. *C*, Reading recognition. The solid line shows coefficient estimates of the effect of temperature on the day of the test of cognitive performance, with 95% confidence intervals in dotted lines. The regression includes indicators for each 2°C, linear controls for precipitation, pressure, wind speed, humidity, dummy variables for day of week, month, year, and state by year, and child fixed effects. The predicted effect using degree days above and below 21°C is shown in the dashed line and is based on a regression using the same set of controls.

Table 3. Migration and the Short-Run Temperature Relationship

| | Math | Reading Comprehension | Reading Recognition |
|-----------------------------------|--------------------|-----------------------|---------------------|
| A. Controlling for moving status: | | | |
| Degree days ≥ 21 | -.219** [.0867] | .0712 [.116] | .0465 [.0814] |
| Degree days < 21 | -.0745 [.0878] | .0501 [.181] | .0104 [.117] |
| Observations | 24,361 | 20,439 | 24,229 |
| R-squared | .737 | .779 | .802 |
| B. Never movers: | | | |
| Degree days ≥ 21 | -.185* [.101] | .0944 [.116] | .0686 [.0849] |
| Degree days < 21 | -.0818 [.0882] | .062 [.188] | .0132 [.124] |
| Observations | 22,922 | 19,375 | 22,809 |
| R-squared | .739 | .78 | .802 |

Note. The above coefficients reflect estimates of the relationship between temperature on the day of the test and cognitive performance. Standard errors clustered on state in brackets. All regression models linearly control for precipitation, pressure, wind speed, humidity, and include child fixed effects, dummy variables for day of week, month, year, and state by year. Panel A includes an indicator if the child moved between survey waves, and panel B limits the sample to never movers.

* $p < .05$.

** $p < .01$.

may rush through the assessment in order to play outside, which lowers her performance.¹⁸ Given the differential effect by subject, such a mechanism seems unlikely in this setting. Nonetheless, we probe this channel using data on the speed at which children complete tests. If children hurry to finish the assessment in warmer weather, then time to completion will fall as temperatures rise. Such data were collected in the 1994, 1996, and 1998 waves, thus providing a useful test for detecting changes in the child's effort on the day of the test.

Consistent with the differential effects by subject, we fail to find evidence to support this channel. Column 1 of table 4 fails to find a statistically significant relationship between temperature and assessment completion time for all three assessments. Allowing for a more flexible specification for temperature, shown for math in figure 2, we see a generally flat relationship between temperature and time to completion, with only a decline at the highest temperature bin, though it is not statistically significant. This drop at the highest bin, while large in magnitude, does not coincide with the general pattern of temperature effects on performance, suggesting that completion time is unlikely to explain our short-run findings.

18. In the conceptual framework, this would amount to $\delta b_t / \delta w_t < 0$.

Table 4. Robustness Checks for Estimates of Short-Run Temperature Relationship

| | Time to Completion (1) | Assessment Completed (2) | Start Time (3) |
|---------------------------|---------------------------|-----------------------------|-------------------|
| A. Math: | | | |
| Degree days ≥ 21 | -.011 [.054] | -.0012 [.0012] | .0138 [.0447] |
| Degree days < 21 | .004 [.052] | -.0015 [.0013] | -.0282 [.0418] |
| Observations | 8,620 | 26,091 | 8,621 |
| R-squared | .663 | .566 | .707 |
| B. Reading comprehension: | | | |
| Degree days ≥ 21 | -.038 [.064] | -.0008 [.0019] | |
| Degree days < 21 | -.025 [.069] | -.0038* [.0019] | |
| Observations | 7,092 | 26,062 | |
| R-squared | .725 | .52 | |
| C. Reading recognition: | | | |
| Degree days ≥ 21 | -.011 [.023] | -.0010 [.0013] | |
| Degree days < 21 | -.051* [.024] | -.0015 [.0014] | |
| Observations | 8,597 | 26,089 | |
| R-squared | .65 | .55 | |

Note. The above coefficients reflect estimates of the relationship between temperature on the day of the test and the following outcomes: column 1 is the number of minutes it takes the child to complete the assessment; column 2 is whether the assessment is completed; and column 3 is the start time of the assessment, only relevant for math since it is always first. Standard errors clustered on state-week in brackets. All regression models control for precipitation, pressure, wind speed, humidity, and include dummy variables for day of week, month, year, and state by year, and child fixed effects.

* $p < .05$.

** $p < .01$.

Table 4 also provides two additional robustness checks. First, the different sample sizes across the subjects (as seen in table 1), which indicates that some scores are unavailable for children, is a potential concern. In particular, one might worry about sample selection bias if the missing test scores correlate with warmer temperatures, perhaps because families cancel the visit or the child scored below a certain value, making a standardized score infeasible. To assess this, we regress our weather variables on score availability. As shown in column 2 of table 4, we find that probability of completing the assessment is unrelated to warmer weather, suggesting that sample composition across subjects is unlikely to bias our results. Second, one might be concerned

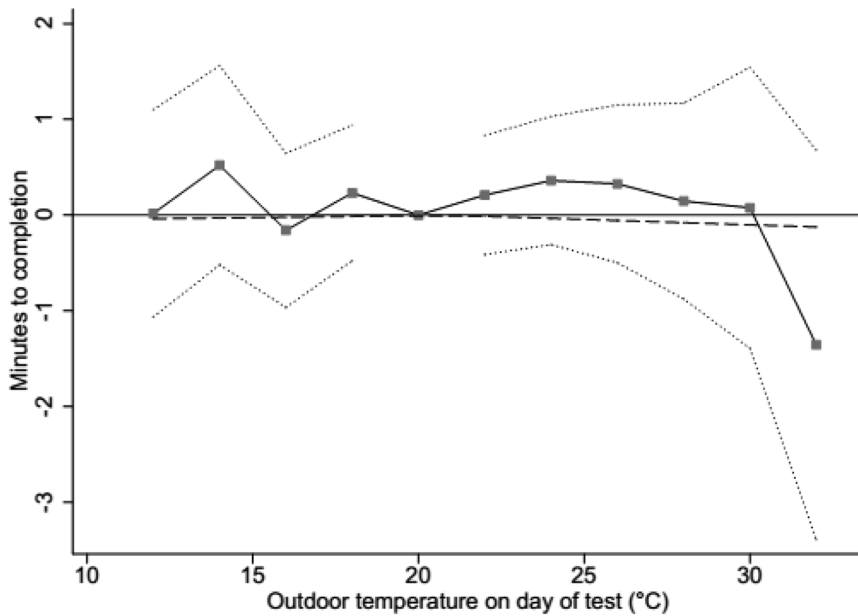


Figure 2. Relationship between short-run temperature and time to completion for math assessment. The solid line shows coefficient estimates of the effect of temperature on the day of the test, with 95% confidence intervals in dotted lines. The regression includes indicators for each 2°C, linear controls for precipitation, pressure, wind speed, humidity, dummy variables for day of week, month, year, and state by year, and child fixed effects. The predicted effect using degree days above and below 21°C is shown in the dashed line and is based on a regression using the same set of controls.

that exams are shifted to cooler times of the day to avoid peak exposure, thereby minimizing the effect on performance. In column 3, we show results using the start time of the assessment as the dependent variable, and find the start time is unrelated to the temperature on the day of the test.¹⁹

Our analysis has thus far focused solely on the effects of weather on the day of the assessment, thereby ignoring potential lagged effects. While the neurological mechanisms discussed in section 1 suggest a rather immediate effect from exposure, lagged exposure has been shown to affect health and thus might also affect performance.²⁰ Figure 3 presents results when we add three lags of temperature, and also one lead of

19. We only show results for math because it is always the first test given.

20. Furthermore, since we do not know the exact time assessments were given for all years, we may be assigning weather with error. A lagged specification may better capture exposure for those with, for example, morning exams. We note, however, that for the sample years in which we observe assessment times, the average start time is 2:41 p.m.

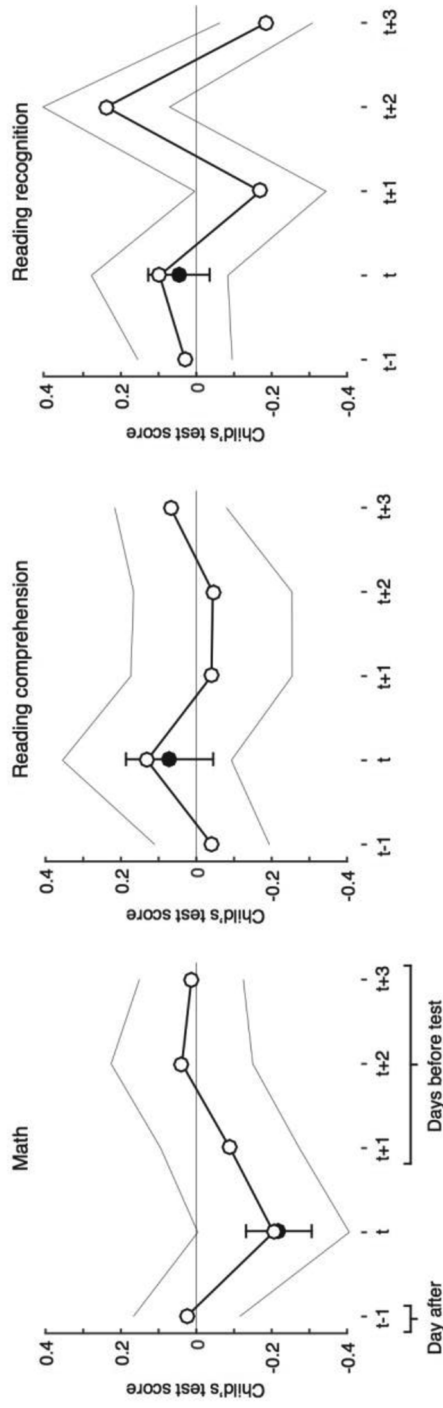


Figure 3. Timing of exposure and test performance. The open circles and thick solid line show coefficient estimates of the effect of degree days ≥ 21 on cognitive performance for three days before, the day of, and one day after each test (thin solid lines are 90% confidence intervals). The regression also controls for degree days < 21 three days before, the day of, and one day after the test as well as linear controls for precipitation, pressure, wind speed, humidity. They also include dummy variables for day of week, month, year, and state by year, and child fixed effects. The estimate for the contemporaneous only model (table 2, col. 4) is shown with the solid circle, along with the 90% confidence interval.

temperature as a falsification test. The coefficient on contemporaneous temperature is largely unchanged (though precision is compromised) and the coefficients on lagged temperature are considerably smaller than our main estimate for math.²¹ The wide confidence intervals are likely the result of multicollinearity since weather is highly correlated across consecutive days.²² The absence of an effect for lead temperature offers further assurance that our results are not driven by unobserved confounding. For the two reading outcomes, we fail to find a statistically significant effect from any period. In the end, and despite moderate to high levels of air-conditioning penetration, these results provide strong evidence for a contemporaneous and negative effect of warmer temperatures on mathematical performance.²³

Before turning to our long-run analysis, we simulate the potential long-run effects based on our short-run estimates to give a sense for the potential magnitude of the effects we might find. To do so, we need to make some assumptions about the human capital accumulation process (the function $g(\cdot)$ in our conceptual model). First, we assume that children's percentile performance on the test is equivalent to a ranking in human capital. Absent negative shocks, children accumulate human capital at comparable rates so that their rank remains unchanged. Second, exposure to a simulated distribution of weather shocks, measured in 2°C bins, leads to a reduction in rank. We assume this reduction is permanent. Further, we assume that a change in performance, absent compensatory behaviors, amounts to a change in learning as described in sections 1 and 2. Although our short-run estimates do not directly test for learning effects, the ergonomics literature reviewed above demonstrates effects of heat on memory, information retention, and information processing, suggesting the plausibility of such an effect (Hyde et al. 1997; Hocking et al. 2001; Vasmatazidis et al. 2002). Given uncertainties about the magnitude of this learning effect, we take the simplest approach and scale the change in performance by λ , under a variety of parameter values $\lambda = \{0.1, 0.3, 0.5\}$. In the context of our conceptual framework, this means $g(y_t) = \lambda * y_t$. For example, $\lambda = 0.5$ implies that a 10-percentage-point decrease in test performance translates into a 5-percentage-point decrease in learning. We then accumulate the shocks between tests based on the realized weather exposure for each child to

21. The effect of the prior day's temperature is roughly half the size of our main estimate, suggesting that temperature impacts on cognitive function might accumulate over short periods or that hot nights might interfere with sleep, although this effect is not significant.

22. In fact, the correlation between today's temperature and any of the lags is never less than 0.78. Adding additional lags to our specification leads to an inconsistent pattern of results that is the hallmark of a multicollinearity problem. We are reassured by the fact that including lagged monthly temperature, which is much less correlated than daily lags, has a minimal impact on our estimates (results not reported).

23. As noted earlier, approximately 80% of US households had some form of air conditioning in 2001, with the rate of ownership much higher in warmer regions (Graff Zivin and Neidell 2014).

compute the change in human capital. Our simulation based on our short-run estimates under the quite conservative assumption that $\lambda = 0.1$ implies that children's exposure between tests would reduce performance by 6.2 percentile points on average. At $\lambda = 0.3$ and 0.5 , this rises to 18.6 and 31.0 percentile points, respectively, thereby implying quite large long-run effects. If test performance is determined by factors unrelated to learning, such that the two measures are completely independent of one another, this is tantamount to assuming that $\lambda = 0.0$.

5. THE LONG RUN: CLIMATE AND HUMAN CAPITAL

The analysis thus far focused on the contemporaneous impacts of temperature on performance. In this section, we turn our attention to the long-run impacts of climate on human capital. We estimate two distinct models.

In our first approach, we estimate a "long difference" model of the following form:

$$y_{i,t} - y_{i,t-1} = f\left(\beta^{\text{LD}}, \sum_{t=1}^t T_{c(i)}\right) + \eta_1(X_{it} - X_{it-1}) + \eta_2 \sum_{t=1}^t Z_{c(i)} + \pi(t, s(i)) + (\varepsilon_{i,t} - \varepsilon_{i,t-1}). \quad (6)$$

The dependent variable is the change in performance over time, which reflects the accumulation of human capital between tests. The variable $\sum_{t=1}^t T_{c(i)}$ reflects our measure of climate, which is a summary measure of temperature between successive tests. We continue to define $T_{c(i)}$ in degree days and indicator bins as before. Given the different structure of this model, the interpretation of β^{LD} now takes a slightly different form. When we use degree days, we interpret β^{LD} as the increase in human capital from a 1°C degree day increase in temperature across all days between tests. When we use indicator bins, we interpret β^{LD} as the increase in human capital from a 1% increase in the number of days that the temperature falls in a certain bin (relative to 20°C–22°C) between tests. For example, we would interpret the coefficient on the 30°C–32°C bin as the effect from shifting 1% of all days between successive tests from 20°C–22°C to 30°C–32°C. To better align with intuition we also use seasonal average temperature (separately for January–February and July–August). In this case, the coefficients reflect the impacts of a 1°C increase in the mean July–August (or January–February) temperature between tests on human capital accumulation. The other meteorological variables (Z) are defined analogously, while the variables X_{it} and $\pi(t, s(i))$ remain unchanged from equation (5). Recall that $\pi(t, s(i))$ includes a state-by-year dummy variable, which controls for any differential trends in warming across states.²⁴

24. This does not completely eliminate concerns about how shared national trends in warming might influence our results, since part of our identifying variation comes from comparing different cohorts of children. Nonetheless, that threat should be small since the United States warmed by less than 0.5 degree Celsius, on average, during our study sample period.

By defining the model in long differences (LD), the model may capture a wider range of adaptive responses (Dell et al. 2014), where the coefficient $\beta^{LD} = dy_2/dw_1 + dy_2/dw_2$ from equation (4). For example, if parents respond to poor performance in school with compensatory investments, regardless of whether they know the source of the poor performance, our estimate for β^{LD} is net of this investment. The model also remains well identified because we are controlling for all time invariant characteristics of the child.

In our second approach, we assign climate as the accumulated temperature from birth until the date of the test, hence providing an even longer-term measure of climate exposure. This necessitates the use of cross-sectional models, which leaves greater scope for omitted variable bias since parents choose where to raise their children and thus climate exposure may be correlated with other characteristics that influence human capital attainment. To address this concern, we exploit the unusual richness of the NLSY to control for a wide range of background factors in the human capital production function (Black et al. 2005). In particular, we estimate the following regression specification:

$$y_{i,t} = f\left(\beta^{CS}, \sum_0^t T_{c(i)}\right) + \eta_1 X_{it} + \eta_2 \sum_0^t Z_{c(i)} + \eta_3 X_{m(i)} + \pi(t, s(i)) + \varepsilon_{i,t}. \quad (7)$$

Climate ($\sum_{t=1}^t T_{c(i)}$) is now measured as lifetime exposure from birth until the time of the test, and we continue to specify this in terms of degree days, indicator bins, and seasonal averages. The interpretation of coefficients is similar to the “between-test” model except they now reflect the effect from birth until the time of the test. The other meteorological variables (Z) are defined analogously.

Given the greater concern for omitted variable bias in this specification, we also add several measures that reflect the child’s potential human capital endowment at birth. The term X_i now includes the child’s birth weight, an important measure of intellectual endowments (Black et al. 2007), which we control for flexibly by including a series of indicator variables for each pound. It also includes the child’s sex, birth order indicators, and maternal age at birth. The term $X_{m(i)}$ includes the mother’s scores on the armed forces qualifying test (AFQT), completed years of schooling, a measure of self-esteem, height, weight, race/ethnicity, foreign language, the religion she was raised, and her spouse’s level of education. We also include flexible controls by allowing for all two-way interactions between these variables and third-order polynomials for all continuous variables. Including grandparent characteristics (grandmother and grandfather’s years of schooling, Duncan Socioeconomic Index (SEI), foreign born) in $X_{m(i)}$ further helps to minimize concerns about omitted variable bias. The term Z is also extended to include numerous county-level characteristics, including the age of housing stock, birth rate, death rate, infant mortality rate, physicians per capita, hospital beds per capita, education per capita, household income per capita, fraction below poverty, geographic size of the county, maximum elevation, and whether it borders an ocean or great lake.

Table 5. Long Difference Estimates of Relationship between Temperature between Tests and Math Performance

| | (1) | (2) | (3) |
|-----------------------|-----------------|------------------|------------------|
| Degree days ≥ 21 | -.630 [.344] | -.250 [.466] | |
| Degree days < 21 | | -.414* [.203] | |
| July–August | | | -.196 [.134] |
| January–February | | | -.0905 [.135] |
| Observations | 16,304 | 16,304 | 16,304 |
| R-squared | .034 | .035 | .035 |

Note. The above coefficients reflect estimates of the relationship between temperature exposure between tests and math performance. Standard errors clustered on state-week in brackets. All regression models control for between-test measures of precipitation, pressure, wind speed, and humidity; dummy variables for day of week, month, year, and state by year; and child fixed effects.

* $p < .05$.

** $p < .01$.

Table 5 presents our long-difference results based on the “between test” specification. We focus solely on mathematical performance since this is the only outcome where we find an effect in the short run. In contrast to the short-run results, however, we do not find a statistically significant relationship between climate and human capital. Column 1, which only includes degree days above 21°C (DD > 21), reveals a statistically insignificant estimate of -0.63 . This estimate indicates that a 1 degree day increase in temperature across all days between two tests, a rather substantial change, decreases math performance by only 0.63 percentile points. This suggests that parents are offsetting a substantial fraction of these effects because in the absence of such offsetting behavior we would expect substantially larger losses of human capital. For example, under the conservative case where $\lambda = 0.1$, a 1 cooling-degree-day increase for all days between tests would amount to an average decline of 8.48 percentage points between tests, since children are tested with 2.15 years between tests on average.²⁵ This scenario would imply that parents are offsetting roughly a 7.85 percentile point accumulated decrease in human capital due to warmer temperature exposures, equal

25. The calculation is $8.48 \text{ points} = -0.219 \text{ points}/^{\circ}\text{C} \times 1^{\circ}\text{C} \times 0.1 \times 180 \text{ school days/yr} \times 2.15 \text{ days}$. This calculation is simplified by the fact that scores are in percentiles, so students can be assumed to maintain their rank ordering in the absence of any relative changes in human capital between tests.

to 92.6% of the loss due to temperature. If λ is assumed larger, then the associated estimate for parental offsetting grows. The “residual” damage to human capital suggested by the long difference result is small relative to these potential accumulated losses.

While the long-run effects are much smaller than those found in our short-run analysis, it is still instructive to examine the unmitigated damage to human capital in the long run. Converting to standardized units,²⁶ a 1 cooling-degree-day Celsius increase reduces test scores by 0.017 standard deviations (SD) on average, while a large, but not implausible, 3°C increase in temperature translates into a 0.051 SD effect.²⁷ Thus, even in the long run, we cannot rule out economically meaningful effects on human capital under these extremes.

When we add degree days below 21°C (DD < 21), shown in column 2, our estimate for warmer temperatures is even closer to zero at -0.25 and remains statistically insignificant. Focusing on mean winter and summer temperatures yields estimates that are again statistically insignificant and considerably smaller than the simulated long-run estimates under even highly conservative assumptions. For example, our estimate of -0.196 for July–August suggests that a 1°C increase for every day in those two months decreases math performance by 0.196 percentile points. Overall, across these models there is variation in the magnitude of the point estimate for the effect of warm days, although this variation is consistent with sampling variability given the size of our estimated confidence intervals.

In table 6, we show results using lifetime temperature exposure. Given that these estimates rely on cross-sectional models, we assess the sensitivity of results to slowly adding more controls, continuing to use different assumptions about the functional form for temperature. In the first panel, which only uses degree days above 21°C, we see in column 1 that temperature is negatively associated with math performance. Adding simple controls for maternal human capital, a strong predictor of children’s human capital attainment (Black et al. 2005), raises that coefficient to -0.463 as

26. We use the percentile measure of PIAT math scores because rank ordering of students should remain fixed in the absence of relative human capital changes, making interpretation of these effects intuitive. Percentile scores have an approximately uniform distribution between 0 and 100. Conversion to standardized PIAT math scores, which are approximately normally distributed, is straightforward. Near the mean score, percentile scores are linear in standardized scores, rising 0.375 percentile points per standardized point. One standard deviation in standardized scores is 13.94 points, implying a corresponding change in percentile scores of 37.13 point. Computing a simple standard deviation of percentile scores recovers a standard deviation of 27.47 points.

27. Since the absence of a long-term impact is somewhat imprecise, it is also useful to look at the lower bound of the 95% confidence interval. Focusing on the estimates from column 1 of table 5, the lower 95% confidence interval of our estimates implies that a 1°C increase reduces test scores by 0.035 of an SD, and a 3°C increase reduces test scores by .11 of an SD.

Table 6. Cross-Sectional Estimates of Relationship between Lifetime Temperature Exposure and Math Performance

| | (1) | (2) | (3) | (4) |
|-------------------------------------|--------------------|------------------|------------------|-----------------|
| A. Cooling degree days: | | | | |
| Degree days ≥ 21 | -1.211 [.977] | -.463 [1.116] | -.440 [1.087] | -.329 [.999] |
| Observations | 24,294 | 24,294 | 24,294 | 24,294 |
| R-squared | .151 | .269 | .271 | .281 |
| B. Cooling and heating degree days: | | | | |
| Degree days ≥ 21 | .347 [.962] | .0368 [1.009] | .102 [.970] | .158 [.918] |
| Degree days < 21 | -1.694** [.548] | -.550 [.430] | -.600 [.440] | -.543 [.420] |
| Observations | 24,294 | 24,294 | 24,294 | 24,294 |
| R-squared | .152 | .269 | .271 | .281 |
| C. Seasonal temperatures: | | | | |
| July–August | -.691* [.297] | -.496 [.291] | -.422 [.312] | -.349 [.302] |
| January–February | -.285 [.281] | .237 [.237] | .139 [.235] | .125 [.230] |
| Observations | 24,294 | 24,294 | 24,294 | 24,294 |
| R-squared | .155 | .270 | .273 | .282 |
| County characteristics | Y | Y | Y | Y |
| Geography | Y | Y | Y | Y |
| Child characteristics | Y | Y | Y | Y |
| Maternal human capital | N | Y | Y | Y |
| Grandparent human capital | N | N | Y | Y |
| $f(\text{maternal human capital})$ | N | N | N | Y |

Note. The above coefficients reflect estimates of the relationship between temperature exposure from birth until the time of test and math performance. Standard errors clustered on state-week in brackets. All regression models control for from-birth measures of precipitation, pressure, wind speed, humidity, and dummy variables for day of week, month, year, and state by year. County characteristics includes county-level measures of age of housing stock, birth rate, death rate, infant mortality rate, physicians per capita, hospital beds per capita, education per capita, household income per capita, and fraction below poverty. Geography includes county square miles, maximum elevation, borders seas, and borders great lake. Child characteristics includes sex, birth order dummies, maternal age at birth, and age of child at time of test. Maternal human capital includes mother's years of schooling, AFQT, self-esteem, height, weight, race, foreign language, and religion, dummy variables to indicate when schooling, AFQT and self-esteem were imputed. Grandparent human capital includes grandmother's and grandfather's years of schooling, Duncan SEI, foreign born, and dummies if schooling missing. $f(\text{maternal human capital})$ includes third-order polynomial for all continuous maternal human capital variables and all two-way interactions.

* $p < .05$.

** $p < .01$.

shown in column 2. As we include more control variables, this estimate remains fairly stable and far from statistically significant—although the confidence interval is fairly wide, spanning from -2.33 to 1.67 points per cooling-degree day in the most saturated model. Notably, however, all of these values suggest that the long-run effects are much smaller than would be expected if the short run effects were simply to accumulate. Panel B adds degree days below 21, while panel C uses winter and summer temperatures. In both cases, the estimates show the same general pattern: statistically insignificant estimates that are much smaller than those implied by the simulation exercise. Again, variation in these point estimates is substantial in percentage terms but remains within the expected range given sampling variability.

In figure 4 we show results allowing for the flexible specification in temperature for the “between test” and “from birth” models that matches the indicators used in the short-run model. As with the previous long-run results, we do not find statistically significant estimates. Moreover, we do not find a pattern in the estimates that comports with the short-run results—estimates are relatively flat over the entire temperature distribution. While the precise mechanism that underlies these results is not known, the significantly smaller coefficients across our long-run models are consistent with the notion that individuals engage in nontrivial amounts of adaptation to minimize the effects of high temperature days on the human capital accumulation of children.²⁸

6. CONCLUSION

In this paper, we merge rich data from the NLSY with meteorological data to provide the first economic analysis of the relationship between temperature/climate and human capital. We find that short-run changes in temperature lead to statistically significant decreases in cognitive performance on math (but not reading) beyond 26°C (78.8°F). Notably, these results obtain despite quite high levels of air conditioning penetration in our study region, suggesting that in the short run, individuals do not completely insulate themselves from climatic factors. In contrast, our long-run analysis reveals a noisy and significantly much smaller relationship between climate and human capital than that suggested by the short-run estimates.

This set of results is important for several reasons. Our short-run results indicate that analytical thinking is compromised at modest temperatures well below our popular conventions regarding a very hot day. Cognitive performance of this sort is the lifeblood of *homo economicus* and critical for decision making in a wide range of domains. That this temperature range is a regular occurrence in summer across much of

28. An analysis of our short-run estimates based on historical climate is also consistent with adaptation. Appendix table 2 shows that the detrimental effect of warm weather on test performance appears larger in cooler counties than warmer ones, although the estimates are not statistically different from each other.

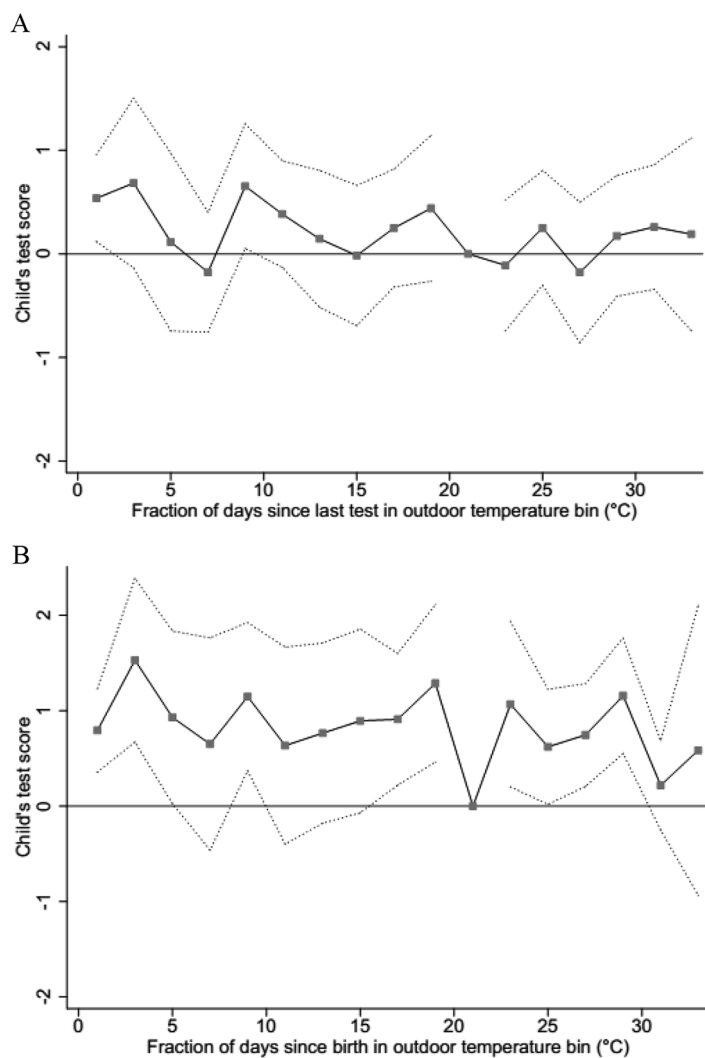


Figure 4. Relationship between long-run temperature and math performance. *A*, Temperature exposure between successive tests. *B*, Temperature exposure from birth until time of test. The solid line shows coefficient estimates with 95% confidence intervals in dotted lines. Panel *A* focuses on measures of temperature between successive tests. Panel *B* focuses on measures of temperature from birth until the time of the test. The regressions include indicators for the fraction of days the temperature was in each 2°C bin, linear controls for precipitation, pressure, wind speed, humidity (measured analogously), dummy variables for day of week, month, year, and state by year. Panel *B* includes the full set of controls as used in column 4 of table 6.

the globe and all year long in parts of the tropics portends potentially sizable impacts on economic well-being. These findings also appear to have strong implications for the optimal timing of cognitively demanding tasks, such as financial decision making and significant health choices.

While cognitive performance and decision making may be compromised by warmer weather, our long-run results demonstrate that these insults have no demonstrable effect on human capital attainment in the long run. Since permanent adaptation strategies are largely held fixed in our comparisons across our short- and long-run specifications, we argue that the difference between these results is driven by compensatory behavior. An interesting feature of this behavior is that it requires no knowledge of the “harmful” effects of temperature since it is an *ex post* adaptive strategy. The feedback from poor test performance may be sufficient to induce individuals to increase investments in learning.

It is important to note, however, that there may be an alternative explanation for the absence of a long-run effect. Test scores are a composite measure of knowledge and performance, and the intertemporal dependencies of one on the other are largely unknown. Thus, it is possible that short-run changes in performance simply do not add up to sizable long-run changes in learning. Given our parameter estimates and simulations, the plausibility of this explanation hinges on a near-zero relationship between the two. Additional research is needed to disentangle the precise mechanisms that underlie the differences between our short- and long-run results and the degree to which they are replicable in other settings.

REFERENCES

- Auffhammer, Maximilian, and Anin Aroonruengsawat. 2011. Simulating the impacts of climate change, prices and population on California’s residential electricity consumption. *Climatic Change* 109 (1): 191–210.
- Baker, Paula C., and Frank Mott. 1989. *NLSY child handbook 1989: A guide and resource document for the National Longitudinal Survey of Youth 1986 Child Data*. Columbus: Ohio State University, Center for Human Resource Research.
- Barreca, Alan, Karen Clay, Olivier Deschênes, Michael Greenstone, and Joseph S. Shapiro. 2013. Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the 20th century. NBER Working Paper 18692, National Bureau of Economic Research, Cambridge, MA.
- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. Why the apple doesn’t fall far: Understanding the intergenerational transmission of education. *American Economic Review* 95:437–49.
- . 2007. From the cradle to the labor market? The effect of birth weight on adult outcomes. *Quarterly Journal of Economics* 122 (1): 409–39.
- Bowler, K., and R. Tirri. 1974. The temperature characteristics of synaptic membrane ATPases from immature and adult rat brain. *Journal of Neurochemistry* 23 (3): 611–13.
- Burke, Marshall, and Kyle Emerick. 2012. Adaptation to climate change: Evidence from US agriculture. Available at SSRN 2144928.

- Burke, Marshall B., Edward Miguel, Shanker Satyanath, John A. Dykema, and David B. Lobell. 2009. Warming increases the risk of civil war in Africa. *Proceedings of the National Academy of Sciences* 106 (49): 20670–74.
- Deboer, Tom. 1998. Brain temperature dependent changes in the electroencephalogram power spectrum of humans and animals. *Journal of Sleep Research* 7 (4): 254–62.
- 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 (3): 66–95.
- . 2014. What do we learn from the weather? The new climate-economy literature. *Journal of Economic Literature* 52 (3): 740–98.
- Deschênes, Olivier, and Michael Greenstone. 2007. The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. *American Economic Review* 97:354–85.
- . 2011. Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US. *American Economic Journal: Applied Economics* 3:152–85.
- Deschênes, Olivier, and Enrico Moretti. 2009. Extreme weather events, mortality, and migration. *Review of Economics and Statistics* 91 (4): 659–81.
- Doya, Kenji, Shin Ishii, Alexandre Pouget, and Rajesh P. N. Rao. 2007. Bayesian brain: Probabilistic approaches to neural coding (computational neuroscience). In *Bayesian brain: Probabilistic approaches to neural coding*, ed. Kenji Doya. Cambridge, MA: MIT Press.
- Graff Zivin, Joshua, and Matthew Neidell. 2013. Environment, health, and human capital. *Journal of Economic Literature* 51 (3): 689–730.
- . 2014. Temperature and the allocation of time: Implications for climate change. *Journal of Labor Economics* 32:1–26.
- Feng, Shuaizhang, Alan B. Krueger, and Michael Oppenheimer. 2010. Linkages among climate change, crop yields and Mexico–US cross-border migration. *Proceedings of the National Academy of Sciences* 107 (32): 14257–62.
- Fine, Bernard J., and John L. Kobrick. 1978. Effects of altitude and heat on complex cognitive tasks. *Human Factors: Journal of the Human Factors and Ergonomics Society* 20 (1): 115–22.
- Froom, Paul, Yeheskial Caine, Igal Shochat, and Joseph Ribak. 1993. Heat stress and helicopter pilot errors. *Journal of Occupational and Environmental Medicine* 35 (7): 720–32.
- Hocking, Chris, Richard B. Silberstein, Wai Man Lau, Con Stough, and Warren Roberts. 2001. Evaluation of cognitive performance in the heat by functional brain imaging and psychometric testing. *Comparative Biochemistry and Physiology Part A: Molecular and Integrative Physiology* 128 (4): 719–34.
- 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 (35): 15367–72.
- Hyde, Dale, John R. Thomas, John Schrot, and W. F. Taylor. 1997. *Quantification of special operations mission-related performance: Operational evaluation of physical measures*. No. NMRI-97-01. Bethesda, MD: Naval Medical Research Institute.
- IPCC. 2007. *Climate change 2007: The physical science basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, ed. S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller. Cambridge: Cambridge University Press.
- Kiyatkin, Eugene A. 2007. Brain temperature fluctuations during physiological and pathological conditions. *European Journal of Applied Physiology* 101 (1): 3–17.

- Libecap, Gary D., and Richard H. Steckel, eds. 2011. *The economics of climate change: Adaptations past and present*. Chicago: University of Chicago Press.
- Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw. 1994. The impact of global warming on agriculture: A Ricardian analysis. *American Economic Review* 84:753–71.
- Moore, Christopher I., and Rosa Cao. 2008. The hemo-neural hypothesis: On the role of blood flow in information processing. *Journal of Neurophysiology* 99 (5): 2035–47.
- Nelson, Richard R., and Edmund S. Phelps. 1966. Investment in humans, technological diffusion, and economic growth. *American Economic Review* 56:69–75.
- Nordhaus, William D. 2010. The economics of hurricanes and implications of global warming. *Climate Change Economics* 1 (1): 1–20.
- Nybo, Lars, and Niels H. Secher. 2004. Cerebral perturbations provoked by prolonged exercise. *Progress in Neurobiology* 72 (4): 223–61.
- Raichle, Marcus E., and Mark A. Mintun. 2006. Brain work and brain imaging. *Annual Review of Neuroscience* 29:449–76.
- Rodgers, Joseph Lee, David C. Rowe, and Kim May. 1994. DF analysis of NLSY IQ/achievement data: Nonshared environmental influences. *Intelligence* 19 (2): 157–77.
- Romer, Paul M. 1986. Increasing returns and long-run growth. *Journal of Political Economy* 94:1002–37.
- Schiff, Steven J., and George G. Somjen. 1985. The effects of temperature on synaptic transmission in hippocampal tissue slices. *Brain Research* 345 (2): 279–84.
- Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher. 2006. The impact of global warming on US agriculture: An econometric analysis of optimal growing conditions. *Review of Economics and Statistics* 88 (1): 113–25.
- Schlenker, Wolfram, and Michael J. Roberts. 2009. Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of Sciences* 106 (37): 15594–98.
- Sinha, Paramita, and Maureen L. Cropper. 2013. The value of climate amenities: Evidence from US migration decisions. NBER Working Paper 18756, National Bureau of Economic Research, Cambridge, MA.
- Todd, Petra E., and Kenneth I. Wolpin. 2007. The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital* 1 (1): 91–136.
- Varshney, Lav R. 2011. Performance of LDPC codes under faulty iterative decoding. *Information Theory, IEEE Transactions on* 57 (7): 4427–44.
- Vasmatazidis, Ioannis, Robert E. Schlegel, and Peter A. Hancock. 2002. An investigation of heat stress effects on time-sharing performance. *Ergonomics* 45 (3): 218–39.
- Yablonskiy, Dmitriy A., Joseph J. H. Ackerman, and Marcus E. Raichle. 2000. Coupling between changes in human brain temperature and oxidative metabolism during prolonged visual stimulation. *Proceedings of the National Academy of Sciences* 97 (13): 7603–8.