Hot Days, the Ability to Work and Climate Resilience: Evidence from a Representative Sample of 42,152 Indian Households*

Anthony Heyes
University of Ottawa
University of Exeter

Soodeh Saberian
University of Manitoba

November 7, 2021

Abstract

The ability of people to work underpins most economic outcomes. Using data from the nationally representative India Human Development Survey (IHDS-II), with pre-scheduling of interview locations ensuring plausibly random assignment of temperature treatment to respondent, we evidence the impact of short-term (within-month) high temperatures on self-evaluated ability to work, and how that impact depends on individual living conditions. Other things equal a hot day (one in which maximum daytime temperature exceeds 37.7°C (100°F)) increases inability to work across the month by about 7%, or 1/20th of a day. Electricity to the home and cooler ownership have important but partial protective effects, we find no such evidence for piped water supply. **Keywords:** Temperature, effective labor supply, climate impacts, climate resilience, mitigation.

*Heyes is a Tier 1 Canada Research Chair (CRC) in Environmental Economics. Heyes acknowledge financial support for this project from the CRC and from SSHRC under Insight Grant project #435-2020-0100 entitled "Heat and the Indian Economy". We are grateful to John Haisken-Denew, Sandeep Kapur, Myra Yazbeck, Richard Tol, Alan Winters, Dana Andersen, Nicolas Ziebarth, Ian Mackenzie, David Byrne, Quentin Grafton and seminar participants at University of Melbourne, Australian National University, University of Queensland, University of Manitoba, University of Alberta and University of Sussex, two referees and Editor of this journal Andrew Foster for very constructive feedback and advice. Errors are ours.

1 Introduction

The extent to which high outdoor temperatures impact the ability to work - and how far those impacts can be mitigated by appropriate use of technology - are central building blocks in understanding the link from climate to individual and societal economic outcomes.[1]

Important recent research has explored some of the links between temperature and economic performance at various levels including individual (Graff Zivin and Neidell (2014)), plant (Zhang et al. (2018)) and jurisdiction (Dell et al. (2012), Deryugina and Hsiang (2014)). In a highly cited paper Dell et al. (2012), for example, show that high annualized temperatures negatively impact both output levels and rates of output growth at national level. Hsiang (2010) relates hot days in a year to annualized measures of value added in various economic sectors in 28 Caribbean and Central American countries, and speculates on the potentially important role played by compromised labor productivity as a mechanism.

A long-standing strand of research in ergonomics relates the temperature to which an individual is contemporaneously exposed to various measures of productivity and performance, including that of employees and students. As early as 1946 research showed that telegraph operators made more mistakes in high temperatures (Mackworth (1946)). Parsons (2007) provides a good introduction to some of this earlier research and Pilcher et al. (2002) a meta-analytic review of the ergonomics literature.

The contribution closest to ours in spirit is Somanathan et al. (2021) who find mixed effects of short-term outdoor temperature on recorded absence in a series of case studies. Weekly heat increases absence slightly in eight workshops belonging to a garment manufacturer in Delhi and among 198 workers at a single iron rail production plant in Bhilhai.[2] However they find no such effect among 147 workers in three weaving workshops in Gujarat. Outdoor temperature also reduces productivity of those workers who are in attendance across most settings. Studying garment factories in Bangalore, Adhvaryu et al. (2020) find that the installation of LED lights in a workshop - which

[1] The background against which we conduct our research is a world in which both average temperatures, and frequency of very hot days, are expected to increase (Rohini et al. (2016), Stern et al. (2006)). According to the Hadley (HadCM3) climate model, for example, under a business-as-usual scenario the number of days when daily average temperature in India exceeds 95 degrees Fahrenheit will increase from five in the period 1957 to 2000 to seventy-five in an average year between 2075 and 2100. The ability of people to work underpins most of what we study in economics - from production, to consumption, to government revenue, and all of the things that flow from those.

[2] The evidence from the garment setting are somewhat weak, with significance in the main linear specification obtained at the 10% level. The main focus of their paper is on worker productivity at the intensive margin (i.e. once a worker is at the workplace), and consistent with previous studies they find such effects in most of the workplaces that they study. As they working with administrative data derived from the employer they do not speak to how temperature-sensitivity varies with living circumstances.

in addition to providing light also generates less heat than the compact fluorescent lamps that they replace - raises productivity on hot days.

While delivering numerous insights Somanathan et al. (2021), like other contributions, observe only a comparatively small sample of workers, engaged in a narrowly defined set of work tasks in a small number of work locations, settings chosen on availability of data and an employer willing to share. The extent to which insights can be extrapolated from the patchwork of niche settings to the broader economy is unclear. One important attempt to develop understanding across the broader workforce Graff Zivin and Neidell (2014) use diary-based data from the US unusual heat can reduce hours worked, particularly in outdoor occupations.

In this paper we take this line of inquiry forward, examining the self-evaluated ability to work of a large sample of workers dispersed across a very wide set of types of occupations and locations.

Our measure of inability to work is derived from the India Human Development Survey (IHDS-II). The IHDS-II, a collaboration between the University of Maryland and the National Council of Applied Economic Research (NCAER) in New Delhi, is a nationally representative survey of 42,152 households in 1503 villages and 971 urban neighborhoods dispersed widely across India. The survey asks detailed questions about household members and their living conditions. Respondents that work are employed or self-employed in a wide array of jobs.

The question from which we derive our independent variable asks respondents to report on how many of the 30 days prior to date of interview they were "unable to work or carry out normal duties". We relate the answer to what we know about temperature in the vicinity of the respondent's home address in that same 30 day period. Obviously there is no single "correct way" to capture heat exposure over a month in a single or small number of regressors. We use several including frequency of 'hot days', defined to be a day in which the maximum temperature reached 37.7°C (which equals 100°F), the moving average of daily maximum temperatures, and counts of the number of days in which maximum temperature falls into each of a series of bins 5°C in width.[3]

The 59,621 workers in our central sample experience between them 466,757 hot days, defined against the 37.7°C threshold, in the 30 day windows before their respective interviews. While this is around 8.4 days on average there is much variation

[3]While the multiple bins approach provides more dimension in terms of main results we prefer the count of hot days for reasons of tractability and the ease with which it allows us to conduct the various subsample analyses that constitute the second half of the paper. It should come as little surprise that the various measures of heat are strongly positively correlated. In robustness exercises we also report results based on; (a) alternative definitions of a hot day (higher and lower thresholds) and, (b) the Heat Index, and algorithm developed by the US National Weather Service that combines temperature and humidity into a single number that captures the impact of heat on the human body. Further discussion of heat metrics in our setting will be presented later.

between individuals across both space and time. Controlling for location and time of year, random assignment of temperature treatment to respondent is implied by the fieldwork approach. We additionally report results that confirm that treatment is not related systematically to numerous observable household or individual characteristics. In our base specifications we find that (1) a hot day causes a roughly 7% increase in inability to work; (2) other things equal the effect is more pronounced in older workers and females, (3) while the effects vary across type of employment, with largest effects in construction, which involves predominantly outdoor and physically-demanding work, we still find evidence suggestive of an effect on office-based jobs.

The nationally-representative character of the sample surveyed allows us to explore heterogeneity of effect not just across job categories, but also geographically, something not possible in a case study. We find that the effect of an additional hot day is largest in typically cooler locations, and find no significant impact in places where such hot days are common, consistent with worker adaptation to local climate. However, the effects are observed in places subject to each of three of the major climate zones, namely Arid, Humid and Subtropical, pointing to a broad geographic basis for the effect found in the whole sample.

A second important advantage of our setting is that we have very detailed information about the conditions in which individual respondents live.[4] While some studies (including Somanathan et al. (2021), Adhvaryu et al (2020)) explore the role that cooling technology in the *workplace* might play in moderating the temperature-productivity and temperature-labor supply relations, the IHDS-II data allows us to generate what we believe to be the first systematic evidence on the protective role of conditions at *home*. We find that a household being connected to the electricity grid has an important causal role in reducing its susceptibility to heat, as does ownership of cooling technology. Surprisingly we find no evidence that the quality of water supply to the house playing a significant mitigative role. This sort of evidence is important in assessing the contribution of public investment in domestic water and power infrastructure in India and elsewhere to enhanced climate-resilience.

Before proceeding to the empirical analysis it is worth thinking explicitly about how we should interpret the results. The outcome variable, a self-assessment by each respondent of the frequency with which they were unable to perform normal duties over a period, is a novel one. One advantage to this is that the interviewee who physically attended work, or went to his self-owned workshop, but still felt unable to achieve much meaningful, is able to express that. However the 'normal duties' element provides a challenge in interpretation. Is the productivity during a day when

[4]It is plausible that inability to work through physiological, for example through illness or loss of sleep, is particularly sensitive to living conditions.

a respondent does not perform normal duties reduced to zero? By 50%? Or by some other percentage? in this sense the deliberately open-ended nature of the question IHDS-II poses is both a strength and a weakness, compared to conventional metrics, for example a dichotomous measure of physical absence as recorded in a personnel file. That it is *different*, however, provides another way in which the results here can complement research using other, for example administrative, measures.

The remainder of the paper is laid out as follows. In Section 2 we outline data sources and methods. Section 3 presents our central results. Section 4 lays out results relating to the protective effects of electricity, water and air-cooling. Section 5 reports the results of a variety of robustness tests and falsification exercises. Section 6 concludes.

2 Data

Our central analysis links data on number of days an individual was unable to work with what we know about weather conditions at the location of each individual on the date in question.

2.1 India Human Development Survey

Our main data comes from the India Human Development Survey-II (IHDS-II) in 2011-12, which covers 204,569 individuals living in 42,152 households located in 1,503 villages and 971 urban neighborhoods spread across India. The IHDS includes economic, social, and health modules and has been used by researchers to explore topics such as human capital accumulation (Azam et al. (2013)), maternal effects of socio-economic status (Mohanty and Gebremedhin (2018)) and the nutritional effects of migration (Atkin (2016)).

For a detailed description of the sampling and interview methods the reader is encouraged to consult Data Sharing for Demographic Research (DSDR).[5] The sample is designed to be nationally representative along a number of dimensions (geographic, religious, ethnolinguistic, etc.). Households are selected according to standardized sampling criteria within a randomly-selected set of villages and urban neighborhoods. The interviews are conducted by researchers from the National Council of Applied Economic Research (NCAER) and the University of Maryland. The quality of data is high by standard metrics (Wang et al. (2014)). Travel schedules between of the interview teams between locations are planned in advance and the order of interviews

[5] A data guide on IHDS-II can be found at `https://www.icpsr.umich.edu/icpsrweb/content/DSDR/idhs-II-data-guide.html`.

within village or urban neighborhood is randomized. Interviews were conducted by a number of enumeration teams between January 2011 and December 2012, with most in the later year.

Our dependent variable throughout is the number of days in the past 30 days that an individual reports being unable to work or carry out normal duties prior to the date of interview (Question 8.11 on Page 11 on questionnaire). 97.7% of respondents provided an answer to this question. Anyone with no response recorded was dropped. In column 1 of Table A.1 we show that we are unable to reject the null hypothesis that absence of response is uncorrelated with the temperature variable.

The usual caution with regard to self-reported data applies. It is likely that respondents vary in how they interpret "unable to work". Provided the way in which they interpret the question is not itself sensitive to the treatment (temperature) then the measurement error introduced by this is unlikely to bias estimates. Insofar as any respondents regard 'normal duties' as dependent on temperature, for example somebody whose normal duties require cool conditions, this could complicate inference. Later in the paper we conduct an exercise that probes the extent to which this is a substantive problem.

Note that our is different from *absence*, which we regard as a strength. The interviewee who attended work, or went to his self-owned workshop, but still felt unable to achieve anything meaningful, is able to express that. A further point to note is that the dependent variable is measured over a 30 day period. One advantage of this is that it allows us to absorb most within-month lagged effects. If some hot days cause a subject to be unable to work not just contemporaneously but for a few days after the temperature cools, our measure will pick that up, provided the effect does not spill out of the 30 day window.

The database contains date of interview and location of each household at district level which enable us to merge the IHDS-II data with weather information. As our focus is on work we also dropped subjects who have as primary activity status retired, housework, student or unemployed. We exclude a tiny number of respondents who had an interview that was not completed on a single day. The final data consists of 59,621 'working' individuals.

The IHDS-II is a rich source of controls. For each individuals we collect age, gender, highest education level, religion and type of work. At household-level we know number of household members, and have metrics for income and assets that allow us to control for economic circumstances. In addition we know about household access to water and the reliability of supply, whether the household is connected to mains electricity and if so the hours per day, and ownership of cooling technology. Throughout the paper we use the provided survey weights to make the sample nationally representative.

2.2 Weather

The spatial and temporal coverage of ground-based meteorological stations in India is inadequate for our purposes. We follow Schlenker and Lobell (2010), Schlenker and Roberts (2009), Auffhammer et al. (2013) and others and use reanalysis data from the ERA-Interim archive. This is collated by researchers at the European Centre for Medium-Term Weather Forecasting (ECMWF). Concretely we collect the ERA-Interim calendar daily maximum temperature, minimum temperature, precipitation, relative humidity and solar radiation data for each cell in a 1° x 1° latitude-longitude grid for each day from 1 January 2011 to 31 December 2012 (this is the spatial level at which this data is presented).

We do not know the exact home address of a respondent but, rather, the district in which they live. Districts are the jurisdictional sub-subdivision in India. The country is divided into 29 states each of which is divided into about 25 districts (723 districts in totla). District boundaries do not correspond to the latitudinal-longitudinal grid cells. To generate weather variables for each district, we construct an inverse-distance weighted average of all the weather grid points within 20-miles of the centroid of that district. On this basis we construct a district-by-day panel of maximum temperature, precipitation, average wind speed, solar radiation and relative humidity.

As noted we use three main ways to capture monthly heat exposure at a location. (1) The number of days in district d in the 30 days prior to interview date t that were hot, *i.e.* daily maximum temperature exceeded 37.7°C (which equals 100°F). This follows Deschênes and Greenstone (2011), Cohen and Dechezleprêtre (2021) and others. (2) The 30 day mean of recorded daily maximum temperatures. (3) The count of days in the 30 day window that daily maximum temperature fell into each of a series of 'bins'. Summary statistics are presented in Table 1. Figure 1 shows the variation in treatment across respondents.

2.3 Methods

In exploring the relation between temperature and inability to work we will adopt three main approaches. Our main specifications incorporate a comprehensive suite of fixed effects.

First, we estimate the following:

$$y_{itd} = \beta_0 + \beta_1 temp_{td} + W_{td}\beta_2 + Z_i\beta_3 + H_i\beta_4 + F_i\beta_4 + \Psi_{td} + \theta_t + \epsilon_{itd} \qquad (1)$$

where y_{itd} is number of days that individual i living in district d reported being unable to work in the 30 days prior to interview date t.

With a count variable as dependent variable we estimate using the Poisson Pseudo-Maximum Likelihood (PPML) regression with high-dimensional fixed effects model developed for application in Stata by Correia et al. (2020). This method refines Poisson estimation for settings that include multiple high-dimensional fixed effects.

The dependent variable includes frequent zero values - around 80% of total responses - people reporting no days in the previous 30 when they were unable to work. As a robustness exercise we estimate the preferred specification using the zero-inflated-Poisson (ZIP) model and find that it delivers qualitatively similar results, though somewhat larger implied effect sizes.

The variable of interest is the measure of temperature that following Deschênes and Greenstone (2011) and Cohen and Dechezleprêtre (2021) is constructed as the number of days in the treatment window, the 30 days prior to interview date t, that the maximum temperature exceeded 37.7°C (equals 100°F). As such $temp_{td}$ is a count variable that takes an integer value between zero and 30 inclusive. β_1 is our coefficient of interest.

To allow for the possibility that dimensions of weather other than temperature might impact ability to work we include a vector of non-temperature weather controls, W_{td}. It contains average daily precipitation, wind speed, solar radiation and relative humidity in the same 30 day period.

The count temperature measure that we use also has many zero values. In other words respondents who had been treated to no hot days in the 30 prior to interview. While our main results are estimated on the whole sample, we re-estimate the preferred specification on the 'intensive margin', by which we mean only those respondents treated by at least 1 hot day in the treatment window.

Z_i contains a set of controls for individual respondent characteristics. In particular age, gender, highest education level and religion.

The vector H_i includes controls for household characteristics, in particular the number of household members, and whether the household has access to electricity and piped water.

The vector F_i contains controls for household financial status, specifically for household income and for wealth (total household assets).

The equation also includes spatial and temporal fixed effects. Ψ_{td} contains district-by-month fixed effects that absorb time-varying differences in the dependent variable common across households within a district. For example, local holidays. The vector θ_t includes year fixed effects to control for any trends common across states that might exist in the data.

ϵ_{itd} is the error term. In our preferred specification standard errors are clustered by district to account for spatial correlation within districts, consistent with Abadie et al.

(2017). Later in the paper we establish that our preferred approach is conservative and that inference is qualitatively robust to a variety of other ways of calculating standard errors, including methods that adjust explicitly for spatial correlation.

Second, we estimate equation (1) but with the temperature regressor replaced by the mean of the daily high temperatures obtained in the 30 day window prior to interview.

Third, following Burgess et al. (2017) and others we conduct a non-parametric exercise. We estimate the following flexible model using PPML:

$$y_{itd} = \beta_0 + \sum_{j=1}^{7} \beta_j temp_{td} + W_{td}\beta_2 + Z_i\beta_3 + H_i\beta_4 + F_i\beta_4 + \Psi_{td} + \theta_t + \eta_s + \epsilon_{itd} \tag{2}$$

where the variable $temp_{itd}$ denotes the number of days in district d in the 30 days prior to interview date t that the daily maximum temperature fell in the jth of of seven temperature 'bins'. The bottom and top bins capture days on which maximum temperature was less than 15°C and more than 40°C respectively, with the five bins in-between each of width 5°C. We then estimate separate coefficients β_j for each of these temperature bins using the bin for temperatures between 20 to 25°C as the reference category.

Throughout the analysis the identifying assumption underpinning causal inference is that having controlled for location and time fixed effects, the realized value of the temperature regressor is subject to variation that is as good as random.

With respect to locational sorting, it is sensible to think that location is endogenous - where individuals choose to live may depend on local climate. For example, an individual who knows themselves to be particularly susceptible to heat might choose to live in a place where hot days occur less frequently.[6] However, that our design exploits variations in realized temperatures *conditional* on location, means that this margin of adjustment is already accounted for. What we identify are the impacts of temperature after any such locational adjustment is accounted for. In other words, if everyone in our sample has relocated to a place where the distribution of temperature that suits their particular preferences and characteristics, our analysis then explores whether there remains an effect of high temperature realizations after such locational adaptation.

[6]Jessoe et al. (2018) are among the studies that show how extreme heat can induce both internal (rural to urban) and international migration, in their case among Mexicans.

3 Results

3.1 Main

The main, whole-sample results are reported in Tables 2 through 4.

Table 2 uses as temperature regressor the count of hot days. Column 1 reports the sparsest specification including state, year and district-month fixed effects. The estimated coefficient on the temperature regressor is 0.064, which is statistically significant at a level much higher than 1%.

Moving right-wards across the table we include, in turn, controls for individual characteristics (column 2); individual and household characteristics (column 3); individual, household and financial characteristics (column 4). The inclusions have no substantive impact on the value our coefficient of interest, though the precision of the estimates is improved.

In column 5 we return to the sparse specification but include the vector of non-temperature weather covariates. Again results are not overly disturbed.

In column 6 we include all of the controls above in addition to a vector of type of job FEs. The later will capture systematic variation in inability to work across different types of work. There are 54 job descriptors in the IHDS-II (for example: nurse, shop-worker etc.). Data from the survey is incomplete with regard to job for about 20,000 respondents such that column 6 is estimated on a correspondingly smaller sample. The overall effect of including these controls is to decrease the estimated coefficient of interest but only slightly.

Column 7 is the preferred specification in this table. It is the same as that in column 6 but excludes the type of job controls. The designation of this specification as the preferred is pragmatic, since it allows us to return to the full sample of 59,621 without the coefficient estimate of interest being much disturbed. While retaining sample size is not crucial for the purposes of this table, much of the rest of the paper reports various sub-sample exercises, and maintaining the larger sample will allow us maintain power within smaller sub-categories.

The coefficient in column 7 is 0.071, significant at a level much higher than 1%. The outcome variable and explanatory variable of interest are defined over the window of 30 days before interview. One extra day increases days of inability to work by 7.3%, since $\exp(0.071) = 1.073$. The mean number of days unable to work in the estimating sample is 0.63 such that one extra hot day implies an increase of 0.046 days.

Column 8 of Table 2 reports the outcome of re-estimating the preferred specification at the intensive margin, by which we mean only on respondents who had been treated by at least one hot day. The estimated coefficient is larger in value than that estimated on the whole sample, at 0.093.

The results in Table 2 rely on Pseudo Poisson estimation using techniques developed by Correia et al. (2020) to accommodate high-dimensional fixed effects. For completeness in Table A.2 we re-estimate the preferred specification using OLS, negative binomial and zero-inflated Poisson (ZIP) methods. As can be seen results are qualitatively robust.

Table 3 reports the same set of specifications as Table 2 but with the count of hot days replaced by a simple 30 day mean of daily maximum temperatures. This alternative temperature metric follows Park et al. (2020), for example, who use this (calculated over a year) metric for cumulative heat exposure. Results are consistent in sign and significance with Table 2, though significance at conventional levels is not quite achieved in the restricted sample in column 6. The estimated coefficient in column 7 implies that a 1°C increase in average daily maximum across the 30 day window causes a 13.3% increase in days unable to work (since $\exp(0.125) = 1.133$), which corresponds to 0.71 days for the mean respondent.

Table 4 summarizes analogous specifications but relates to Equation 2, incorporating temperature as the count of days in which maximum temperature fell into each of a series of bins as regressors (20 to 25°C as reference category). The pattern of coefficients is consistent across columns. The coefficients from column 6 and presented graphically in Figure 2. The close-to-linear trend is confirmed. This validates our focus on the linear case in most of the paper. There is suggestive visual evidence pointing to 'flattening out' of the effect towards the right-hand side of Figure 2, though that evidence is weak, with little mass of data in the < 15°C bin delivering a very wide confidence interval.[7]

The various results here reflect that there is no single "best" way to measure and incorporate heat exposure in an analysis like this. It is unsurprising that all of the various measures are strongly correlated, and the results presented in Tables 2 through 4 complement one another. We prefer the simple count of hot days - which binarizes each day to either being hot or not - as it facilitates the various subsample analyses which constitute the rest of the paper. Naturally the multiple-bins approach just presented gives more nuanced main results, but is less tractable for future purposes. As such among the results in these three tables the specification in column 7 of Table 2 is the preferred.

[7]Furthermore, inserting a $temp_{itd}-$squared variable into an otherwise unchanged Equation 1 delivers a very small and statistically insignificant associated coefficient and disturbs other coefficients only minutely. Results of this exercise are not reported.

3.2 Geography

India is a large country with diverse local climates, so respondents vary in the weather conditions that they usually experience. All of our specifications include district fixed effects, as such we control for any time invariant characteristics of a respondent's location, including climate. However, one advantage of using a large-scale survey conducted across a widely-dispersed set of locations is that it allows us to probe how effects vary according to typical local conditions.

We do this in two different ways, reporting the results in Table 5.

First we divide the sample into locations that vary into how unusual hot days, as per our, are. More precisely we count the number of such days across a 36 month period (January 2010 through December 2012). Columns) through 4 in Table 5 summarize the main coefficients derived from estimating the preferred specification but on subsamples. Column 5 estimates on interviews executed at locations where hot days are comparatively common, more than 100 per annum. Column 4 where there are 60 to 100, which contains most of the sample, and column 2 those places where hot days are rarer (20 to 60 per annum).

The sample sizes vary substantially across the columns, so care is needed in interpretation, but it can be seen that the estimated coefficient is largest in column 2, smallest and far short of statistical significance at conventional levels in column 4. An extra hot day in the month at a location in column 2 increases number of days unable to work by 32.7%, whereas in column 3 that number is 6.4%.

These results suggest a degree of adaptation in the population.

To complement this we allocate respondent locations according to which of the four major climate zones they fall into, namely Arid, Montana, Humid and Subtropical. Columns 5 through 8 in Table 5 report the results of estimating the preferred specification on subsamples drawn from each of the climate zones in sequence. As can be seen we find significant positive effects in each case, except for Montana. The latter result should not be over-interpreted given the small share of the population of India, and hence the IHDS sample, live under such a climate. The former results show that the phenomenon studied in this paper - the heat sensitivity of ability to work - has a broad geographical base across the varying local climates that Indian people inhabit.

3.3 Age

Previous research points to the impact of high temperatures on mortality (Deschênes and Greenstone (2011) and Deschenes (2018)), morbidity (Barreca and Shimshack (2012) and Liss et al. (2017)), mental function (Dai et al. (2016)) being greater on older adults. Here we explore whether the effect of hot days on inability to work is similarly sensitive to age.

To do this we divide our sample into three bins, designed to contain workers who are young (under 20 years), prime-age (20 to 60 years) and old (older than 60 years).[8] Column 2 in Table 6 reports the result of re-estimating the preferred specification but with the addition of regressors interacting the temperature variable with dummies taking the value 1 if a respondent falls into each age category bin. The middle bin is the omitted category. The coefficient estimate on the bin containing older respondents, 0.023***, points to older workers being substantially more sensitive to hot temperatures in terms of self-reported ability to work.

3.4 Sex

Previous research also points to the impact of high temperatures on mortality (Deschenes (2018)), morbidity (Barreca and Shimshack (2012)), mental function (Heyes and Saberian (2019)) being greater on females than males.

In addition note that our outcome variable will pick-up a variety of reasons for inability to work. While that includes directs impact on the respondent him or herself we conjecture that another important mechanism might work via needs of a dependent child. If a child is unable to attend school (say because of temperature-induced sickness) and requires supervision at home it is plausibly the mother who is more likely to be unable to work in order to respond to that familial need.

To probe this in our setting we re-estimate the preferred specification but adding a regressor interacting the temperature variable with a a dummy that takes the value 1 if the respondent is female, zero otherwise. The result of this exercise is reported in column 3 in Table 6. It can be seen that the coefficient on the interaction term is positive and significant at a level higher than 1%. Being female increases the treatment effect by around a third.

3.5 Workplace

We explore how heat effects ability to work varies across job category and/or work task. We do this by adding to the preferred specification, one at a time, a series of regressors that interact the temperature measure with dummies that take the value 1 if the job of a respondent falls into a particular category, zero otherwise.

IHDS provides fine-grained information on job types in 99 categories. For the purposes of constructing adequately-populated subsamples these divisions are too narrow.

[8] Recall that throughout the study we are limiting attention to respondents with jobs, so while the first sub-sample does include some children (under 18, the age of majority in India) these are largely young adults.

We construct three broader categories that we call "construction" (to include job titles such as construction laborer, carpenter, plumbers, stone cutter, etc.), "agriculture" and "office" (clerical assistant, office administrator, etc.). Our categories are not exhaustive, containing 24,332 (58.3% of those who report job title).

Columns 4 and 5 in Table 6 report the results of the interaction exercise with respect to respondents who work in agriculture and construction respectively. These frequently imply work that is both outdoor and physically-demanding in character. Column 6 isolates respondents in work tasks characterized as "office". Perhaps surprisingly the coefficient is small and not statistically significant, such that we find no evidence that respondents who work in agriculture are any more or less sensitive to heat than non-agricultural respondents. However, consistent with expectations, respondents working in construction are significantly more sensitive, and those in office-based jobs significantly less sensitive than the wider sample.

Columns 7 and 8 report two further exercises. While the primary mechanisms that we have in mind linking extreme heat to self-evaluated ability to work are physiological in character, it is also possible that some of the effect could driven either by, (a) some work tasks require electricity for their execution, and extreme heat makes electricity shortage more likely or, (b) heat causing there to be less work to be available. While nothing in the data allows us to assess these directly we consider two sub-groups of worker who may provide suggestive evidence for against such causal channels. First, we construct a set of job types that we judge unlikely to require electricity to execute (drivers, trawler-men, etc..) and define a dummy variable that takes the value 1 if the respondent works in one of those jobs, zero otherwise. Second, to investigate (b), we construct a set of jobs that we judge work supply to be unlikely to be reduced in event of heat (typists, mail delivery workers, etc..). We then re-estimate the preferred specification but adding a regressor that interacts the temperature measure with each of these dummy variables in turn. The categories that we construct contain 3,652 and 3,987 respondents respectively. The coefficients on each interaction term can be seen to be very small in value and in neither case come close to statistical significance at conventional levels, providing no evidence in favour of these alternative mechanisms. We acknowledge however that we are unable to pin down mechanisms persuasively in our setting, and leave that for future research.

The substantial differences in effect size between categories points to sharply different effects of hot weather on effective supply of work in different sectors and work tasks.

4 Climate-resilience: The protective effects of electricity, cooling and water

While the central ambition of climate policy remains to influence the path of climate (for example, reducing the rate of increase of average temperatures and extremely hot days in places like India), a second strand is focussed on adaptation. Key to policy formulation is understanding measures that increase resilience in the face of evolving weather patterns (Barreca et al. (2016)). This is particularly important in poorer countries, where the health and other impacts of increasing frequency of hot days is expected to be most pronounced (Hansen et al. (2012)).

A major strength of our setting is that we know a lot about the living conditions of individual respondents. We exploit this here. Some existing studies explore the role of protective technology, such as air-conditioning, in the workplace itself, but none (to the best of our knowledge) had the data needed to examine the role of home circumstances. As such are results complement earlier research that look, for example, at absence at workplace level where nothing is known about the home conditions of the worker population (Somanathan et al. (2021)).

Having presented evidence of the effects of hot weather on self-reported inability to work, in the rest of the paper we explore the extent to which those effects are moderated by: (a) a household having access to electricity (and, if so, whether the quality of that electricity); (b) ownership of cooling technology and, (c) quality of water supply.

Of course, household living circumstances are not generally randomly assigned, so care is needed in interpreting correlations. However we offer supplementary analyses that will reinforce (cautious) causal interpretation.

We explore these adaptation questions in a series of exercises. In each we add to the preferred specification a regressor that interacts the temperature variable with a binary variable capturing one of the adaptation measures above. The results of these are collated in Table 7.

4.1 Electricity

Household access to electricity is frequently regarded as an important outcome of economic development (Reiss and White (2005)). Access has been shown to affect household-level economic outcomes (Khandker et al. (2013), Burlig and Preonas (2016)), proneness to infection (Barron and Torero (2017)) and other health outcomes (Spalding-Fecher (2005)), health attitudes (Manning et al. (2015)), study habits and educational outcomes for children (Khandker et al. (2013) and Barron and Torero

(2017)), media consumption and social attitudes (Lee et al. (2020)), amongst other things.

With regard to climate change, electricity-access has been argued to offer significant protective effects to household health (Barbier (2014)). This might work through diverse channels. One obvious is through its allowing the use of cooling technology. In a well-known study using decadal, state-level data, Barreca et al. (2016) find that penetration of air-cooling technology predicts long-term decline in the temperature-mortality relationship in the United States across the course of the twentieth century. Electricity also allows for effective refrigeration of food, medicines, etc.. in hot climates.

The IHDS-II reports whether the home of each respondent is connected to the electricity grid. About 14% of our sample was not connected, with that percentage much higher outside urban areas.

In column 2 of Table 7 the binary adaptation variable takes the value 1 if the respondent lives in a household with electricity connection. The negative and significant coefficient on the interaction term, -0.017**, implies an important protective effect of connection.

The main challenge to causal inference here is that electricity connection is not randomly-assigned. In particular, we might expect access to electricity to be greater for 'better off' households - those with higher income and/or assets - implying that what we are picking up here is the protective effect of being well-off rather than the effect of electricity connection *per se*. The primary defence against spurious attribution is the inclusion of financial controls, and the detailed nature of IHDS-II allows us to include fine-grained household controls for income and wealth. However, to address residual concerns that the controls do not adequately isolate electricity-connection from financial circumstances we conduct an additional exercise. Following Graff Zivin and Neidell (2012), Currie and Stabile (2006) and others we repeat the regression reported in column 2 (and each subsequent regression reported in this Section) in Appendix Table 3 but dropping altogether the vector of financial controls. If we have failed to control adequately for financial circumstances in our main specification we would expect dropping this vector to disturb substantially the estimated coefficient on the temperature variable. Although we remain cautious in causal interpretation, that the estimated coefficient on the interaction term is no different in Appendix Table A.3 to that in Table 7 is consistent with our having isolated the causal protective role of connection.

In economically-advanced countries a household being connected to the grid implies that it has more or less continuous access to electricity. This is not true in India, and most other low and middle-income countries, where reliability is a significant challenge and the quality of supply even to those households with a connection can be highly variable. Consistency of supply has been shown to have significant effects on outcomes such as non-agricultural household income (Samad and Zhang (2016), Chakravorty et al. (2014)) and infant health (Lewis (2018)). India has invested heavily in recent years in improving supply quality, not just grid coverage (Chakravorty et al. (2014)), partly with the ambition of increasing the climate-resilience of the economy.

For each household in the IHDS, in addition to knowing whether the household has connection to the grid, we have a self-reported estimate of quality, captured by how many hours in a typical day electricity is available. The binary adaptation variable in the specification in column 3 takes the value 1 if the household in which the respondent resides benefits from electricity supply for at least 12 hours per day (our designation of good quality electricity), zero otherwise. The negative and significant coefficient on the interaction term in column 3 points again to electricity in the home having an important protective benefit.

4.2 Cooling

Much emphasis has been put on the role that cooling equipment might play in protecting individuals and populations against various impacts of high temperature including morbidity (O'Neill et al. (2009)), mortality (Barreca et al. (2016)), learning (Park et al. (2020)) and student exam performance (Park (2017)). Barreca et al. (2016)), for example, show that state-by-decade rates of air-conditioning penetration in the US predict the decline in the excess mortality caused by hot days. We explore the analogous question, but with inability to work as the outcome interest.

In the period of this study (and today) in India technological protection against heat typically involves installation of an air cooler rather than an air conditioner. Air coolers pump air through water, reducing temperature via evaporative cooling, while air-conditioning uses refrigerants to absorb heat. Air coolers are typically cheaper and easier to maintain. In our sample only 1,241 (1.92%) of respondents report ownership of an air conditioner (typically apartment-dwellers in the major cities) whereas that number is 10,730 (16.57%) for air cooler.[9] Given the small number of individuals with air conditioning we amalgamate those who have air coolers or air conditioners (or both) into a single sub-sample ('cooling').

For the purpose of column 4 in Table 7 the binary adaptation variable takes the value 1 if the household reports owning an air cooler or air conditioner, zero otherwise. The coefficient on the interaction term, -0.008** points to a statistically significant but

[9] The variation in ownership is important to us in developing results. This contrasts with the near-universality of air-conditioning in the US. Almost 90% of US homes are treated by air-conditioning, and in the hotter states (Florida, Arizona, etc.) that number is close to 100% (Isaac and Van Vuuren (2009)).

fairly small protective benefit of ownership. Column 5 repeats this exercise, but with the triple interaction between ownership of the treatment variable, cooling technology and location of the household in a hot location (defined as one of the locations in the sample that had a greater than median number of hot days over the 36 month period starting January 2010). The coefficient on that interaction is very similar in size to that in column 4.

Ownership of air cooling technology is clearly not randomly assigned, so care is needed in making causal claims.[10] That coefficient estimates are robust qualitatively when the whole suite of financial controls are dropped from these regressions (see Table A.3) assuages concerns that we have failed to control with sufficient precision for household financial circumstances and reinforces cautious causal interpretation.

In summary we find important mitigative benefit from cooling technology, and also reiterate that the technology observed in this setting is likely much less advanced than state-of-the-art air-conditioning typically found in, for example, the United States.

4.3 Water

Finally we explore the role of water supply in protecting individuals and households against the impacts of extreme temperature. Improving water supply has been a mainstay of development policy for decades (Hunter et al. (2010)). In India an estimated 550 billion USD was spent on enhancing water supply infrastructure between 2006 and 2018, an important motivation for which was as an investment in building climate resilience (Larsen et al. (2016)).

The IHDS-II provides detailed household-level information on water in two different ways: (a) the primary source of water for each household (this is presented in nine categories, but we are going to focus on those with and without the 'gold standard' of piped water supply of the sort almost universal in high income countries); (b) the answer (yes/no) by the respondent when asked whether he or she regards water supply to the household as 'normally adequate'.

Column 6 and 7 of Table 7 summarize the results of two separate exercises relating to water.

First, we divide respondents according to the main source of water for the household in which they live. In particular we divide those who have piped water as their main source from those that do not (the latter use wells, collect rainwater, water from rivers, etc.). The sample splits roughly 41% to 59% between these two categories. The binary

[10] Davis and Gertler (2015) represents perhaps the most careful empirical analysis of the air-conditioning adoption decision. They use micro-data from Mexico to show how purchase of air-conditioners is sensitive to local temperature distributions. The effects we identify should be interpreted as already accounting for households having made cooling equipment ownership decisions appropriate to their own setting, including typical local climate.

adaptation variable embedded in the interaction term in column 6 in Table 7 takes the value 1 if the water source of the household is piped, zero otherwise.

For the specification in column 7 we divide the sample according to whether the respondents self-evaluate their household as having, in general "adequate" water supply. The response is binary (yes/no) and subjective. No priming or guidance is given to respondents as to what 'adequacy' should be taken to mean. That 48,878 respond yes implies that many (indeed most) without piped water regard their supply as adequate, for example. The binary adaptation variable in this column takes the value 1 if the response is yes, no otherwise.

Though the signs of the estimated coefficient values in columns 6 and 7 are negative they do not achieve statistical significance, so we fail to find evidence of a protective benefit of piped water access or self-evaluated water quality.

5 Robustness

In this section we report on several robustness checks and falsification exercises. The appendix contains a number of additional robustness results.

5.1 Alternative heat metrics

In executing this study a number of modeling decisions were made. One important decision had regard to the most appropriate way to specify the temperature variable which has served as our treatment variable of interest. As that single variable is tasked with capturing the temperature across a 30 day period its design is a challenge. With our focus on extreme heat, at the location of each respondent we considered counts of hot and not hot days in the 30 days prior to interview. The average of daily maximum temperatures over the 30 day window was used as an alternative. The former approach follows closely the methods used by Graff Zivin et al. (2018), Blakeslee and Fishman (2018) and Carleton and Hsiang (2016) in their investigations of monthly or quarterly level temperature effects. The latter closer to, for examples, Ranson (2014) and Park et al. (2020), who characterize monthly-average temperature effects.

In columns 3 and 4 in Table 8 we repeat the count variant exercise (Table 2 of our main results) but using alternative thresholds to define a hot day. In column 3 a day is defined as hot if the maximum temperature at any point on that day exceeds 35°C (95°F), in column 4 if it exceeds 40.5°C (105°F). In aggregate our sample of 59,621 individuals experienced on average 5.68 and 3.82 of such days, respectively, in the 30 day window prior to interview. The results as they compare to the central specification are as would be expected. When our definition of a hot day is less (more) demanding, the implied effect on inability to work is correspondingly smaller (bigger).

The role of humidity, in addition to temperature, on various outcomes has been noted in the literature (for example, Baylis (2020) and Heyes and Saberian (2019)). While we control for humidity at the location of interview a suspicion might remain that we have failed properly to capture the independent effects of temperature and air water content. To probe this we conduct a further exercise in which we assign days to bins based not on dry-bulb air temperature but on daily heat index (HI). The heat index combines air temperature and relative humidity according to a non-linear equation developed by the US National Weather Service such as to capture human perception of heat.[11] The HI measure that we develop is a count of the number of days in the 30 days before interview that the HI at the location of interview reached 100°. Reported in column 5 the estimated coefficient implies an impact of a hot day so defined increasing inability to work by 15.8%. Note that this coefficient is not directly comparable to the others in this table, though reassuringly sign and significance are consistent.

Read as a whole the results in this table, and others that we have conducted but not reported, point to a significant and substantial positive relation between temperature and inability to work, which is not conditional on the particular metric for temperature adopted.

5.2 Additional exercises

Table 9 reports the results of several additional exercises, based again on the preferred count variant of the analysis.

Large states In column 2 we report estimating the preferred specification on the 5 largest states by contribution to sample (Uttar Pradesh, Madhya Pradesh, Maharashtra, Karnataka and Rajasthan). With the exception of Karnataka these are more northerly states and supply around 42% of sample. The estimated coefficient here remains positive and strongly significant despite the much smaller sample size. The estimated coefficient is larger than that derived from the whole sample, though not significantly so.

Urban-rural Around a third of Indians are urban dwellers (31.2% in the 2011 census) and that is mirrored in our sample (30.9% urban). Columns 3 and 4 in Table 9 summarize the results of running our preferred specification on the urban and rural subsamples respectively.

[11] For example, a temperature of 90°F combined with 70% relative humidity delivers a heat index of 105. A temperature of 98°F combined with a relative humidity 40% also delivers a heat index of 105, implying that according to this metric the 'feeling' of heat would be deemed equal under those two combinations. It is similar in intent to Humidex, a trade-marked metric developed in Canada and used there and elsewhere.

North-South Columns 5 and 6 report the results of estimating the preferred specification on respondents who live in states that lie wholly to the North or wholly to the South of the Tropic of Cancer, which cuts the country roughly in half in terms of surface area. The estimated coefficient is similar across columns, pointing to the effect identified having a broad-based geography.

Outliers In column 7 we investigate the role of outliers by re-estimating the preferred specification but excluding respondents associated with the top and bottom deciles of treatment (in other words the hottest and coldest deciles) from the estimating sample. The result is little disturbed making consistent with our main estimate not driven overly driven by outliers or extreme values.

Winter extremes Though the most populous areas of India are hot for all or most of the year, an additional concern could be that results are driven by unusually hot days occurring in cooler times of year. To explore this in column 8 of Table 9 we re-estimate the specification in column 1 but excluding the 20.1% of respondents whose interviews fell during the winter months (December, January or February) with no discernible impact on results.

Dependent variable ceiling The dependent variable asked interviewees to self-report the number of days they were unable to work in the 30 days before interview. We might be concerned about the role of the 6,852 (10.9% of sample) for whom response was 30. These might be long- to medium-term disabled, for example, for whom no temperature realization could have delivered an answer different from 30. To address this in column 9 we re-estimate the preferred specification but excluding those. Again, inference is not meaningfully disturbed, though the coefficient now has slightly different interpretation, namely the marginal effect of an extra hot day on number of days unable to work *conditional on* being in that subset of subjects who were able to work on at least one day in the past month.

5.3 Rain

A potential challenge in a study of this sort is disentangling the role of temperature from precipitation. Either might have direct impacts on ability to work (for example through challenges to travel) and both can influence frequency of illness. Indeed in our setting, where the most prevalent type of low-level morbidity is diarrhea, the confounding role of precipitation could be pronounced (Singh et al. (2001) and Vargas et al. (2004)).

While we control for precipitation in all of our specifications (and established in Tables 2 through 4 that dropping the whole vector of non-temperature weather controls did not substantially change our central results) in this section we further test the possibility that we are spuriously attributing rain effects to temperature.

We do this by placing respondents into six bins according to the amount of precipitation experienced at their location in the 30 days before interview. If there were an important confounding influence of precipitation beyond what our controls are picking up, we would expect to see different estimated coefficients on the temperature variable in (say) very wet versus very dry months.

The results of this exercise are reported in Table A.4. The key thing is to observe is not that estimated coefficients are positive and significant in each subsample but rather that the estimates are very consistent across columns. This includes those 29,232 respondents for which precipitation in the 30 days before interview was essentially zero (less than half a millimetre). In other words, and recalling that the rain control is included in almost all earlier regressions, the implied effect of high temperature is roughly the same when the period in question has been wet as when it has been dry suggesting no important confounding effect of precipitation on our inference.

5.4 Alternative standard errors

Our central results included standard errors clustered at district level. We believe this to be the most natural approach, for reasons already noted, and somewhat conservative. Furthermore, given that the temperature variable was constructed at district level is consistent with the advice of Abadie et al. (2017) that the preferred level of clustering is primarily a study design issue and should be chosen to correspond to the level at which the treatment variable is assigned.

Notwithstanding this we consider four alternative approaches to calculating standard errors that some readers might prefer, in each case applied to our whole sample estimates. These include alternative location and time clusters and Eicker-White heteroskedasticity-consistent standard errors. Columns 8 and 9 report Conley standard errors with cut-offs at 15 miles and 250 miles. Conley standard errors are commonly used to account for spatial correlation, that is adjust for the potential dependence between respondents based on spatial proximity.

The results of these exercises are reported in Table A.5. Better than 1% significance is maintained regardless of the approach adopted. The preferred approach (column 1) delivers standard errors that are *larger* than those in seven of the eight other columns, consistent with our inference with regard to statistical significance being conservative.

5.5 Placebos

Placebo exercises are a popular test of design (Hartman and Hidalgo (2018)). If when we replace the independent variable of interest series with a placebo series that we know to be miss-assigned or irrelevant and still obtain significant results then we can infer miss-specification.

We conduct three different such tests. The placebo series in each case are (a) temperature at location of respondent 100 days *after* date of interview (falsely-assigned date); (b) temperature at location of respondent 100 days *before* date of interview (falsely-assigned date); (c) temperature on date of interview but at the location of the respondent in the sample most distant from the respondent in question (falsely-assigned location). The first of these two are commonly-employed in the literature on the effects of daily variations in pollution or heat on diverse outcomes (e.g. Ebenstein et al. (2016) and Rivers et al. (2020)) while in a geographically-large country like India (c) is arguably a more compelling option (e.g. Archsmith et al. (2018) and Heyes et al. (2016)).

The results of these exercises are reported in Table A.6. The estimated coefficients on the placebo temperature variable are much smaller than that from the non-placebo estimation and in no case come close to statistical significance at conventional levels.

6 Conclusions

The frequency of very hot days - in India and elsewhere - is projected to increase substantially over the next 50 years (Stern et al. (2006) and Barnett et al. (2005)). Extreme temperature can impact individuals and economies through diverse channels. In our research we have focussed on identifying a causal link from hot days to self-evaluated inability to work.

The results are stark. Though we present a variety of specifications the consistent message is that a day on which the maximum temperature reached 100°F, not a particularly unusual event in most of India, increases self-reported inability to work substantially. In our preferred specification one extra hot day in the 30 day period increases inability to work across the same period by about 7% or 1/20th of a day. Our measure means that this captures both the contemporaneous effect as well as an within-month spillover effects on other days. The results prove resilient when exposed to a wide battery of robustness tests and falsification exercises.

Two great advantages of our setting are; (1) we have a large sample of respondents, constructed by IHDS-II to be representative along several important dimensions, working in a very wide set of different occupations in different locations and observed across different times of the year. (2) We know a lot about the living circumstances of individual respondents allowing us to explore the protective effects of elements of those circumstances, allowing us to probe in some detail how climate resiliency is influenced by home environment. The results extend to a more aggregate level the recent work of Somanathan et al. (2021) who report mixed evidence on shorter-term heat and absenteeism in a series of case studies.

We are also the first to be able to talk about mitigation with respect to this outcome variable and with respect to the living circumstances of the workers in question. We find evidence of important protective effects of electricity connection and ownership of air-cooling appliances, though not from mains water connection.

Two final points.

First, the nature of the outcome variable based, as it is, on a self-report. We have already observed that respondents will undoubtedly have interpreted this question in different ways. In terms of estimation this is not a significant hurdle provided that interpretation is not itself sensitive to the temperature treatment. Rather it implies that stated response is a noisy but unbiased proxy for the true inability, reducing precision of our estimates. We regard it as a strength that the data is not an administrative count of absence. Many in our sample are self- or informally-employed and are given scope to report not being able to work effectively even though present in the workplace. The 30 day measure also allows for intra-month lags, for example if heat on one day makes inability to work more likely the following.

Second, the analysis is not able to speak to the effect of heat-*waves*. Our central temperature metric is a count of the number of hot days in a 30 day window. That captures nothing about the pattern of heat *within* the window. Recent popular emphasis on heatwaves might suggest that five consecutive 100 degree F days could be more potent than would be the case if those five days were dispersed in time. Our analysis does not pick that up, though we are probing this issue in related research in another setting.

At the same time as being cautious not to over-interpret the results, we also should not under-play their potential importance. The ability of people to work underpins most of what we study in economics - from production, to consumption to government revenue, and all of the things that flow onwards from those. How changing temperature patterns might change that ability is an important step in understanding the climate-economy link, and remains an important topic for further research.

Bibliography

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. (2017). When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research # 24003.

Adhvaryu, A., Kala, N., and Nyshadham, A. (2020). The light and the heat: Productivity co-benefits of energy-saving technology. *Review of Economics and Statistics*, 102:779–792.

Archsmith, J., Heyes, A., and Saberian, S. (2018). Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation. *Journal of the Association of Environmental and Resource Economists*, 5:827–863.

Atkin, D. (2016). The caloric costs of culture: Evidence from Indian migrants. *American Economic Review*, 106:1144–81.

Auffhammer, M., Hsiang, S. M., Schlenker, W., and Sobel, A. (2013). Using weather data and climate model output in economic analyses of climate change. *Review of Environmental Economics and Policy*, 7:181–198.

Azam, M., Chin, A., and Prakash, N. (2013). The returns to English-language skills in India. *Economic Development and Cultural Change*, 61:335–367.

Barbier, E. B. (2014). Climate change mitigation policies and poverty. *Wiley Interdisciplinary Reviews: Climate Change*, 5:483–491.

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438:303–310.

Barreca, A., Clay, K., Deschenes, O., Greenstone, M., and Shapiro, J. S. (2016). Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century. *Journal of Political Economy*, 124:105–159.

Barreca, A. I. and Shimshack, J. P. (2012). Absolute humidity, temperature, and influenza mortality: 30 years of county-level evidence from the United States. *American Journal of Epidemiology*, 176:114–122.

Barron, M. and Torero, M. (2017). Household electrification and indoor air pollution. *Journal of Environmental Economics and Management*, 86:81–92.

Baylis, P. (2020). Temperature and temperament: Evidence from twitter. *Journal of Public Economics*, 184:104–161.

Blakeslee, D. S. and Fishman, R. (2018). Weather shocks, agriculture, and crime evidence from India. *Journal of Human Resources*, 53:750–782.

Burgess, R., Deschenes, O., Donaldson, D., and Greenstone, M. (2017). Weather, climate change and death in India. Technical report, London School of Economics.

Burlig, F. and Preonas, L. (2016). Out of the darkness and into the light? Development effects of rural electrification. Technical report, Energy Institute at Haas #268R.

Carleton, T. A. and Hsiang, S. M. (2016). Social and economic impacts of climate. *Science*, 353:98–137.

Chakravorty, U., Pelli, M., and Marchand, B. U. (2014). Does the quality of electricity matter? Evidence from rural India. *Journal of Economic Behavior and Organization*, 107:228–247.

Cohen, F. and Dechezleprêtre, A. (2021). Mortality inequality, temperature and public health provision: Evidence from Mexico. *American Economic Journal: Economic Policy*, Forthcoming.

Correia, S., Guimarães, P., and Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. *The Stata Journal*, 20:95–115.

Currie, J. and Stabile, M. (2006). Child mental health and human capital accumulation: The case of ADHD. *Journal of Health Economics*, 25:1094–1118.

Dai, L., Kloog, I., Coull, B. A., Sparrow, D., Spiro III, A., Vokonas, P. S., and Schwartz, J. D. (2016). Cognitive function and short-term exposure to residential air temperature: A repeated measures study based on spatiotemporal estimates of temperature. *Environmental Research*, 150:446–451.

Davis, L. W. and Gertler, P. J. (2015). Contribution of air conditioning adoption to future energy use under global warming. *Proceedings of the National Academy of Sciences*, 112:5962–5967.

Dell, M., Jones, B. F., and Olken, B. A. (2012). Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics*, 4:66–95.

Deryugina, T. and Hsiang, S. M. (2014). Does the environment still matter? Daily temperature and income in the United States. Technical report, National Bureau of Economic Research #20750.

Deschenes, O. (2018). Temperature variability and mortality: Evidence from 16 Asian countries. *Asian Development Review*, 35:1–30.

Deschênes, O. and Greenstone, M. (2011). Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US. *American Economic Journal: Applied Economics*, 3:152–85.

Ebenstein, A., Lavy, V., and Roth, S. (2016). The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics*, 8:36–65.

Graff Zivin, J., Hsiang, S. M., and Neidell, M. (2018). Temperature and human capital in the short and long run. *Journal of the Association of Environmental and Resource Economists*, 5:77–105.

Graff Zivin, J. and Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102:3652–3673.

Graff Zivin, J. and Neidell, M. (2014). Temperature and the allocation of time: Implications for climate change. *Journal of Labor Economics*, 32:1–26.

Hansen, J., Sato, M., and Ruedy, R. (2012). Perception of climate change. *Proceedings of the National Academy of Sciences*, 109:2415–2423.

Hartman, E. and Hidalgo, F. D. (2018). What's the alternative? An equivalence approach to balance and placebo tests. *American Journal of Political Science*, 62:1000–1013.

Heyes, A., Neidell, M., and Saberian, S. (2016). The effect of air pollution on investor behavior: Evidence from the S&P 500. Technical report, National Bureau of Economic Research #22753.

Heyes, A. and Saberian, S. (2019). Temperature and decisions: Evidence from 207,000 court cases. *American Economic Journal: Applied Economics*, 11:238–265.

Hsiang, S. 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.

Hunter, P. R., MacDonald, A. M., and Carter, R. C. (2010). Water supply and health. *PLoS Medicine*, 7:100–136.

Isaac, M. and Van Vuuren, D. P. (2009). Modeling global residential sector energy demand for heating and air conditioning in the context of climate change. *Energy Policy*, 37:507–521.

Jessoe, K., Manning, D. T., and Taylor, J. E. (2018). Climate change and labour allocation in rural Mexico: Evidence from annual fluctuations in weather. *The Economic Journal*, 128:230–261.

Khandker, S. R., Barnes, D. F., and Samad, H. A. (2013). Welfare impacts of rural electrification: A panel data analysis from Vietnam. *Economic Development and Cultural Change*, 61:659–692.

Larsen, T. A., Hoffmann, S., Lüthi, C., Truffer, B., and Maurer, M. (2016). Emerging solutions to the water challenges of an urbanizing world. *Science*, 352:928–933.

Lee, K., Miguel, E., and Wolfram, C. (2020). Experimental evidence on the economics of rural electrification. *Journal of Political Economy*, 128:1523–1565.

Lewis, J. (2018). Infant health, women's fertility, and rural electrification in the United States, 1930–1960. *Journal of Economic History*, 78:118–154.

Liss, A., Wu, R., Chui, K. K. H., and Naumova, E. N. (2017). Heat-related hospitalizations in older adults: An amplified effect of the first seasonal heatwave. *Scientific Reports*, 7:39–81.

Mackworth, N. H. (1946). Effects of heat on wireless operators. *British journal of industrial medicine*, 3:143.

Manning, D. T., Means, P., Zimmerle, D., Galvin, K., Loomis, J., and Paustian, K. (2015). Using contingent behavior analysis to measure benefits from rural electrification in developing countries: An example from Rwanda. *Energy Policy*, 86:393–401.

Mohanty, I. and Gebremedhin, T. A. (2018). Maternal autonomy and birth registration in India: Who gets counted? *PloS One*, 13:194–195.

O'Neill, M. S., Carter, R., Kish, J. K., Gronlund, C. J., White-Newsome, J. L., Manarolla, X., Zanobetti, A., and Schwartz, J. D. (2009). Preventing heat-related morbidity and mortality: New approaches in a changing climate. *Maturitas*, 64:98–115.

Park, J. (2017). Hot temperature, human capital, and adaptation to climate change. Technical report, Harvard working paper.

Park, R. J., Goodman, J., Hurwitz, M., and Smith, J. (2020). Heat and learning. *American Economic Journal: Economic Policy*, 12:306–39.

Parsons, K. (2007). *Human thermal environments: The effects of hot, moderate, and cold environments on human health, comfort and performance*. CRC press.

Pilcher, J. J., Nadler, E., and Busch, C. (2002). Effects of hot and cold temperature exposure on performance: A meta-analytic review. *Ergonomics*, 45:682–698.

Ranson, M. (2014). Crime, weather, and climate change. *Journal of Environmental Economics and Management*, 67:274–302.

Reiss, P. C. and White, M. W. (2005). Household electricity demand, revisited. *Review of Economic Studies*, 72:853–883.

Rivers, N., Saberian, S., and Schaufele, B. (2020). Public transit and air pollution: Evidence from Canadian transit strikes. *Canadian Journal of Economics/Revue canadienne d'économique*, 53:496–525.

Rohini, P., Rajeevan, M., and Srivastava, A. (2016). On the variability and increasing trends of heat waves over India. *Scientific Reports*, 6:26–53.

Samad, H. and Zhang, F. (2016). *Benefits of electrification and the role of reliability: Evidence from India*. The World Bank.

Schlenker, W. and Lobell, D. B. (2010). Robust negative impacts of climate change on African agriculture. *Environmental Research Letters*, 5:14–46.

Schlenker, W. and Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of sciences*, 106:15594–15598.

Singh, R. B., Hales, S., De Wet, N., Raj, R., Hearnden, M., and Weinstein, P. (2001). The influence of climate variation and change on diarrheal disease in the Pacific Islands. *Environmental Health Perspectives*, 109:155–168.

Somanathan, E., Somanathan, R., Sudarshan, A., and Tewari, M. (2021). The impact of temperature on productivity and labor supply: Evidence from Indian manufacturing. *Journal of Political Economy*, 129:1797–1827.

Spalding-Fecher, R. (2005). Health benefits of electrification in developing countries: A quantitative assessment in South Africa. *Energy for Sustainable Development*, 9:53–62.

Stern, N., Peters, S., Bakhshi, V., Bowen, A., Cameron, C., Catovsky, S., Crane, D., Cruickshank, S., Dietz, S., Edmonson, N., et al. (2006). *Stern Review: The economics of climate change*, volume 30. HM Treasury London.

Vargas, M., Gascon, J., Casals, C., Schellenberg, D., Urassa, H., Kahigwa, E., Ruiz, J., and Vila, J. (2004). Etiology of diarrhea in children less than five years of age in Ifakara, Tanzania. *American Journal of Tropical Medicine and Hygiene*, 70:536–539.

Wang, W., Xie, G., and Hao, L. (2014). Rural panel surveys in developing countries: A selective review. *Economic and Political Studies*, 2:151–177.

Zhang, P., Deschenes, O., Meng, K., and Zhang, J. (2018). Temperature effects on productivity and factor reallocation: Evidence from a half million Chinese manufacturing plants. *Journal of Environmental Economics and Management*, 88:1–17.

Table 1: Summary Statistics

	Mean	Std. Dev.
Number of days unable to work	0.630	2.766
Zero unable to work day (%)	80.661	-
Number of hot days	7.168	10.721
Zero hot day (%)	42.460	-
Wind speed (km/h)	2.804	1.043
Precipitation (mm)	4.202	8.658
Solar radiation (W/m^2)	19.612	4.773
Relative humidity (%)	0.526	0.222
Income (Million rupees)	0.145	0.247
Assets (Million rupees)	15.408	6.480
Number of persons in household	5.615	2.730
Water availability (%)	92.340	26.600
Electricity availability (%)	88.020	32.475
Age (Year)	39.340	14.032
Highest education (Year)	8.490	4.980
Female (%)	21.181	-
Hindu (%)	81.970	-
Muslim (%)	11.661	-
Sikh (%)	2.381	-

Notes: All proportions use the household weights provided by the IHDS.

Table 2: Main results - Count of hot days

	(1) No controls	(2) Individual controls	(3) Household controls	(4) Financial controls	(5) Weather controls	(6) Type of job controls	(7) Preferred (count)	(8) Preferred (Intensive margin)
Temperature	0.064***	0.061***	0.063***	0.064***	0.076***	0.059**	0.071***	0.093***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.022)	(0.024)	(0.021)	(0.027)
Observations	59,621	59,621	59,621	59,621	59,621	39,883	59,621	25,314
Individual controls	N	Y	Y	Y	N	Y	Y	Y
Household controls	N	N	Y	Y	N	Y	Y	Y
Financial controls	N	N	N	Y	N	Y	Y	Y
Weather controls	N	N	N	N	Y	Y	Y	Y
Job FEs	N	N	N	N	N	Y	N	N
Year FEs	Y	Y	Y	Y	Y	Y	Y	Y
District-month FEs	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Dependent variable is number of days that a worker has been unable to work in 30 days prior to the interview date. Temperature is count of hot days in 30 days prior to the interview date. Individual controls include age, highest education level, religion and gender. Household controls are number of persons in a household, water and electricity availability. Financial controls are household total income and assets. Weather covariates include wind speed, precipitation, solar radiation and humidity. All weather covariates are 30-day average. All regressions include year and district-month. Each specification contains controls as indicated. Column 8 runs our preferred specification on workers that experienced at least 1 hot day in 30 days prior to interview date. Standard errors clustered on district in parentheses. * significant at 10% ** significant at 5% *** significant at 1%.

Table 3: Main results - Average daily maximum temperature

	(1) No controls	(2) Individual controls	(3) Household controls	(4) Financial controls	(5) Weather controls	(6) Type of job controls	(7) Preferred
Temperature	0.100***	0.088***	0.088***	0.100***	0.156***	0.066	0.125**
	(0.030)	(0.028)	(0.027)	(0.030)	(0.050)	(0.057)	(0.049)
Observations	59,621	59,621	59,621	59,621	59,621	39,883	59,621
Individual controls	N	Y	Y	Y	N	Y	Y
Household controls	N	N	Y	Y	N	Y	Y
Financial controls	N	N	N	Y	N	Y	Y
Weather controls	N	N	N	N	Y	Y	
Job FEs	N	N	N	N	N	Y	N
Year FEs	Y	Y	Y	Y	Y	Y	Y
District-month FEs	Y	Y	Y	Y	Y	Y	Y

Notes: Dependent variable is number of days that a worker has been unable to work in 30 days prior to the interview date. Temperature is 30-day average of daily maximum temperature. Standard errors clustered on district in parentheses. See notes to Table 2 for full list of controls. Each specification contains controls as indicated. * significant at 10% ** significant at 5% *** significant at 1%.

Table 4: Main results - Days binned by maximum temperature

	(1) No controls	(2) Individual controls	(3) Household controls	(4) Financial controls	(5) Weather controls	(6) Type of job controls	(7) Preferred
Maximum temp <15	-0.006	-0.021	-0.020	-0.006	-0.014	0.108*	-0.028
	(0.027)	(0.025)	(0.025)	(0.027)	(0.035)	(0.063)	(0.032)
15≤ Max temp <20	0.001	-0.000	0.002	0.001	-0.008	0.009	-0.001
	(0.026)	(0.026)	(0.028)	(0.026)	(0.028)	(0.022)	(0.029)
20≤ Max temp <25	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
25≤ Max temp <30	0.006	0.003	0.009	0.006	0.010	0.007	0.015
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)
30≤ Max temp <35	0.026*	0.020	0.020	0.026*	0.022	0.008	0.018
	(0.014)	(0.014)	(0.013)	(0.014)	(0.016)	(0.018)	(0.015)
35≤ Max temp <40	0.034*	0.027	0.032*	0.034*	0.040**	0.021	0.034*
	(0.018)	(0.019)	(0.019)	(0.018)	(0.019)	(0.022)	(0.019)
40≤ Max temp	0.056***	0.050***	0.053***	0.056***	0.071***	0.042	0.062**
	(0.018)	(0.017)	(0.017)	(0.018)	(0.026)	(0.030)	(0.027)
Observations	59,621	59,621	59,621	59,621	59,621	39,883	59,621
Individual controls	N	Y	Y	Y	N	Y	Y
Household controls	N	N	Y	Y	N	Y	Y
Financial controls	N	N	N	Y	N	Y	Y
Weather controls	N	N	N	N	Y	Y	Y
Job FEs	N	N	N	N	N	Y	N
Year FEs	Y	Y	Y	Y	Y	Y	Y
District-month FEs	Y	Y	Y	Y	Y	Y	Y

Notes: Dependent variable is number of days that a worker has been unable to work in 30 days prior to the interview date. Temperature variables are count of hot days in each of the corresponding bins. Reference category is the 20 to 25 °C bin. Standard errors clustered on district in parentheses. See notes to Table 2 for full list of controls. Each specification contains controls as indicated. * significant at 10% ** significant at 5% *** significant at 1%.

Table 5: Geography

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Count of hot days (2010-2012)			Climate zone			
	Preferred	20-60	60-100	+100	Arid	Montana	Humid	Subtropical
Temperature	0.071***	0.283***	0.062**	0.045	0.080***	-0.117	0.162**	0.081*
	(0.021)	(0.066)	(0.026)	(0.047)	(0.027)	(0.086)	(0.063)	(0.042)
Observations	59,621	5,927	35,916	7,258	15,676	3,722	15,990	11,469

Notes: Dependent variable is number of days that a worker has been unable to work in 30 days prior to the interview date. Standard errors clustered on district in parentheses. See notes to Table 2 for full list of controls. Column 1 coincides with column 7 from Table 2, our preferred specification. Columns 2, 3 and 4 run our preferred specification on subsample of districts that their annual number of hot days from 2010 to 2012 are between 20 to 60, 60 to 100 and more than 100. Columns 5, 6, 7 and 8 run our preferred specification on states that are located in arid, Montana, humid and subtropical climate zones. * significant at 10% ** significant at 5% *** significant at 1%.

Table 6: Heterogeneity

				Type of work				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Preferred	Age	Sex	Agriculture	Construction	Office	Electricity	Unique task
Temperature	0.071***	0.068***	0.064***	0.059**	0.061***	0.057**	0.059**	0.059**
	(0.021)	(0.021)	(0.021)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Age (<20) × Temp	-	-0.011	-	-	-	-	-	-
	-	(0.009)	-	-	-	-	-	-
Age (>60) × Temp	-	0.023***	-	-	-	-	-	-
	-	(0.009)	-	-	-	-	-	-
Female × Temp	-	-	0.017***	-	-	-	-	-
	-	-	(0.004)		-	-	-	-
Agriculture × Temp	-	-	-	-0.005	-	-	-	-
	-	-	-	(0.004)	-	-	-	-
Construction × Temp	-	-	-	-	0.013**	-	-	-
	-	-	-	-	(0.005)	-	-	
Office × Temp	-	-	-	-	-	-0.018**	-	-
	-	-	-	-	-	(0.008)	-	-
No electricity × Temp	-	-	-	-	-	-	-0.001	-
	-	-	-	-	-	-	(0.007)	-
Unique task × Temp	-	-	-	-	-	-	-	-0.006
	-	-	-	-	-	-	-	(0.009)
Observations	59,621	59,621	59,621	39,883	39,883	39,883	39,883	39,883
Individual controls	Y	Y	Y	Y	Y	Y	Y	Y
Household controls	Y	Y	Y	Y	Y	Y	Y	Y
Financial controls	Y	Y	Y	Y	Y	Y	Y	Y
Weather controls	Y	Y	Y	Y	Y	Y	Y	Y
Job FEs	N	N	N	N	N	N	N	N
Year FEs	Y	Y	Y	Y	Y	Y	Y	Y
District-month FEs	Y	Y	Y	Y	Y	Y	Y	Y
Interaction terms	N	Age	Female	Agriculture	Construction	Office	No electricity	Unique task

Notes: Dependent variable is number of days that a worker has been unable to work in 30 days prior to the interview date. Column 1 coincides with column 7 from Table 2, our preferred specification. See notes to Table 2 for full list of controls. In each column, interaction terms are the specified dummy (which is 1 if respondent characteristics falls into that category, zero otherwise) times count of hot days. Standard errors clustered on district in parentheses. * significant at 10% ** significant at 5% *** significant at 1%.

Table 7: Adaptation

		Electricity		Cooling		Water	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Preferred	Connection	Quality	Ownership	Ownership	Piped	Quality
Temperature	0.071***	0.081***	0.073***	0.071***	0.060***	0.067***	0.081***
	(0.021)	(0.023)	(0.022)	(0.021)	(0.017)	(0.021)	(0.022)
Electricity × Temp	-	-0.017**	-	-	-	-	-
	-	(0.008)	-	-	-	-	-
Electricity quality × Temp	-	-	-0.010*	-	-	-	-
	-	-	(0.006)	-	-	-	-
Cooling × Temp	-	-	-	-0.008*	-	-	-
	-	-	-	(0.004)	-	-	-
Hot × Cooling × Temp	-	-	-	-	-0.009**	-	-
	-	-	-	-	(0.005)	-	-
Piped water × Temp	-	-	-	-	-	-0.004	-
	-	-	-	-	-	(0.005)	-
Adequate water × Temp	-	-	-	-	-	-	-0.011
	-	-	-	-	-	-	(0.008)
Observations	59,621	59,621	59,621	59,621	59,621	59,621	59,621
Individual controls	Y	Y	Y	Y	Y	Y	Y
Household controls	Y	Y	Y	Y	Y	Y	Y
Financial controls	Y	Y	Y	Y	Y	Y	Y
Weather controls	Y	Y	Y	Y	Y	Y	Y
Job FEs	N	N	N	N	N	N	N
Year FEs	Y	Y	Y	Y	Y	Y	Y
District-month FEs	Y	Y	Y	Y	Y	Y	Y
Interaction terms	N	Electricity	Hours per day	Cooling	Hot × Cooling	Piped	Adequate

Notes: Dependent variable is number of days that a worker has been unable to work in 30 days prior to the interview date. Column 1 coincides with column 7 from Table 2, our preferred specification. See notes to Table 2 for full list of controls. In each column, the interaction term is the specified dummy times count of hot days. Dummies are defined as follows: Column 2 an indicator that takes value 1 for electricity connection, zero otherwise. Column 3 an indicator that takes value 1 if the household has electricity supply at least 12 hours per day, zero otherwise. Column 4 an indicator that takes value 1 if the household reports ownership of an air conditioner or air cooler, zero otherwise. Column 5 cooling ownership indicator times a dummy that takes value 1 if count of hot days is on average more than 76 (median), zero otherwise. Column 6 an indicator that takes value 1 if household has piped water. Column 7 an indicator that takes value 1 if household has normally adequate water supply, zero otherwise. * significant at 10% ** significant at 5% *** significant at 1%.

Table 8: Alternative heat metrics

	(1) Preferred (count)	(2) Preferred (average)	(3) >95	(4) >105	(5) Heat Index
Temperature	0.071***	0.125**	0.068***	0.090***	-
	(0.021)	(0.049)	(0.035)	(0.026)	-
HI	-	-	-	-	0.146***
	-	-	-	-	(0.038)
Observation	59,621	59,621	59,621	59,621	59,621

Notes: Dependent variable is number of days that a worker has been unable to work in 30 days prior to the interview date. Standard errors clustered on district in parentheses. See notes to Table 2 for full list of controls. Column 1 coincides with column 7 from Table 2, our preferred specification. Column 4 repeats column 7 from Table 3. Column 3 and 4 estimate the preferred specification replacing temperature with count of hot days when maximum temperature exceeds 95 and 105°F. Column 5 replaces temperature with heat index. * significant at 10% ** significant at 5% *** significant at 1%.

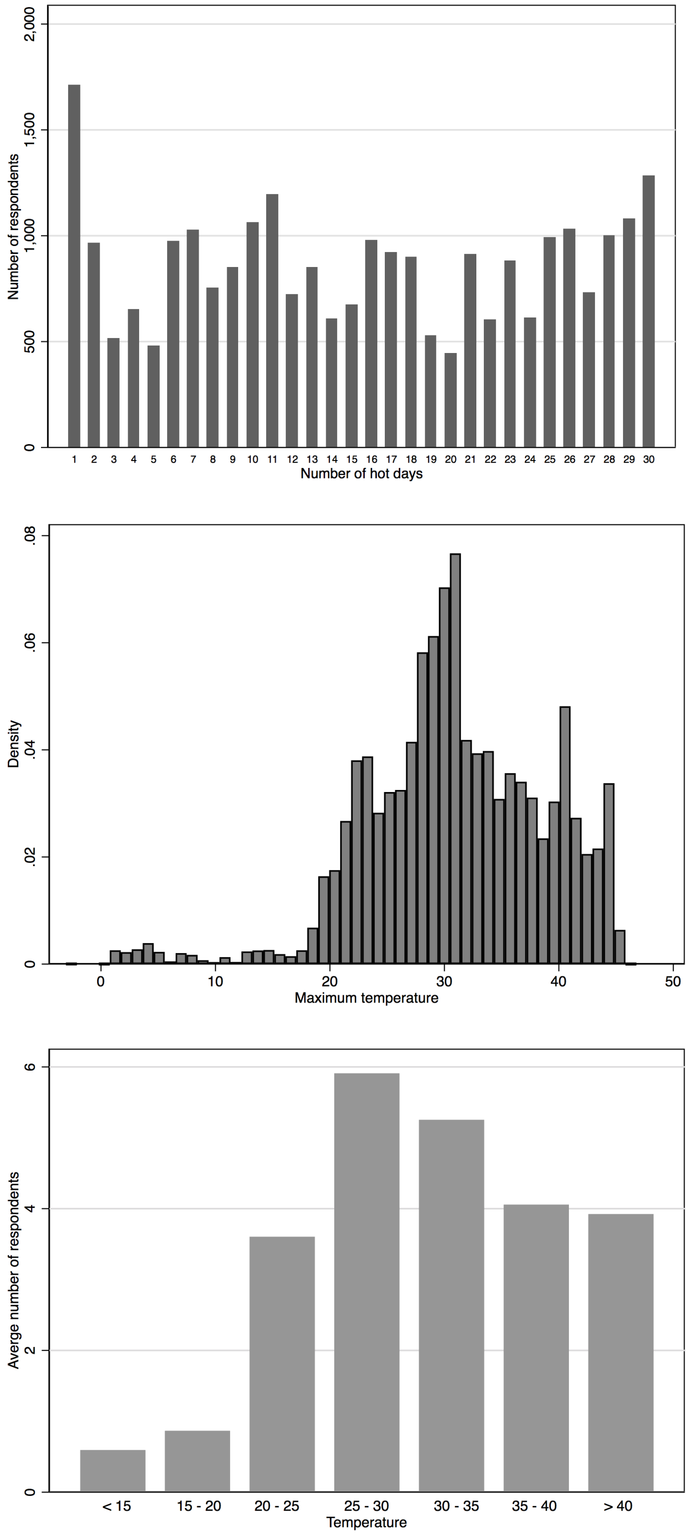
Table 9: Additional exercises

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Preferred	5 Largest	Urban	Rural	North	South	Exclude hottest & coldest decile	Exclude winter	Exclude unable days=30
Temperature	0.071***	0.079*	0.118***	0.066***	0.085***	0.089**	0.078***	0.071***	0.060***
	(0.021)	(0.043)	(0.032)	(0.023)	(0.033)	(0.036)	(0.028)	(0.021)	(0.019)
Observations	59,621	25,277	18,291	40,418	17,750	24,570	29,062	58,456	59,388

Notes: Dependent variable is number of days that a worker has been unable to work in 30 days prior to the interview date. Standard errors clustered on district in parentheses. See notes to Table 2 for full list of controls. Column 1 coincides with column 7 from Table 2, our preferred specification. Column 2 estimates the preferred specification on the five largest states. Columns 3 and 4 estimate the preferred specification on urban and rural subsamples. Column 5 and 6 estimate the preferred specification on the subsample of states located north and south of the Tropic of Cancer. Column 7 excludes district in the hottest and coldest deciles according to their number of hot days. Column 8 excludes respondents that were interviewed during the winter. Column 9 excludes respondents who reported being unable to work for the whole 30 days prior to the interview date. * significant at 10% ** significant at 5% *** significant at 1%.

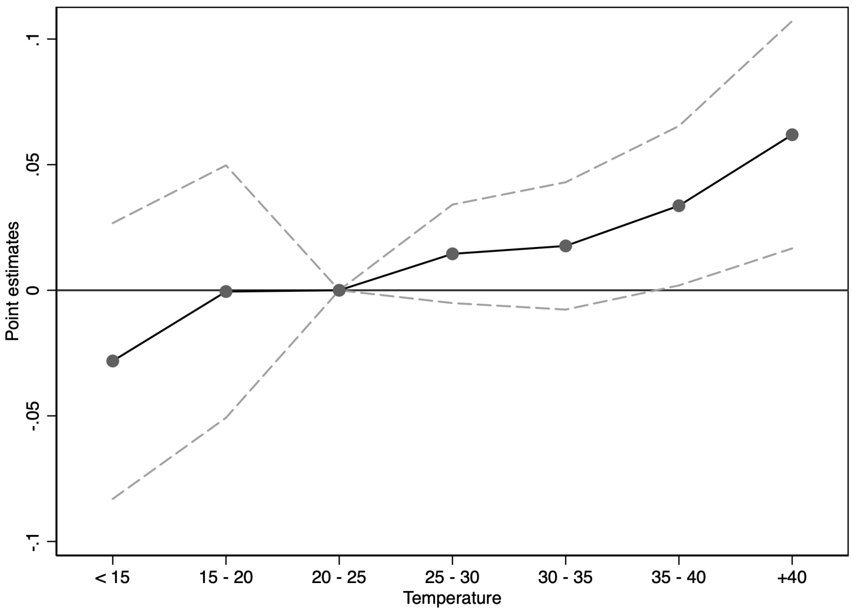
Figures

Figure 1: Variation in treatment



Notes: The upper panel of this figure plots counts of hot days in 30 days prior to the interview date, excluding zero. The second panel depicts density of 30 days average daily maximum temperature.The lower panel plots average number of respondents in each temperature bin for the period of 2011-2012.

Figure 2: Nonlinear estimates



Notes: This figure shows estimated impact of temperature on inability to work relative to a day when maximum daily temperature is between 20 to 25°C. The solid line reports 7 coefficient estimates with circle markers being the effect of a single day in each of the corresponding bins, relative to the effect of a day between 20 to 25°C. Dash line represents 95% confidence interval based on standard errors clustered on district. Regression includes individual, household, financial and weather controls. Year and district-month dummies are included in the regression. All weather covariates are 30 day average. See notes to Table 2 for full list of controls.

A Appendix

Table A.1: Randomization test

	(1) No response	(2) Income	(3) Assets	(4) Number of persons	(5) Sex	(6) Electricity	(7) Electricity quality	(8) Water quality	(9) Cooling ownership
Temperature	0.005	-5863.579	-0.192	0.096	-0.000	-0.000	0.128	0.000	0.011**
	(0.008)	(3640.078)	(0.917)	(0.060)	(0.000)	(0.000)	(0.161)	(0.000)	(0.005)
Observation	62,071	59,621	59,621	59,621	59,621	59,621	59,621	59,621	59,621

Notes: Standard errors clustered on district in parentheses. See notes to Table 2 for full list of controls. All regressions have the same specification as our preferred but with different outcome variable in each column. The outcome variable in column 1 is a dummy that takes value 1 if the respondent does not provide an answer to the question of our interest, zero otherwise. Households income is the outcome variable in column 2. Column 3 estimates the effect of temperature on households total assets. Column 4 estimates the effect of temperature on the number of persons in a household. In column 5 the outcome variable is a dummy that takes value 1 if the respondent is a female, zero otherwise. In column 6 the outcome variable is an indicator for household's electricity connection. In column 7 the outcome variable is a dummy taking value 1 if the household has more than 12 hours electricity connection per day, zero otherwise. In column 8 the outcome variable is a dummy that takes value 1 if the household has adequate water supply, zero otherwise. In column 9, the outcome variable is an indicator for household's cooling (AC or cooler) ownership. * significant at 10% ** significant at 5% *** significant at 1%.

Table A.2: Alternative estimation

	(1) Preferred	(2) OLS	(3) Negative binomial	(4) Zero-inflated poisson
Temperature	0.071***	0.008***	0.080***	0.149***
	(0.021)	(0.002)	(0.025)	(0.032)
Observation	59,621	59,621	59,621	59,621

Notes: Dependent variable is number of days that a worker has been unable to work in 30 days prior to the interview date. Standard errors clustered on district in parentheses. See notes to Table 2 for full list of controls. Column 1 coincides with column 7 from Table 2, our preferred specification. In columns 2, 3 and 4 we re-estimate preferred specification using OLS, negative binomial and zero-inflated Poisson methods. * significant at 10% ** significant at 5% *** significant at 1%.

Table A.3: Adaptation: No financial controls

		Electricity		Cooling		Water	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Preferred	Connection	Quality	Ownership	Ownership	Piped	Quality
Temperature	0.071***	0.090***	0.075***	0.075***	0.064***	0.070***	0.085***
	(0.021)	(0.022)	(0.022)	(0.021)	(0.017)	(0.022)	(0.022)
Electricity × Temp	-	-0.017**	-	-	-	-	-
	-	(0.008)	-	-	-	-	-
Electricity quality × Temp	-	-	-0.001	-	-	-	-
	-	-	(0.006)	-	-	-	-
Cooling × Temp	-	-	-	-0.015***	-	-	-
	-	-	-	(0.004)	-	-	-
Hot × Cooling × Temp	-	-	-	-	-0.016***	-	-
	-	-	-	-	(0.005)	-	-
Piped water × Temp	-	-	-	-	-	-0.006	-
	-	-	-	-	-	(0.007)	-
Adequate water × Temp	-	-	-	-	-	-	-0.012
	-	-	-	-	-	-	(0.008)
Observations	59,621	59,621	59,621	59,621	59,621	59,621	59,621
Individual controls	Y	Y	Y	Y	Y	Y	Y
Household controls	Y	Y	Y	Y	Y	Y	Y
Financial controls	N	N	N	N	N	N	N
Weather controls	Y	Y	Y	Y	Y	Y	Y
Job FEs	N	N	N	N	N	N	N
Year FEs	Y	Y	Y	Y	Y	Y	Y
District-month FEs	Y	Y	Y	Y	Y	Y	Y
Interaction terms	N	Electricity	Hours per day	Cooling	Hot × Cooling	Piped	Adequate

Notes: Dependent variable is number of days that a working individual has been unable to work in 30 days prior to the interview date. This table repeats the same exercise as Table 7 but excludes financial controls.

Table A.4: Rain

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Preferred	$<$0.5	$<$1	$<$2	$<$3	$<$4	$<$5
Temperature	0.071***	0.081***	0.083***	0.090***	0.084***	0.083***	0.078***
	(0.021)	(0.025)	(0.026)	(0.026)	(0.026)	(0.025)	(0.025)
Observations	59,621	29,232	33,786	39,076	41,806	43,800	44,990

Notes: Dependent variable is number of days that a working individual has been unable to work in 30 days prior to the interview date. Standard errors clustered on district in parentheses. See notes to Table 2 for full list of controls. Column 1 coincides with column 7 from Table 2, our preferred specification. Column 2 through 7 estimate the preferred specification on respondents whose average of precipitation in 30 days prior the interview date is less than 0.5 mm, 1 mm, 2 mm, 3 mm, 4 mm and 5 mm. * significant at 10% ** significant at 5% *** significant at 1%.

Table A.5: Alternative standard errors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	District preferred	District and date	State and year	District and year	Climate zone and month	Year-month	Eicker-White	Conley 15 mi	Conley 250 mi
Temperature	0.071^{***}	0.071^{***}	0.071^{***}	0.071^{***}	0.071^{***}	0.071^{***}	0.071^{***}	0.071^{***}	0.071^{***}
	(0.021)	(0.021)	(0.003)	(0.002)	(0.011)	(0.008)	(0.016)	(0.016)	(0.020)
Observation	59,621	59,621	59,621	59,621	59,621	59,621	59,621	59,621	59,621

Notes: Dependent variable is number of days that a working individual has been unable to work in 30 days prior to the interview date. See notes to Table 2 for full list of controls. Column 1 coincides with column 7 from Table 2, our preferred specification. Standard errors in parentheses are two-way clustered on district and date in column 2, state and year in column 3, district and year in column 4 and climate zone and month in column 5. Standard errors in parentheses are clustered on year-month in column 6. Eicker-White standard errors reported in parentheses in column 7. Conley robust standard error with 15 and 250 miles distance cut-off and 5 days lag are in columns 8 and 9. * significant at 10% ** significant at 5% *** significant at 1%.

Table A.6: Placebos

	(1)	(2)	(3)	(4)
	Base	+100 days	-100 days	Furthest
Temperature	0.071***	0.000	0.009	0.044
	(0.021)	(0.008)	(0.006)	(0.059)
Observation	59,621	59,621	59,621	59,621

Notes: Dependent variable is number of days that a working individual has been unable to work in 30 days prior to the interview date. Standard errors clustered on district in parentheses. See notes to Table 2 for full list of controls. Column 1 coincides with column 7 from Table 2, our preferred specification. Columns 2 and 3 falsely-assign temperature from 100 days after and before interview date. Column 4 assigns temperature farthest from the district of a working individual. * significant at 10% ** significant at 5% *** significant at 1%.