

Temperature and Decisions: Evidence from 207,000 Court Cases*

Anthony Heyes Soodeh Saberian
University of Ottawa University of Ottawa

April 19, 2017

Abstract

If decisions with lasting consequences are influenced by extraneous or transient factors then welfare can be damaged. We analyze the impact of outdoor temperature on high-stakes decisions (immigration adjudications) made by professional decision-makers (US immigration judges). In our preferred specification, which includes spatial, temporal and judge fixed effects, and controls for various potential confounders, a one standard deviation increase in case-day temperature reduces positive decisions by 8.56%. This is despite judgements being made indoors, ‘protected’ by climate-control. Results are consistent with established links from temperature to mood and risk appetite and have important implications for evaluating the welfare-burden of climate change.

Keywords: Decision-making - temperature - climate change impacts - adaptation - biology and economics.

*Anthony Heyes is also part-time Professor of Economics at the University of Sussex. Heyes holds the Canada Research Chair (CRC) in Environmental Economics at University of Ottawa and acknowledges financial support from the CRC Program and from SSHRC under Insight Grant # 435-2012-472. We are grateful to Maya Papineau, Matthew Neidell, John List, Patrick Baylis, Pierre Brochu, Sandeep Kapur, Jason Garred, Abel Brodeur and diverse seminar participants for helpful conversations. Errors are ours.

1 Introduction

Decisions - those made by consumers, managers, investors and other sorts of agent - are pivotal to almost all economically and socially important outcomes. Textbooks are full of agents making judgements and trade-offs. It is therefore not surprising that a central ambition of economics (and other behavioral sciences) has been to understand how individuals make decisions. The canonical model is rationality, and welfare analysis typically assumes that agents successfully ‘solve’ the constrained optimization problems that confront them.

However, this assumption is increasingly challenged by evidence that factors not accounted for, and apparently irrelevant, in standard models can cause decision-making to depart substantially from the optimizing ideal. For examples, Mani et al. (2013) show that poverty, by occupying scarce mental resources or ‘bandwidth’, reduces cognitive function and reduces decision quality. Hunger negatively influences mental function (Weaver and Hadley (2009) and Weinreb et al. (2002)) and perception of risk (Ferrarelli (2016)). Tiredness reduces cognitive function (Tchen et al. (2003), Abd-Elfattah et al. (2015)), increases risk-taking (Viner et al. (2008)) and reduces self-control (Kahol et al. (2008)).¹ A wider set of behavioral research, consistent with introspection, points to the importance of transitory emotions and mind-states in influencing decisions with long-term consequences (see Loewenstein (1996) for an early survey). For instance, while Ariely and Loewenstein (2006) show that sexual arousal can impact sexual decision-making, Jahedi et al. (2016) show that it can also influence a wider set of economic decisions by temporarily distorting risk attitudes. All of these fit into the ‘biology and economics’ agenda that seeks to model the agents that populate economic textbooks as biological organisms (‘wet machines’) - sensitive to the environment in which they function.

If decisions with durable consequences are systematically influenced by irrelevant or transient factors the potential for welfare loss is obvious. Our focus is on the role of temperature, and fits into this strand of research. Temperature profiles vary across space. Some places are hotter than others, usually cold places sometimes have hot days, etc.. Furthermore, climate change points to temperature patterns changing over time, both in terms of averages and variability (Stern (2007)). The question we investigate is the following: do decision outcomes, the substance of

¹There is a philosophical debate about how to conduct welfare analysis in these settings (Diamond and Vartiainen (2012)). Typically preferences (say with respect to risk) are regarded as having some longevity. If a person who has lost a night of sleep due to construction noise acts “as if” they have a higher risk appetite than they otherwise would then emerging practice would be to treat the mis-decisions made as welfare-reducing (O’Brien and Mindell (2005) and Halleröd and Larsson (2008)).

which have nothing to do with contemporaneous temperature, depend causally on how hot it is outside on the day the decision is made? Our answer is a resounding yes - with high significance and robustness, and a substantial effect size. As such we evidence a subtle and pernicious channel through which variations in climate (through space and time) can damage wellbeing: By influencing decisions.

We analyze the universe of files (just under 207 000) evaluated over a four year period by the 266 immigration court judges at the 43 US Federal Immigration Courthouse locations spread across most major US cities. Four things make this an ideal test-bed for the theories that we investigate;

(1) The decisions that we observe are socially and economically significant and the appropriate choice self-evidently has nothing to do with contemporaneous temperature. As such any influence of temperature on decisions necessarily implies inefficiency and welfare burden;

(2) Our subjects are experienced decision-makers. While the precise characteristics of any individual file are unique, the setting in which they work and the broad parameters of case files are not novel. Furthermore, the setting mirrors the sort of repetitive-but-idiosyncratic decisions that agents such as consumers and managers face in the main economic models;

(3) The decision-makers that we observe work *indoors* and protected in their workplace by climate-control at a level typical of good-quality US Federal government buildings in the twenty-first century. In terms of protection, then, close to full application of the most obvious technological solution to mitigate temperature effects is already accounted for in the results. With regard to biological adaptation to prevailing conditions, the judges do not move around - they are attached to a single court location - meaning that they are ‘used to’ the prevailing temperature patterns in the city in which we observe them. Further, because location and dates of work are determined externally and in a way not sensitive to short-term temperature realizations we can ignore issues of displacement that might be important in other settings - the work task is fixed in substance, space and time.²

(4) The quality of data and the procedural details of the immigration system allow us to avoid a plethora of identification challenges, allowing for clean, persuasive causal inference.

Our main approach uses high frequency data to estimate a linear probability model with a variety of fixed effects, though we also provide some non-parametric results. We also develop variants in which (a) the dependent variables of interest are

²For example, in some professions an employee might choose to defer work from a hot day to a cooler day (or work in the evening), or decide at short notice to work from home on a particular day.

the Heat Index (a measure used by the US National Weather Service that combines temperature and humidity non-linearly into a metric designed to capture how hot it ‘feels’) and, (b) the variation between realized temperature on a particular date and local norms for that date. Our main identifying assumption is minimal: That temperature realizations are as good as random after accounting for spatial and temporal fixed effects.

The analysis uncovers a substantial effect of short-term (daily) variations in temperature on decision outcomes. In our preferred specification, which includes year-by-month, location and judge fixed effects, as well as controls for case characteristics and other potential environmental confounders, same-day outdoor temperature has an impact on decision outcomes. Our results suggest that a one standard deviation increase in temperature reduces the likelihood of a decision favorable to the applicant by 1.4% which is equivalent to 8.56% decrease in the grant rate (the grant rate in the data as a whole is 16.4%). Consistent with some existing studies of temperature susceptibility by gender (Yu et al. (2010), Xiong et al. (2015)), the effect is particularly pronounced for female judges. To allay concerns that there might be something unique to the immigration setting that is driving the results we repeat the exercise for decisions made in 18 461 Parole Suitability Hearings at the 39 locations of the California Department of Corrections and Rehabilitation (CDCR), arriving at parallel conclusions.

Why are these results important? As a straight piece of law and economics the research contributes to an assessment of the consistency of US immigration (and California parole) practices. The Sixth Amendment to the US Constitution lays out ‘fair trial’ as a fundamental right. Administrative Procedure Act (APA) (1946) determines that any adjudication or decision by an agent of the US government should not be “arbitrary or capricious”. Agency decisions should be “...rationally connected to the facts before it” (Committee on Capital Markets Regulation (CCMR), 2016, p.2). The immigration court system is ‘about’ decisions, and natural justice - as well as the law - dictates that decisions on a particular file should be based solely on the merits of the case (“the facts and nothing but the facts”). There is no plausible reason why a particular file should have any different prospect of success if evaluated on a day unusually warmer for that location at that month-of-the-year, than on a day with a different temperature realization.³

However, as our opening paragraphs suggest, cautiously we propose that the analysis provides a *prima facie* case that temperature may damage decision con-

³There has been a very long and much broader body of debate on arbitrariness in legal systems in the US and elsewhere. Oakley and Coon (1986), Danziger et al. (2011).

sistency and quality in a much wider set of settings. If experienced, professional judges, working in an environment in which they are protected from outdoor temperature with high-quality climate control technology, are as subject to influence as our analysis suggests, what should we think might be the impact of temperature on the wider population of agents (consumers, investors, managers, etc.) making diverse decisions with long lasting implications for welfare?

We are careful not to over-interpret the results, but it is tempting to juxtapose the findings with what we know about differences in temperature profiles across locations and through time. That we do not observe ‘right’ and ‘wrong’ decisions, even ex post, precludes definitive welfare analysis. Given that the correct arbitration does not depend on contemporaneous temperature *the sensitivity of outcomes to changes in temperature in itself implies inefficiency*. However it is not possible for us to point to particular type 1 and 2 errors.⁴ Notwithstanding this we can develop ballpark estimates for “excess” wrong decisions based on an additional *assumption*, grounded in existing research, that human comfort and performance is optimized at a particular temperature range. Other things equal does the analysis imply that immigration system decisions are on average ‘better’ when made in New York than in Phoenix? Or, to extrapolate further, that consumers, market traders, entrepreneurs and others make better decisions in New York than they do in Delhi? Climate change is expected not only to raise average temperatures in many places - will this diminish through time the quality of myriad decisions made by millions of agents within a city, with concomitant welfare losses? Can this provide a mechanism linking temperature patterns, through decisions, to workplace productivity and economic development?

The rest of the paper is laid out as follows. In Section 2 we provide a sketch of some existing research on the effect of temperature on humans, and the mechanisms that might underpin a link from outdoor temperature to indoor decision-making. Sections 3 and 4 detail data sources and methods. Section 5 presents the results of the main analysis and a series of robustness and falsification checks. Section 6 concludes.

⁴We do not have access to decision appeals which, at least superficially, might help identify errors. However, the rights to appeal and review in this area are much less developed than in those areas of law that relate to US citizens (which be construction immigration law does not). In addition, this is an area in which judges have very wide discretion in interpreting case circumstances, and there is no right to appeal purely against how that discretion is exercised. Appeals (as in most areas of law) relate to procedural mistakes.

2 Literature

While mechanism is not going to be our central focus it is worth highlighting several strands of research that link temperature to mental function, decision-making, risk attitudes and mood.

Several studies have examined the role of *indoor* temperature on some measure of mental or cognitive acuity. The temperature in a space is manipulated by the researcher, who then observes some measure of performance. For example, Hedge (2004) and Fang et al. (2004) examine performance on simple visual tasks and abstract problem solving in a laboratory. Wyon et al. (1996) assess vigilance, again in a temperature-manipulated laboratory setting. Chao et al. (2003) measure a set of more complex tasks in an office. Allen and Fischer (1978) measure student learning in classrooms. Seppanen et al. (2006) conduct a meta-analysis of the 24 papers that a particular search protocol elicits on this topic (including those just listed). Of these, 9 take place in the lab, the rest are in offices or schools, and between them they generate just over 100 effect size estimates. Their systematic review of the literature generates an estimate of the indoor temperature associated with highest productivity being at 21.75 °C (71.5 °F) with a decrement of performance of around 9% when temperature is 30 °C (86.1 °F).⁵ In general heat stress has a much greater influence than does cold stress on the performance of cognitive tasks (see Hancock and Vasmatazidis (2003) for a review).

Turning to decision-making, Cheema and Patrick (2012) present five studies of consumer behavior in which they manipulate laboratory temperatures. In higher temperatures subjects are; (a) less likely to engage in gambles (particularly complex gambles); (b) less likely to choose innovative products over established ones, and; (b) more likely to rely on “system 1” (heuristic or habit-based) processing (Pocheptsova et al. (2009)). In our setting - in which the rejection rate of immigration applications is around 83% such that the granting an applicant leave to stay can plausibly be regarded as the less-habitual, more innovative and more risky choice - this would point to a negative relation between high temperatures and grant rates.

While evidence of the effect of contemporaneous indoor temperature on brain-intensive tasks is suggestive for us, none of it is directly relatable. Studies that

⁵The first of these numbers accords with anecdotal introspection. In a more recent review (Cheema and Patrick, 2012, page 985) note that: “Prior studies find that an ambient temperature of 72 °F, one at which most people appear comfortable, may be most conducive for automatic tasks”. For instance, Allan et al. (1979) find that performance on a paired-association memory task peaks at 72 °F. Other evidence suggests a difference between temperatures that are optimal for comfort and those that are optimal for performance. Specifically, Pepler and Warner (1968) show that people perform office work best at 68 °F, although they report feeling cold.”

cast light on the impact of how daily *outdoor* temperature effects indoor mental performance are rare. Graff Zivin et al. (2015) find that (outdoor) temperature above 79 °F on a particular day damages performance of children on math (but not reading) tasks. Park (2016) investigates the relationship between daily outdoor temperature and high school exit exams in New York city and finds that compare to a 72 °F day, taking exam on a 90 °F day reduces students' performance by 0.19 standard deviation.

Turning away from cognition, separate strands of research evidence; (a) a causal link from ambient temperature (and other dimensions of weather) to 'mood', broadly defined, and then; (b) a causal link from mood to decision-making. Baylis (2015) links temperature to measures of hedonic state (mood) using geo-located Twitter activity. His four sentiment metrics based on phraseology, emoticon use and profanity each become more negative once outdoor temperatures exceed 70 °F (with little to no effect for colder temperatures). Denissen et al. (2008) find a similar effect when they analyze online diary entries of 1233 students. Relatedly, a number of behavioral finance papers (for examples Hirshleifer and Shumway (2003), Cao and Wei (2005), Floros (2011)) link daily variations in weather - typically cloud cover and sunshine, but also temperature and humidity - to stock price movements via changes in emotional state.⁶ It is well-established that emotions inform judgement and regulate thought in a variety of different ways (Clore and Huntsinger (2007) provide an accessible introduction to that literature).

The decision-maker in our setting is protected from outdoor temperature by climate control, but may 'import' the effect of exposure to, for example, an extreme outdoor temperature when they move inside, coming in from the morning commute, or after a break.⁷ Determining the physiological mechanisms through which this happens is beyond the scope of our paper, but competition for glucose in the body

⁶Again physiological processes are not well-understood, but Lambert et al. (2002), for example, point to weather (particularly sunshine) influencing the production of serotonin, an important neurotransmitter which is generally thought to be a contributor to well-being. Temperature and other weather variables have been linked to a variety of emotional and behavioral outcomes including aggression (Baron and Bell (1976)), impatience (Ahn et al. (2010)), generosity (Williams and Bargh (2008)), depression (Pflug et al. (1976)) and suicidal tendencies (Page et al. (2007)). However this remains an ill-understood area which deserves further empirical research.

⁷Unfortunately we do not observe the time at which a particular file is heard, or know the movements of the judge during the day, which might have allowed us to investigate intra-day effects - for example larger effects just after a period outdoors.

is a popular theory.⁸ Importantly glucose depletion can have sustained effects.⁹ Outdoor conditions may also condition a judge’s behavior in a way that makes him fatigue, or influences his mood. For example, if external temperature is very high he may not venture outside during breaks ‘for fresh air’. Anyone who has spent time in a city like Houston or Atlanta during a heat-wave should understand that possibility. Lack of fresh air has been linked to reduced cognitive function (Chen and Schwartz (2009)) and depressive mood (Cunningham (1979)). In this way outdoor temperature could in principle affect the output of the subject even *if she never went outside and was exposed to it*.

3 Data

Our central analysis links US-wide data on outcomes of asylum applications with what we know about environmental conditions at the location of decision on the date in question. We also use state-wide parole decisions from California to probe external validity.

3.1 Immigration

We use case-level administrative data on US asylum applications made to immigration courts from January 2000 through September 2004. Our final dataset includes the universe of 206 924 decisions made over this 58 month period by all 266 immigration judges across the 43 US cities in which courts are located (see Figure 1). Each court serves a specific geographical region. Decision data is merged with hand-collected data on judge gender. In our dataset, 34% of judges are female. The mean grant rate (the rate at which a decision is made that favors the applicant) in the database as a whole is 16%.

Our data comes from asylumlaw.org. Asylumlaw no longer operates but was:

⁸“One of the body’s most important tasks is temperature regulation. When the ambient temperature is unusually hot or unusually cold, we use energy - in the form of glucose - to maintain a healthy internal temperature. These two processes - correcting for excessive heat and unwanted cold - are not equally taxing; cooling the body down requires more energy than warming it up. Warm temperatures, then, are more likely to deplete our resources - as our bodies work to maintain homeostasis, we use up large amounts of glucose. Because glucose is also used for mental processes, it may be that the physical demands imposed by excessive warmth reduce our capacity for cognitive functions, thereby reducing our decision-making abilities” Ward (2013). There is a much larger literature on the effect of ambient temperature on a variety of animal behaviors. We do not survey it here. However - for one example among many - Mathot et al. (2015) find that birds are less likely to engage in risky choices at higher-than-familiar temperatures.

⁹Elsewhere, Graff Zivin et al. (2015), p.2 note the existence of a more general “.. neurological literature that documents the brain’s sensitivity to temperature”.

“A website run by an international consortium of agencies that helps asylum seekers in Australia, Canada, the United States and several countries in Europe. It provides links to legal and human rights resources, experts, and other information valuable for asylum seekers.”¹⁰ The data contains date of hearing, identity of judge, nationality of applicant and category of application.¹¹

Asylum decisions made by immigration judges are decisive and those that are denied asylum are subject to removal. Judges sit alone, and there are no formal quotas with respect to their grant rate. While the activities of judges are subject to the overall supervision of the US Attorney General, this is an area of law in which individual judges are commonly regarded as having a high degree of personal discretion and independence in the way in which they evaluate files (see Ramji-Nogales et al. (2007) and Chen et al. (2016)). Though the characteristics of cases that judges in different locations are likely to see will of course vary, the degree of discretion is supported anecdotally by the wide variation in grant rates of judges both between and within particular courthouses. For instance, over the study period in the Los Angeles courthouse there are five judges that granted asylum to fewer than 4% while three others granted in over 67%.

Judges typically determine multiple cases on a given day. The judge is presented with a file, may (or may not) ask questions of the applicant and/or his legal representatives, then enters an adjudication. Within a court all cases are in principle randomly assigned to the judges (Ramji-Nogales et al. (2007)), however we do not test for random assignment on observables, neither does our approach to identification rely on it. The setting of dates for cases and the rostering of judges is done well in advance. For instance, as of December 2016 more than 533 000 immigration cases had hearing dates scheduled, with the average delay from scheduling to hearing being over a year.

An important question in the evaluation of climate change is the extent to which adaptation might allow the impacts of temperature variations to be mitigated. The most obvious protective measures are building design and climate-control. As such it is useful to note in passing the context in which our subjects work. No data is available regarding internal temperatures, nor the precise engineering and thermal properties of buildings. However, all of the courtrooms represented in the study are contained within climate-controlled buildings, as would be expected for important

¹⁰The dataset was provided by Professor Kelly Shue of the University of Chicago Booth School of Business.

¹¹There are two types of cases in immigration courts: affirmative cases in which the applicant presents in the courts on her/his own and defensive cases in which the applicant is referred by the immigration authorities.

operational spaces of the US Federal government. Figure 2 presents pictures of the 16 biggest locations ranked by contribution of cases to sample - contributing between them 86.4% of the total. While the buildings vary, taken as a set it is apparent that the judges work in good quality space, of the sort experienced by many North American professionals.¹² The effects of external temperature on internal behavior that we identify in this paper should be taken as already being adjusted for that level of adaptation embodied in buildings typical of this class.

3.2 Parole

Data on all parole hearings conducted by the Board of Parole Hearing (BPH) between 3 January 2012 and 18 December 2015 is obtained from the California Department of Corrections and Rehabilitation (CDCR).¹³ The dataset includes 18 461 hearing decisions made by 12 BPH commissioners across the 39 venues in California. Figure 3 maps venue locations.

The Board of Parole is responsible for evaluating the risk to public safety from the release of inmates incarcerated for serious crimes. An affirmative decision by the BPH means that a prisoner is returned to the community, so these are high stakes decisions. Parole hearings are conducted in-person with the inmate and his attorney at the inmate’s prison. Sessions are scheduled one year before an inmate becomes eligible for parole and conducted by a panel of two members, a Board Commissioner and Deputy Commissioner (Kathryne et al. (2016)). The former is a non-expert appointed from a variety of professional backgrounds (law enforcement, academia, the military, politics) while the latter is a civil servant and expert in legal process. Formally the Commissioner is responsible for running the hearing and exercising discretion in determining outcome, while the Deputy Commissioner for legalities and post-release management of successful applicants. Despite this, that the panel comprises two members potentially complicates inference, obscuring *individual* decision-making. The grant rate in the dataset - the fraction of cases in which a decision is made that is favorable to the applicant - is 16.48%.

Our data contains the date of hearing, identity of panel members, inmate unique identifier, location of hearing, hearing type and outcome.¹⁴

¹²In an unreported robustness check we dropped each venue one at a time and re-ran the preferred specification on the remaining sample. In no case did this substantially disturb the resulting estimates.

¹³The data can be obtained from: <http://www.cdcr.ca.gov/>.

¹⁴There are two types of hearing that we consider: (1) Initial parole (which is scheduled one year before eligibility), (2) Subsequent parole that is scheduled if there is any consideration in the initial session.

3.3 Environment

Our main research question is whether the adjudication on a file responds to the outdoor temperature on the day on which it is evaluated. To accomplish this, we combine our decision dataset with temperature and a variety of other environmental controls.

The location of asylum decisions from which we construct our dependent variable is drawn from the 43 US cities in which the US Department of Justice operates immigration courthouses. These are widely dispersed (see Figure 1) and subject to diverse weather conditions.

The exact date and location of each decision is known which allows us to assign environmental measures (pollution and weather) to each. Temperature and other weather data is obtained from two sources. Hourly observations for air temperature, dew point, air pressure, precipitation and wind speed are retrieved from the National Oceanic and Atmospheric Administration (NOAA).¹⁵ Data for cloud cover comes from the Northeast Regional Climate Center (NRCC).¹⁶ Weather information is assigned to each courthouse location from the closest monitoring stations, in no case further than 20 miles away. The average distance between weather monitoring stations and courthouses is 9.35 miles with standard deviation of 6.33.

For our central specifications we work with averages computed for the period 6 AM to 4 PM each day. This is the period over which decision-makers are likely “up and about” - including travel to work, and work day. It excludes exposure that arise after courts close, which logically can have no effect on proceedings. Figure 4 summarizes the annual distribution of 6 AM to 4 PM mean temperatures across ten temperature-day categories over the study period (2000 to 2004) across locations in 10 °F bins. Most existing research on the effects of short-term temperature and pollution on a variety of outcome variables work with calendar-day data and, while we believe this to be an inferior approach, for purpose of comparison we also present analysis on that basis. In a further variant, that we do not report, we also conduct the exercise using 8-hour blocks (Midnight to 6 AM, 6 AM to 4 PM, 4 PM to midnight).

We will also be controlling for air quality conditions. Daily pollution data is published online by United States Environmental Protection Agency (USEPA).¹⁷ The dataset includes daily measures of particulate matter less than 2.5 microns in width ($PM_{2.5}$), carbon monoxide (CO) and ozone (O_3) throughout the United

¹⁵The data is obtained from: <https://www.ncdc.noaa.gov/>.

¹⁶The data is retrieved from: <http://www.nrcc.cornell.edu/>.

¹⁷The data is available at: <https://aqs.epa.gov/api>.

States for the period of 2000 to 2004.

Table 1 presents summary statistics for all variables.

4 Methods

4.1 Empirical strategy

We estimate the following linear probability model:

$$g_{it} = \beta_0 + \beta_1 temp_{it} + W_{it}\beta_2 + P_{it}\beta_3 + X_{it}\beta_4 + \Psi_c + \theta_t + \epsilon_{it} \quad (1)$$

where g_{it} is a binary variable that takes the value one if the judge’s decision in asylum application i on date t is granted, zero otherwise.

The key independent variable is the mean 6 AM to 4 PM temperature on the date the case is considered, $temp_{it}$. For most of our discussion β_1 is the coefficient of interest.

To allow for the possibility that other dimensions of weather rather than temperature might impact decisions, we include a vector of weather controls, W_{it} . It contains dew point temperature (a standard measure of humidity), precipitation, wind speed, air pressure and sky cover on date t , in the vicinity of the courthouse in which application i is adjudicated, all calculated on a 6 AM to 4 PM average basis. Pollution exposure can also influence cognitive function, mood and/or decision-making (Archsmith et al. (2016), Chang et al. (2016), Lavy et al. (2016)). To allow for this possibility we include P_{it} which is a vector of pollution controls. It comprises mean daily measures of ozone (O_3), carbon monoxide (CO) and ultra fine particulate matter ($PM_{2.5}$).

Court and case context can be expected to impact case outcomes (Chen et al. (2016)). We include a vector X_{it} of controls for a number of additional court and application characteristics. More specifically, we control for the identity of judges, category of application and nationality of applicants.¹⁸ The vector θ_t includes time fixed effects; day of week to account for possible changes in decision patterns across the day of week and year-by-month fixed effects. The latter capture both seasonal and temporal trends in the data. Finally, Ψ_c is a vector of city fixed effects to control for any time-invariant unobservables common to each city that might impact

¹⁸The case characteristics that we observe are limited in the dataset. It is clear that other unobserved characteristics are important determinants of case outcomes such that we have omitted variables. However, once we have controlled for location and time fixed effects it is plausible that those omitted characteristics would be uncorrelated with case-day temperature such that the OLS estimate of β_1 would be unbiased and the associated standard error remains undisturbed.

immigration outcomes.

Error terms may be spatially- and serially-correlated. In our preferred specification standard errors are clustered by city-week which serves two purposes: to account for spatial correlation across cities and to allow for autocorrelation in decisions in each week. For the purposes of robustness we establish later that the results are robust to a variety of other ways of calculating standard errors.

As noted we include a rich set of fixed effects. In particular we are identifying off within-location, within-month variation. Our identifying assumption is that once location and time effects are controlled for, the realization of outdoor temperature on any *particular* day - and therefore the assignment of a temperature treatment to any particular decision - is as good as random. That is to say, we can examine cases heard in Atlanta, in the first week of June. Sometimes a case adjudicated on day in the first week of June in Atlanta can be assigned a temperature treatment of 60 °F, other times 90 °F. It is that variation, plausibly exogenous, that we exploit for identification.

5 Results

5.1 Linear

The base results are summarized in Table 2. Column (1) is the preferred specification, incorporating the full suite of controls - time fixed effects, weather and pollution controls.¹⁹

The coefficient in column (1) is -0.000893 implies that a one standard deviation increase in 6 AM - 4 PM temperature on the day a decision is made reduces the likelihood of a grant decision by 1.4%. Recall that the average grant rate in the sample is 16.4%, so this implies a 8.56% decrease in grant rate.

Columns (2) and (3) of Table 2 reports the results of including a lag or a lead. In each case the point estimates on the lagged terms are much smaller in absolute value than those on the main measure, mixed in sign, and never approach significance at conventional levels. Column (4) includes both lead and lag terms. Figure 5 plots results when we add three lags and three leads simultaneously. As can be seen, none of the lags or leads achieve significance.

¹⁹All of our main specifications are estimated on the whole 58 months of data. The terrorist attacks of 11 September 2001 fall during our study period and can be expected to have impacted on the operation of the immigration system in the US. While we do not report them here, we have run the main specifications on the pre- and post-9/11 portions of the data-set, observing consistent patterns across them.

Our main specification incorporates what we believe to be the most natural set of fixed effects (year-month and city). However Table 3 reports the results of other approaches. Column (1) repeats the preferred specification for purposes of comparison, while columns (2) to (4) present three plausible alternatives. Importantly the estimated coefficient retains sign and significance in each case, and values are disturbed comparatively little.

Table 4 explores the sensitivity of results to some alternative but plausible specifications.

Much of the related literature on short-term effects of weather and air quality on human outcomes has used the calendar day as its unit of analysis (for examples, Hirshleifer and Shumway (2003), Lavy et al. (2016) and Park (2016)). While this is not our preferred approach - a substantial portion of each calendar day occurs after the court is closed - for comparability we report in Table 4 column (2) the results of repeating the exercise on a calendar day basis. As would be expected given the introduction of additional imprecision into the way in which the regressor of interest is measured, the estimated coefficients are attenuated somewhat, but retain sign and significance and are similar in magnitude to in Table 2 (-0.0008 instead of -0.0011 for the preferred specifications).²⁰

Decision locations are dispersed widely across the country and in places that exhibit very different weather patterns. This implies that a 90 °F degree day in Phoenix may not have the same effect as such a day in Boston. The inclusion of city and year-month fixed effects should control for unobservable characteristics of that location at that time of year (such as “normal” weather conditions). However to probe this further we estimate a variant in which the independent variable of interest is the deviation of 6 AM - 4 PM temperature on decision day from the average 6 AM - 4 PM temperature for that location in that week of the year. The results of this exercise are summarized in column (3). The point estimate on same-day temperature deviation is negative and significant at 5%.

The results of an additional exercise to address the concern that the impact of a given temperature treatment may vary by location is reported in column (4). Here we re-estimate the preferred specification but now incorporating a vector of *city* and *temperature* interaction terms, with New York chosen as our reference city. Point estimates on 40 out of the 44 interaction terms are insignificant (the exceptions are San Francisco, Los Angeles, Philadelphia and New Orleans). As can be seen, the coefficient of interest is little changed.

²⁰In a further variant we conducted the exercise using 8-hour blocks (Midnight - 8 AM, 8 AM - 4 PM, 4 PM - midnight). We do not report the results of this here, but they parallel those just presented.

Most of the evidence that we present points to the depressing effect of hot days on affirmative decisions (this will be confirmed in the non-parametric results that follow). Much of the US is cold during the winter months, while the whole mainland is mild to hot during the rest of the year. Column (5) reports the results of re-estimating the preferred specification but excluding the winter months. Again, the coefficient on temperature retains sign and significance, though it is now somewhat larger in absolute value.

5.2 Non-linear

In addition to the conventional linear estimate we also examine possible of non-linearity in the relationship between temperature and decision outcomes by re-estimating using temperature bins 5 °F in width, with the 50 - 55 °F bin as the reference category.

The results of this non-parametric exercise from this analysis are presented in column (1) of Table 5 and illustrated in Figure 6. Point estimates are statistically significant when temperature exceeds 55 °F. They are also meaningful in size. Other things equal, taking a case heard on a day where outdoor temperature is between 50 - 55 °F and dropping it instead into a day where the temperature exceeds 85 °F reduces the likelihood of a favorable decision by 5.83%.

The negative effects of temperature appear close-to-linear and most of the robustness checks and other exercises that we conduct below will be centred on the linear results.

5.3 Heterogeneity by gender of judge

Table 6 reports a test of the hypothesis that temperature-sensitivity is particularly pronounced amongst females (Yu et al. (2010), Xiong et al. (2015)).

For this exercise we re-estimate the preferred regression specifications on the subsample of decisions made by female judges (72 229 decisions made by 95 individuals) and male judges (134 695 decisions made by 171 individuals) separately. In Table 6 the results of these exercises are summarized in columns (2) and (3) respectively. In each case the point estimate is negative and significant at the 5% level. However the female coefficient is around 45% bigger in absolute value. The Hausman test (reported in the lower panel of Table 6) confirms that the coefficient values are significantly different at the 5% level (p value 0.013).

In addition to showing consistency with the literature on gender and temperature-sensitivity already cited, the result also goes some way to address a concern that the

patterns that we observe are driven not by the effect of temperature on judgement, but that temperature is instead influencing outcomes by impacting (for example) the comportment of the applicant or his lawyer. If that (or other external-to-judge mechanisms) were the channel we would not expect to see differences based on gender of judge.

5.4 Robustness

Table 7 reports the results of a battery of robustness tests.

5.4.1 Pollution

Recent research points to a possible link from short-term pollution exposure to mood and cognitive function, either of which might influence decision outcomes (Heyes et al. (2016) and Szyszkowicz et al. (2010)). While our main specifications include controls for ambient levels of the main pollutants (O_3 , $PM_{2.5}$ and CO), concern may remain that we have failed to control adequately for air quality effects, and that these are confounding our results. If that were the case then we would expect dropping the whole set of pollution controls to substantially affect our estimate of β_1 . In column (2) we report the result of re-estimating the preferred specification but omitting the vector of pollution controls. The estimated coefficient on temperature retains sign and significance and value changes only a little (-0.0009 instead of -0.0011).

5.4.2 California

Of our 43 venues 6 are located in California (accounting for around 32% of all decisions). To rule out that what we are picking up something idiosyncratic to California - particularly since our external validity exercise is going to rely on Californian parole data - we re-estimate our preferred specification excluding decisions made at courts in that state. This excludes around 71 000 of the 207 000 decisions in sample. The result of this exercise are reported in column (3) of Table 7. Again, when estimated on the restricted sample the estimate of β_1 retains sign and significance and is little-changed in value (-0.0010 instead of -0.0011). So the pattern that we observed in the data is not being ‘driven’ by anything particular to California.

5.4.3 Weather

Columns (4), (5) and (6) probe further the potential confounding role of rain and cloud.

Existing research points to cloud cover as influencing mood (Lambert et al. (2002), Kent et al. (2009) and Hirshleifer and Shumway (2003)). We include a continuous variable that captures extent of cloud cover in our main specification to control for this. However, as a further test we re-estimate the central specification on those decisions made on “clear sky” days - the subset of days when daily cloud cover is less than 5% (results in column (4)). The point estimate of β_1 for the subsample estimation remains negative and significant. Though larger in absolute value (-0.0025 instead of -0.0011) the difference between the two values is not significant at the 5% level.

Similarly rain can influence mood (Denissen et al. (2008)). While a continuous measure of precipitation is included in the vector of weather controls, column (5) reports the result of re-estimating the preferred specification on the subset of decisions (157 681 of them) made on days in which local recorded precipitation is zero. On such days rain cannot plausibly be argued to have influenced outcomes. The estimated coefficient retains sign and significance and is unchanged in value (at -0.0011). Column (6) reports the results of pushing this further by repeating the same exercise this time excluding days on which recorded precipitation on either the day of decision or the day before were non-zero (111 361 decisions). Again the point estimate on the coefficient of interest is somewhat larger in absolute value (-0.0014 instead of -0.0011) but retains sign and significance.

5.4.4 Heat Index (HI)

The way in which temperature is experienced by the human body can itself depend on the water content of the air. Humidity is known to affect both mood and labor productivity (Howarth and Hoffman (1984), Tsutsumi et al. (2007) and Wan et al. (2009)). We therefore investigate the joint effect of temperature and humidity in our setting by dropping temperature and dew point from our preferred specification and replacing it with the Heat Index (HI). The HI is used by the US National Weather Service and combines air temperature and relative humidity, via a non-linear algorithm, into a single metric designed to capture how hot it ‘feels’. It effectively adjusts upwards the dry air temperature for moisture content to provide an index of the discomfort associated with a particular temperature/humidity combination.²¹

²¹Countries including the UK and France have an alternative index - called Humidex - that has the same intention, and is highly correlated with HI, but is calculated by a slightly different

Column (7) reports the results of re-estimating our preferred specification but with HI added, temperature and dew point dropped. Consistent with earlier results we find a negative and significant effect of heat index on decision outcomes. Though the coefficient here is not directly comparable to those from the various other specifications, the point estimate implies that a one standard deviation increase in HI reduces the probability of grant decision by 0.56% (recall that this is against an average grant rate in the sample of 16.4%). However, since HI is primarily regarded as a reliable measure of discomfort only in warm conditions, we also conduct this exercise once more on the subsample of days on which the local heat index exceeds 75 °F in column (8). The estimated coefficient on heat index is negative and significant with an absolute value larger than in column (7), though estimated on a much smaller sample.

In Figure 7a and column (2) of Table 5 we repeat this exercise for the HI variant of the analysis - with dew point temperature and temperature omitted as regressors but HI added. As noted, this provides a plausible way for allowing for the combined effects of temperature and humidity on how heat ‘feels’. Since HI is only regarded as reliable on warmer days, Column (3) of Table 5 and Figure 7b repeat the same exercise for the subsample of days on which HI exceeds 65 °F, with the 65 - 70 °F bin as the reference category. Again the negative impacts of HI exhibit a close to linear pattern with the negative effect become significant for values of HI exceeding 80 °F.

5.4.5 Outlier judges

We note in the data section that judges do not have specific quotas with respect to what their grant rates should be - indeed this is an area of the legal system in which judges, sitting alone, are regarded as exercising a very high degree of personal discretion (Ramji-Nogales et al. (2007)). To convince ourselves that the result that we are claiming is not being driven by ‘extreme’ judges we conduct two outlier analyses.²² In the first we exclude those decisions made by judges who have a grant rate across the whole study period in either the top or the bottom quartile (just retaining the ‘middle half’ of judges when ranked in terms of moderation).²³ Column (9) reports the results of this exercise - again sign and significance is retained and

formula. HI and Humidex references are often heard on media weather broadcasts during warmer times of year. The HI is typically seen as relevant or reliable measure only in warm conditions.

²²For example, suppose there existed a judge who is so extreme that he never found in favor of the applicant (his grant rate was 0%). The grant rate of that judge could not go lower upon exposure to high temperature because he is already at the lower bound. Recall that we already have judge fixed effects in all of our main specifications.

²³This excludes decisions made by judges who have overall grant rates below 9% or above 25%.

the value of the coefficient is little disturbed (-0.0009 instead of -0.0011). In the second we conduct the same exercise but exclude the top and bottom deciles of judges.²⁴ The results of this is reported in column (10). Again the sign and significance is retained and the value of the coefficient is little disturbed (-0.0008 instead of -0.0011).

5.5 Alternative standard errors

The main results in Table 2 - and the significance levels reported there - rely on standard errors clustered on city-week.

We believe this is the most natural approach. However, in Table 8 we present standard errors from six alternative clustering strategies (columns (1) through (5)) and heteroscedasticity-consistent Eicker-White standard errors. In all cases the level of significance of the estimated coefficient is unchanged. While alternative clustering makes little difference the Eicker-White standard errors can be seen to be around 30% smaller.

5.6 Placebos

As further falsification tests we perform three placebo exercises.²⁵ First, we replace the decision-day temperature series with the temperature at the same location 100 days after decision-day, and 100 days before. Second we replace the decision-day temperature in the vicinity of the courthouse in which the decision was made with the temperature on the same day, but taken from the weather monitoring station *most distant from it* “as the crow flies”. For example, for Hartford (Connecticut) the placebo temperature is taken from the NOAA measuring station at Davenport (California) 4238.72 miles away; for Dallas (Texas) the placebo temperature values are taken from Port Angeles (Washington) 2792.42 miles away.

The results of these exercises are reported in Table 9. In each case the absolute value of the estimate of the coefficient of interest is several times smaller, signs are mixed and in no case is statistical significance achieved.

²⁴This excludes decisions made by judges who have overall grant rates below 4.6% or above 31%. Note that while we exclude the top and bottom decile of judges we do not lose exactly 20% of our sample of decisions. This is because different judges are associated with different numbers of decisions across the study period.

²⁵For this exercise we limit analysis to mainland US locations (exclude weather stations in Puerto Rico and Hawaii). We ran a wide variety of other placebos with similar (insignificant) results.

5.7 Parole

Until now we have focused on judges evaluating immigration files. We are not going to claim broad generality of results, though we believe they are highly suggestive of what is likely to be a wider phenomenon. However to probe at least a little into whether the effects that we have identified are unique to the immigration setting we repeat the central linear and non-parametric analysis for decisions made by parole commissioners in the context of Californian parole hearings.

Table 10 presents results that repeat the main part of our analysis on a calendar day basis using results from the universe of hearings for the period of 3 January 2012 to 18 December 2015 (18 461 in total) as dependent variable. More concretely the dependent variable is a dummy that takes the value one if a parole applicant is granted release, zero otherwise.

The pattern of results presented in Table 10 proves similar to those earlier. Decision-day outdoor temperature has a significant, negative effect on likelihood of a decision to release the applicant. The effect is similar in magnitude to the immigration setting. In the preferred specification (column (1)) a one standard deviation increase in outdoor temperature reduces the probability of a grant release decision by 1.2%. Against an average grant rate in the data-set of 16.48% this implies a 7.3% decrease in the rate of affirmative decisions. We also test the implications of adding a single lag or lead, both individually and concurrently (columns (2), (3) and (4)), again finding coefficients on these that are much smaller, mixed in sign, and never achieve significance. That their inclusion or exclusion disturbs the estimated coefficient of interest more than in the immigration case likely reflects the lower day-to-day variation in the mid to southern Californian locations of the hearing venues.

Figure 8 depicts the results of non-parametric analysis for Californian parole hearing decisions. Point estimates are statistically significant at 5% for temperatures exceeding 65 °F. Consistent with the results from the immigration setting, there is close to linear effect of temperature on decision outcomes. Results suggest that compared to a day with average temperature in the 50 to 55 °F bin, the likelihood of releasing an inmate is 2.6% lower on a day when temperature is higher than 85 °F. In the context of an overall grant rate of 16.48% this corresponds to a 14.7% fall in the rate of decisions favoring the applicant - a substantial effect.

6 Conclusions

Temperature patterns are changing - in much of the world average temperatures are rising, as are the frequency of very hot and very cold days. Understanding how such changes are likely to influence a variety of social and economic outcomes is crucial to forming a measured view of the implications of such change.

We present what we believe to be the first evidence - in either a naturally-occurring or artificial setting - that same-day outdoor temperature influences indoor decisions. Effect sizes are large and robust. That we study a naturally-occurring, high-stakes setting populated by experienced subjects adds to the likelihood that the effects identified reflect a broader phenomenon.²⁶ While the evaluation of a file may be sensitive to the case-day behavior of the applicant, and we cannot rule out that part of the effect that we uncover works through induced changes in that, the heterogeneity of effect between male and female judges points to an internal-to-judge effect. If this was purely a story about over-heated applicants changing their comportment, we would not expect the gender of judge in a particular case to matter.

While we don't observe their precise movements nor the particularities of the indoor conditions in which they work, we can say that these professionals work in good quality, climate-controlled environments. Also, presumably, they travel to work and move around their cities in a manner consistent with better off professional workers (have air conditioning in their cars, *etc.*). In other words, the subjects that we study are offered a level of protection against weather variations that most people, even office-based professionals, would find quite comprehensive. That despite this we still observe substantial and robust effects of ambient temperatures outdoors to how these individuals are going about their business indoors, causes us to be sceptical of claims that climate control is likely to be fully-effective in ameliorating climate impacts.

There are different ways to think about the implications of the results. At the broadest, we provide a bridge from local climate to what is happening indoors - where most high value employment is based, and where most important work and non-work decisions are taken - even when the agents, and the buildings in which they work, are adapted to local conditions.

As such we can, amongst other things, provide a plausible link from local climate to workplace productivity. Of course we rarely have persuasive measures of individual, daily productivity in high value employment settings (which is why ex-

²⁶The parole results provided some 'out of sample' testing, and reassurance that the patterns that we see in the immigration data are not unique to that setting.

isting research has focussed on low-grade jobs such as picking fruit and answering routine calls in a job center). Our setting shares that shortcoming since the job of a judge is quality-driven and we do not observe ‘right’ and ‘wrong’ decisions, even ex post. However, given that the correct arbitration self-evidently does not depend on contemporaneous temperature *the sensitivity of outcomes to changes in temperature in itself implies welfare inefficiency*. Insofar as the correct arbitration matters - in other words that this is from a societal perspective of a high-stakes setting - the large effect sizes imply that the welfare losses are, in turn, large.

Away from the world of work, decisions are central to human well-being. We all routinely make decisions about what to buy, how to invest, how to vote, when to quit our jobs, etc. If decisions with durable impacts are systematically affected by irrelevant, transient factors then the potential for individual and welfare loss across many settings is obvious.

One area in which we have been agnostic throughout the paper is channels. Pinning down the mechanism(s) from outdoor temperature to indoor decision processes would be a useful ambition of future work, and probably initially best suited to laboratory or laboratory-in-the-field methods. The two broad channels that we noted in the introductory review that are consistent with the results relate to (1) mood and (2) cognitive acuity. High temperatures may stimulate temper, irritability (for example in Baylis (2015) Twitter users are more likely to use profanity) and other emotions that might induce a judge to be less well-disposed towards a typical applicant. In addition depressive mood has been linked to reduced risk appetite. In both the immigration and parole settings denying a request can be plausibly be regarded as the risk averse course of action. Mental fatigue and other effects of heat can reduce mental acuity which can increase mistakes, and also themselves induce transient increases in risk aversion.

Just as we have sought not to over-sell the results, neither should we over-state the limitations. That outdoor temperature can have a large, significant and apparently robust effect on indoor decisions, even when subjects operate in a climate-controlled setting, has potentially huge ramifications for how we think about the links from climate to human well-being. The bounds on those effects, and the mechanisms underpinning them, are important foci of ongoing research.

Bibliography

- Abd-Elfattah, H. M., Abdelazeim, F. H., and Elshennawy, S. (2015). Physical and cognitive consequences of fatigue: A review. *Journal of Advanced Research*, 6(3):351–358.
- Administrative Procedure Act (APA) (1946). The Administrative Procedure Act (United States): Pub.L. 79–404, 60 Stat. 237. <http://www.legisworks.org/congress/79/publaw-404.pdf>. Accessed: 2017-01-21.
- Ahn, H.-K., Mazar, N., and Soman, D. (2010). Being hot or being cold: The influence of temperature on judgment and choice. *NA-Advances in Consumer Research*, 37(1):85–88.
- Allan, J. R., Gibson, T. M., and Green, R. G. (1979). Effect of induced cyclic changes of deep body temperature on task performances. *Aviation, Space, and Environmental Medicine*, 50(6):585–589.
- Allen, M. A. and Fischer, G. J. (1978). Ambient temperature effects on paired associate learning. *Ergonomics*, 21(2):95–101.
- Archsmith, J., Heyes, A., and Saberian, S. (2016). Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation. Technical report, Unpublished Manuscript UC Davis Economics.
- Ariely, D. and Loewenstein, G. (2006). The heat of the moment: The effect of sexual arousal on sexual decision making. *Journal of Behavioral Decision Making*, 19(2):87–98.
- Baron, R. A. and Bell, P. A. (1976). Aggression and heat: The influence of ambient temperature, negative affect, and a cooling drink on physical aggression. *Journal of Personality and Social Psychology*, 33(3):244–245.
- Baylis, P. (2015). Temperature and temperament: Evidence from a billion tweets. Technical report, Energy Institute at HAAS Working Paper.
- Cao, M. and Wei, J. (2005). Stock market returns: A note on temperature anomaly. *Journal of Banking and Finance*, 29(6):1559–1573.
- Chang, T., Zivin, J. G., Gross, T., and Neidell, M. (2016). The effect of pollution on worker productivity: Evidence from call-center workers in China. Technical report, National Bureau of Economic Research.
- Chao, H., Schwartz, J., Milton, D., Muillenber, M., and Burge, H. (2003). Effects of indoor air quality on office workers’ work performance—a preliminary analysis. *Proceedings of Healthy Buildings*, 3(4):237–243.
- Cheema, A. and Patrick, V. M. (2012). Influence of warm versus cool temperatures on consumer choice: A resource depletion account. *Journal of Marketing Research*, 49(6):984–995.

- Chen, D., Moskowitz, T. J., and Shue, K. (2016). Decision-making under the gambler’s fallacy: Evidence from asylum judges, loan officers, and baseball umpires. *Quarterly Journal of Economics*, 131(3):1181–1241.
- Chen, J.-C. and Schwartz, J. (2009). Neurobehavioral effects of ambient air pollution on cognitive performance in US adults. *Neurotoxicology*, 30(2):231–239.
- Clore, G. L. and Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, 11(9):393–399.
- Committee on Capital Markets Regulation (CCMR) (2016). Nothing but the facts: The regulatory reform process. http://www.capmksreg.org/wp-content/uploads/2016/11NBTF_Regulatory_Reform_.pdf. Accessed: 2017-01-21.
- Cunningham, M. R. (1979). Weather, mood, and helping behavior: Quasi experiments with the sunshine samaritan. *Journal of Personality and Social Psychology*, 37(11):1947–1956.
- Danziger, S., Levav, J., and Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, 108(17):6889–6892.
- Denissen, J. J., Butalid, L., Penke, L., and Van Aken, M. A. (2008). The effects of weather on daily mood: A multilevel approach. *Emotion*, 8(5):662–635.
- Diamond, P. and Vartiainen, H. (2012). *Behavioral economics and its applications*. Princeton University Press.
- Fang, L., Wyon, D., Clausen, G., and Fanger, P. O. (2004). Impact of indoor air temperature and humidity in an office on perceived air quality, SBS symptoms and performance. *Indoor Air*, 14(7):74–81.
- Ferrarelli, L. K. (2016). Hunger signals suppress risk perception. *Science Signaling*, 9(458):292–293.
- Floros, C. (2011). On the relationship between weather and stock market returns. *Studies in Economics and Finance*, 28(1):5–13.
- Graff Zivin, J. S., Hsiang, S. M., and Neidell, M. J. (2015). Temperature and human capital in the short-and long-run. Technical report, National Bureau of Economic Research.
- Halleröd, B. and Larsson, D. (2008). Poverty, welfare problems and social exclusion. *International Journal of Social Welfare*, 17(1):15–25.
- Hancock, P. and Vasmatazidis, I. (2003). Effects of heat stress on cognitive performance: The current state of knowledge. *International Journal of Hyperthermia*, 19(3):355–372.
- Hedge, A. (2004). Linking environmental conditions to productivity. *Power Point presentation. Eastern Ergonomics Conference and Exposition*.

- 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.
- Hirshleifer, D. and Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance*, 58(3):1009–1032.
- Howarth, E. and Hoffman, M. S. (1984). A multidimensional approach to the relationship between mood and weather. *British Journal of Psychology*, 75(1):15–23.
- Jahedi, S., Deck, C., and Ariely, D. (2016). Arousal and economic decision making. *Journal of Economic Behavioral Organization*, 134(11):165–189.
- Kahol, K., Leyba, M. J., Deka, M., Deka, V., Mayes, S., Smith, M., Ferrara, J. J., and Panchanathan, S. (2008). Effect of fatigue on psychomotor and cognitive skills. *American Journal of Surgery*, 195(2):195–204.
- Kathryne, M. Y., Debbie, M., and Favre-Bulle, T. (2016). Predicting parole grants: An analysis of suitability hearings for California’s lifer inmates. *Federal Sentencing Reporter*, 28(4):268–277.
- Kent, S. T., McClure, L. A., Crosson, W. L., Arnett, D. K., Wadley, V. G., and Sathiakumar, N. (2009). Effect of sunlight exposure on cognitive function among depressed and non-depressed participants: A REGARDS cross-sectional study. *Environmental Health*, 8(1):34–36.
- Lambert, G., Reid, C., Kaye, D., Jennings, G., and Esler, M. (2002). Effect of sunlight and season on serotonin turnover in the brain. *Lancet*, 360(9348):1840–1842.
- Lavy, V., Ebenstein, A., and Roth, S. (2016). The impact of short term exposure to ambient air pollution on cognitive performance and human capital formation. *American Economic Journal: Applied Economics*, 8(4):36–65.
- Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, 65(3):272–292.
- Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty impedes cognitive function. *Science*, 341(6149):976–980.
- Mathot, K. J., Nicolaus, M., Araya-Ajoy, Y. G., Dingemanse, N. J., and Kempenaers, B. (2015). Does metabolic rate predict risk-taking behaviour? A field experiment in a wild passerine bird. *Functional Ecology*, 29(2):239–249.
- Oakley, J. B. and Coon, A. F. (1986). The federal rules in state courts: A survey of state court systems of civil procedure. *Washington Law Review*, 61(5):1367–1368.
- O’Brien, E. M. and Mindell, J. A. (2005). Sleep and risk-taking behavior in adolescents. *Behavioral Sleep Medicine*, 3(3):113–133.

- Page, L. A., Hajat, S., and Kovats, R. S. (2007). Relationship between daily suicide counts and temperature in England and Wales. *The British Journal of Psychiatry*, 191(2):106–112.
- Park, J. (2016). Temperature, test scores, and educational attainment. Technical report, University of Harvard Working Paper.
- Pepler, R. D. and Warner, R. (1968). Temperature and learning: An experimental study. *Transactions of the ASHRAE Annual Meeting*, 42(2):211–219.
- Pflug, B., Erikson, R., and Johnsson, A. (1976). Depression and daily temperature. *Acta Psychiatrica Scandinavica*, 54(4):254–266.
- Pocheptsova, A., Amir, O., Dhar, R., and Baumeister, R. (2009). Deciding without resources: Psychological depletion and choice in context. *Journal of Marketing Research*, 46(3):344–355.
- Ramji-Nogales, J., Schoenholtz, A. I., and Schrag, P. G. (2007). Refugee roulette: Disparities in asylum adjudication. *Stanford Law Review*, 60(2):295–411.
- Seppanen, O., Fisk, W. J., and Lei, Q. (2006). Room temperature and productivity in office work. *Proceeding of Healthy Buildings Congress*, 1(1):243–247.
- Stern, N. H. (2007). *The economics of climate change: The Stern review*. Cambridge University Press.
- Szyszkowicz, M., Willey, J. B., Grafstein, E., Rowe, B. H., and Colman, I. (2010). Air pollution and emergency department visits for suicide attempts in Vancouver, Canada. *Environmental Health Insights*, 4(1):79–86.
- Tchen, N., Juffs, H. G., Downie, F. P., Yi, Q.-L., Hu, H., Chemerynsky, I., Clemons, M., Crump, M., Goss, P. E., Warr, D., Tweedale, M., and Tannock, I. (2003). Cognitive function, fatigue, and menopausal symptoms in women receiving adjuvant chemotherapy for breast cancer. *Journal of Clinical Oncology*, 21(22):4175–4183.
- Tsutsumi, H., Tanabe, S.-i., Harigaya, J., Iguchi, Y., and Nakamura, G. (2007). Effect of humidity on human comfort and productivity after step changes from warm and humid environment. *Building and Environment*, 42(12):4034–4042.
- Viner, R. M., Clark, C., Taylor, S. J., Bhui, K., Klineberg, E., Head, J., Booy, R., and Stansfeld, S. A. (2008). Longitudinal risk factors for persistent fatigue in adolescents. *Archives of Pediatrics and Adolescent Medicine*, 162(5):469–475.
- Wan, J., Yang, K., Zhang, W., and Zhang, J. (2009). A new method of determination of indoor temperature and relative humidity with consideration of human thermal comfort. *Building and Environment*, 44(2):411–417.
- Ward, A. (2013). How temperature shapes difficult decisions. <https://www.scientificamerican.com/article/warm-weather-makes-it-hard-think-straight/>. Accessed: 2017-01-21.

- Weaver, L. J. and Hadley, C. (2009). Moving beyond hunger and nutrition: A systematic review of the evidence linking food insecurity and mental health in developing countries. *Ecology of Food and Nutrition*, 48(4):263–284.
- Weinreb, L., Wehler, C., Perloff, J., Scott, R., Hosmer, D., Sagor, L., and Gundersen, C. (2002). Hunger: Its impact on children’s health and mental health. *Pediatrics*, 110(4):41–50.
- Williams, L. E. and Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science*, 322(5901):606–607.
- Wyon, D. P., Wyon, I., and Norin, F. (1996). Effects of moderate heat stress on driver vigilance in a moving vehicle. *Ergonomics*, 39(1):61–75.
- Xiong, J., Lian, Z., Zhou, X., You, J., and Lin, Y. (2015). Investigation of gender difference in human response to temperature step changes. *Physiology and Behavior*, 151(1):426–440.
- Yu, W., Vaneckova, P., Mengersen, K., Pan, X., and Tong, S. (2010). Is the association between temperature and mortality modified by age, gender and socioeconomic status? *Science of the Total Environment*, 408(17):3513–3518.

7 Tables

Table 1: Summary Statistics

	Mean	Std. Dev.
Grant indicator	0.164	0.371
Temperature (°F)	57.37	15.721
Heat index (°F)	57.77	16.423
Air pressure (pa)	29.688	0.759
Dew point (°F)	49.372	17.202
Precipitation (mm)	0.003	0.014
Wind speed (km/h)	4.557	3.441
Sky cover (percent)	55.44	0.276
Ozone (ppm)	0.0220	0.0120
CO (ppm)	0.917	0.496
PM _{2.5} (μ/m^3)	14.957	11.569

Table 2: Fixed effect estimates: 6 AM - 4 PM average

	(1)	(2)	(3)	(4)
	Preferred	1-Day lag	1-Day lead	All
$Temperature_t$	-0.000893* [0.000230]	-0.000676* [0.000288]	-0.00103* [0.000277]	-0.000820* [0.000312]
$Temperature_{t-1}$	- -	-0.000327 [0.000214]	- -	-0.000331 [0.000213]
$Temperature_{t+1}$	- -	- -	0.000133 [0.000199]	0.000251 [0.000197]
Observations	206,924	206,924	206,924	206,924
Time FEs	Y	Y	Y	Y
City FEs	Y	Y	Y	Y
Weather	Y	Y	Y	Y
Pollutants	Y	Y	Y	Y

Notes: Dependent variable is a dummy taking value one if decision is favourable to applicant, zero otherwise. Standard errors clustered on city-week in brackets. Weather covariates include dew point, air pressure, wind speed, precipitation and cloud cover. All weather covariates are 6 AM to 4 PM average. Pollutant covariates include controls for ozone, carbon monoxide and PM25. All pollutant variables are the mean of daily values. Time fixed effects include day of week and year-month dummies. Regressions also control for city fixed effects, judge name, type of application and nationality of applicant. * significant at 10% ** significant at 5% *** significant at 1%.

Table 3: Alternative fixed effects

	(1)	(2)	(3)	(4)
$Temperature_t$	-0.000893** [0.000230]	-0.000652** [0.000280]	-0.00105** [0.000288]	-0.000961** [0.000230]
Observations	206,924	206,924	206,924	206,924
Time FEs	year-month	city-year-month	year-week	city-week
City FEs	city	N	city-week	city
Weather	Y	Y	Y	Y
Pollutants	Y	Y	Y	Y

Notes: Dependent variable is a dummy taking value one if decision is favourable to applicant, zero otherwise. Standard errors clustered on city-week in brackets. Weather covariates include dew point, air pressure, wind speed, precipitation and cloud cover. All weather covariates are 6 AM to 4 PM average. Pollutant covariates include controls for ozone, carbon monoxide and PM25. All pollutant variables are the mean of daily values. Regressions also control for judge name, type of application and nationality of applicant. * significant at 10% ** significant at 5% *** significant at 1%.

Table 4: Sensitivity analyses

	(1)	(2)	(3)	(4)	(5)
	Preferred spec.	Calendar day	Deviation from weekly avg.	City interaction	Winter exclusion
$Temperature_t$	-0.000893** [0.000230]	-0.000715** [0.000208]	-0.000534* [0.000290]	-0.000703** [0.000295]	-0.00111** [0.000283]
Observations	206,924	180,429	206,924	206,924	156,951
Time FEs	Y	Y	Y	Y	Y
City FEs	Y	Y	Y	Y	Y
City*temperature	N	N	N	Y	N
Weather	Y	Y	Y	Y	Y
Co-pollutants	Y	Y	Y	Y	Y

Notes: Dependent variable is a dummy taking value one if decision is favourable to applicant, zero otherwise. Standard errors clustered on city-week in brackets. Weather covariates include dew point, air pressure, wind speed, precipitation and cloud cover. All weather covariates are calendar daily average. Pollutant covariates include controls for ozone, carbon monoxide and PM25. All pollutant variables are the mean of daily values. Time fixed effects include day of week and year-month dummies. Regressions also control for city fixed effects, judge name, type of application and nationality of applicant. * significant at 10% ** significant at 5% *** significant at 1%.

Table 5: Non-linear estimates

	(1) Temperature	(2) Heat Index	(3) HI>65
X \leq 40	0.00232 [0.00541]	0.00216 [0.00539]	- -
X \in [40-45)	0.0115* [0.00638]	0.0112* [0.00634]	-
X \in [45-50)	0.00493 [0.00514]	0.00438 [0.00507]	- -
X \in [50-55)	- -	- -	- -
X \in [55-60)	-0.00978** [0.00446]	-0.00722* [0.00417]	- -
X \in [60-65)	-0.0109** [0.00527]	-0.00523 [0.00482]	- -
X \in [65-70)	-0.0195** [0.00667]	-0.0120** [0.00581]	- -
X \in [70-75)	-0.0185** [0.00790]	-0.00818 [0.00652]	0.00369 [0.00600]
X \in [75-80)	-0.0246** [0.0103]	-0.00385 [0.00593]	-0.00266 [0.00703]
X \in [80-85)	-0.0352** [0.0125]	-0.0269** [0.00936]	-0.0201** [0.00964]
X \in [85-90)	-0.0583** [0.0129]	-0.0309** [0.00905]	-0.0279** [0.00987]
X \in [90-95)	- -	-0.0171 [0.0132]	-0.0123 [0.0148]
X \geq 95	- -	-0.0614** [0.0274]	-0.0537** [0.0269]
Observations	206,924	206,924	67,194
Time FEs	Y	Y	Y
City FEs	Y	Y	Y
Weather	Y	Y	Y
Co-pollutants	Y	Y	Y

Notes: Dependent variable is a dummy taking value one if decision is favourable to applicant, zero otherwise. Standard errors clustered on city-week in brackets. Weather covariates include dew point (every 5-degree indicators), air pressure, wind speed, precipitation and cloud cover. All weather covariates are 6 AM to 4 PM average. Pollutant covariates include controls for ozone, carbon monoxide and PM25. All pollutant variables are the mean of daily values. Time fixed effects include day of week and year-month dummies. Regressions also control for city fixed effects, judge name, type of application and nationality of applicant. * significant at 10% ** significant at 5% *** significant at 1%.

Table 6: Heterogeneity gender of judge

	(1)	(2)	(3)
	Whole sample	Female	Male
$Temperature_t$	-0.000893** [0.000230]	-0.000966** [0.000407]	-0.000813** [0.000276]
Observations	206,924	72,229	134,695
Hausman test	3.65**		
P-value	0.0325		
Time FEs	Y	Y	Y
City FEs	Y	Y	Y
Weather	Y	Y	Y
Co-pollutants	Y	Y	Y

Notes: Dependent variable is a dummy taking value one if decision is favourable to applicant, zero otherwise. Standard errors clustered on city-week in brackets. Weather covariates include dew point, air pressure, wind speed, precipitation and cloud cover. All weather covariates are 6 AM to 4 PM average. Pollutant covariates include controls for ozone, carbon monoxide and PM25. All pollutant variables are the mean of daily values. Time fixed effects include day of week and year-month dummies. Regressions also control for city fixed effects, judge name, type of application and nationality of applicant. * significant at 10% ** significant at 5% *** significant at 1%.

Table 7: Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Preferred	Pollution exclusion	CA exclusion	Clear sky days	Zero precipitation	Zero precipitation including lag	HI	HI (>75)	Quartiles exclusion	Deciles exclusion
$Temperature_t$	-0.000893* [0.000230]	-0.000710* [0.000227]	-0.000876* [0.000369]	-0.00190* [0.000727]	-0.000879* [0.000254]	-0.000996* [0.000269]	-	-	-0.00116* [0.000360]	-0.000923* [0.000251]
$Heatindex_t$	-	-	-	-	-	-	-0.000343* [0.000139]	-0.00174* [0.000705]	-	-
Observations	206,924	206,924	135,184	13,981	157,681	111,361	206,921	29,659	108,199	156,598
Time FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
City FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Weather	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Co-pollutants	Y	N	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Dependent variable is a dummy taking value one if decision is favourable to applicant, zero otherwise. Standard errors clustered on city-week in brackets. Weather covariates include dew point, air pressure, wind speed, precipitation and cloud cover. All weather covariates are 6 AM to 4 PM average. Pollutant covariates include controls for ozone, carbon monoxide and PM25. All pollutant variables are the mean of daily values. Time fixed effects include day of week and year-month dummies. Regressions also control for city fixed effects, judge name, type of application and nationality of applicant. * significant at 10% ** significant at 5% *** significant at 1%.

Table 8: Alternative standard errors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	City-week	Year-month	City-year	City	Judge	City and week	Eicker-White
$Temperature_t$	-0.000893** [0.000233]	-0.000893** [0.000247]	-0.000893** [0.000292]	-0.000893** [0.000271]	-0.000893** [0.000214]	-0.000893** [0.000249]	-0.000893** [0.000164]
Observations	206,924	206,924	206,924	206,924	206,924	206,924	206,924
Time FEs	Y	Y	Y	Y	Y	Y	Y
City FEs	Y	Y	Y	Y	Y	Y	Y
Weather	Y	Y	Y	Y	Y	Y	Y
Co-pollutants	Y	Y	Y	Y	Y	Y	Y

Notes: Dependent variable is a dummy taking value one if decision is favourable to applicant, zero otherwise. Weather covariates include dew point, air pressure, wind speed, precipitation and cloud cover. All weather covariates are 6 AM to 4 PM average. Pollutant covariates include controls for ozone, carbon monoxide and PM25. All pollutant variables are the mean of daily values. Time fixed effects include day of week and year-month fixed dummies. Regressions also control for city fixed effects, judge name, type of application and nationality of applicant. * significant at 10% ** significant at 5% *** significant at 1%.

Table 9: Placebos

	(1)	(2)	(3)	(4)
	Preferred	+100 days	-100 days	Furthest monitor
$Temperature_t$	-0.000893** [0.000230]	-0.0000253 [0.000136]	-0.000104 [0.000124]	0.000138 [0.000133]
Observations	206,924	206,924	206,924	206,924
Time FEs	Y	Y	Y	Y
City FEs	Y	Y	Y	Y
Weather	Y	Y	Y	Y
Co-pollutants	Y	Y	Y	Y

Notes: Dependent variable is a dummy taking value one if decision is favourable to applicant, zero otherwise. Standard errors clustered on city-week in brackets. Weather covariates include dew point, air pressure, wind speed, precipitation and cloud cover. All weather covariates are 6:00 AM to 4:00 PM average. Pollutant covariates include controls for ozone, carbon monoxide and PM25. All pollutant variables are the mean of daily values. Time fixed effects include day of week and year-month dummies. Regressions also control for city fixed effects, judge name, type of application and nationality of applicant. * significant at 10% ** significant at 5% *** significant at 1%.

Table 10: Parole estimates: Calendar day

	(1)	(2)	(3)	(4)
	Preferred	1-Day lag	1-Day lead	All
$Temperature_t$	-0.000928** [0.000388]	-0.00164** [0.000675]	-0.000911* [0.000628]	-0.00180* [0.000999]
$Temperature_{t-1}$	- -	0.000837 [0.000693]	- -	0.000867 [0.000729]
$Temperature_{t+1}$	- -	- -	-0.0000195 [0.000703]	0.000151 [0.000735]
Observations	18461	18461	18461	18461
Time FEs	Y	Y	Y	Y
Institution FEs	Y	Y	Y	Y
Weather	Y	Y	Y	Y
Co-pollutants	Y	Y	Y	Y

Notes: Dependent variable is a dummy for favourable judgment. Standard errors clustered on institution-week in brackets. Weather covariates include controls for dew point, air pressure, wind speed, precipitation and cloud cover. Co-pollutant covariates include controls for ozone, carbon monoxide and PM25. Time fixed effects include day of week and year-month dummies. Regressions also control for institution fixed effect, judges' name, type of application and inmate's name. All environmental variables are the mean of daily values. * significant at 10% ** significant at 5% *** significant at 1%

Figures

Figure 1: Location of immigration courts (excluding Honolulu)



Figure 2: US immigration courts



Note: Buildings that house the 16 largest courts ranked by contribution to sample. From left-to-right, top row: New York, Los Angeles, Miami, San Francisco, Chicago, Arlington, Orlando, Baltimore, Boston, Detroit, Philadelphia, Memphis, Atlanta, Houston, San Diego, Seattle.

Figure 3: Location of hearing institutions

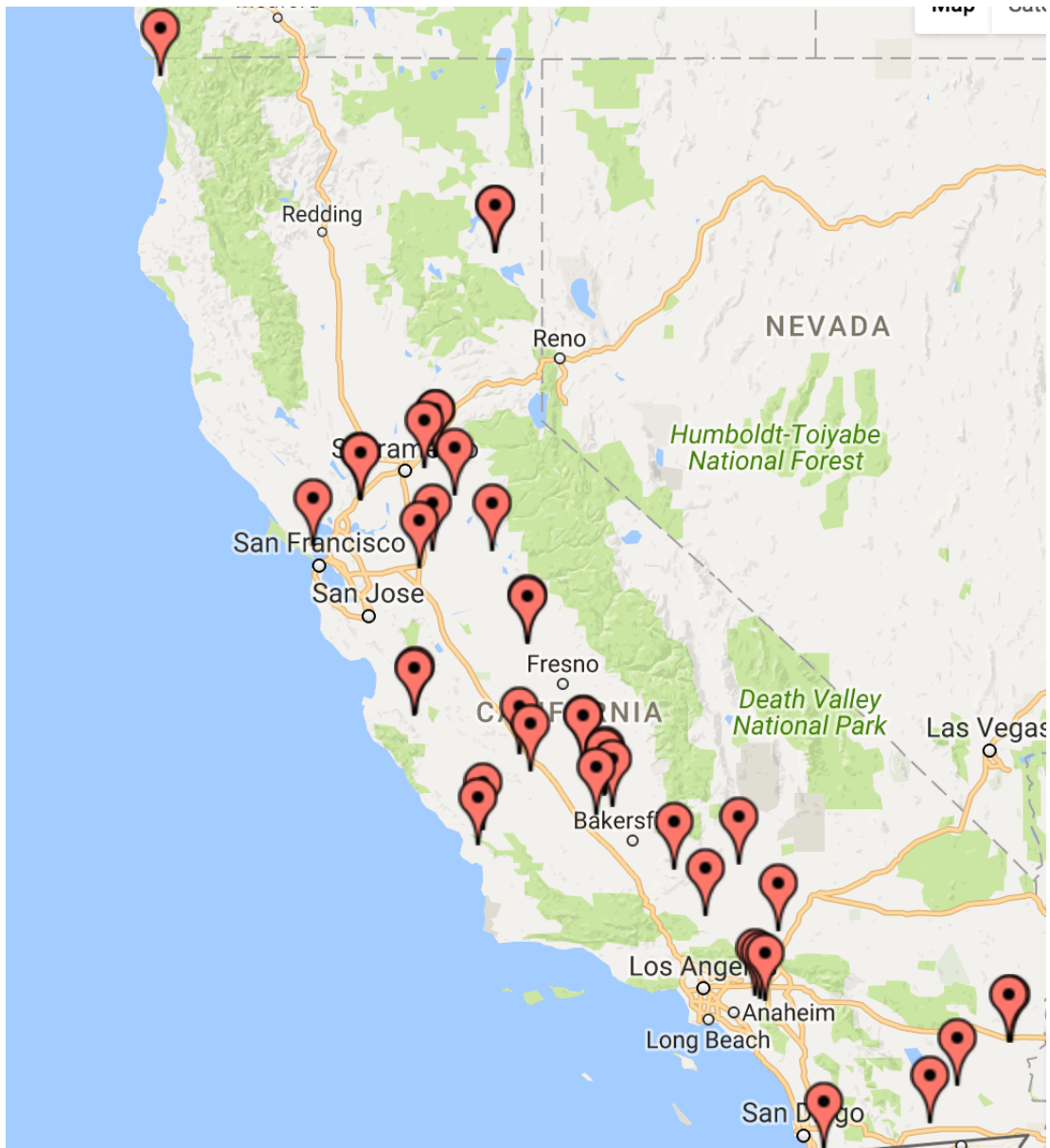


Figure 4: Distribution of 6 AM - 4 PM temperature, 2000-2004

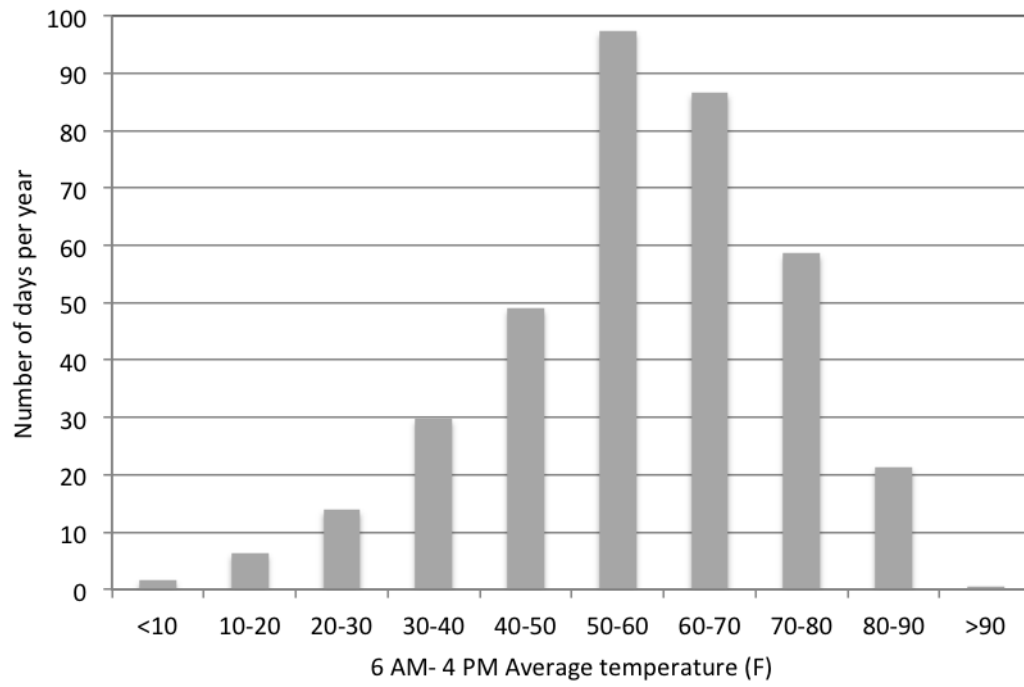
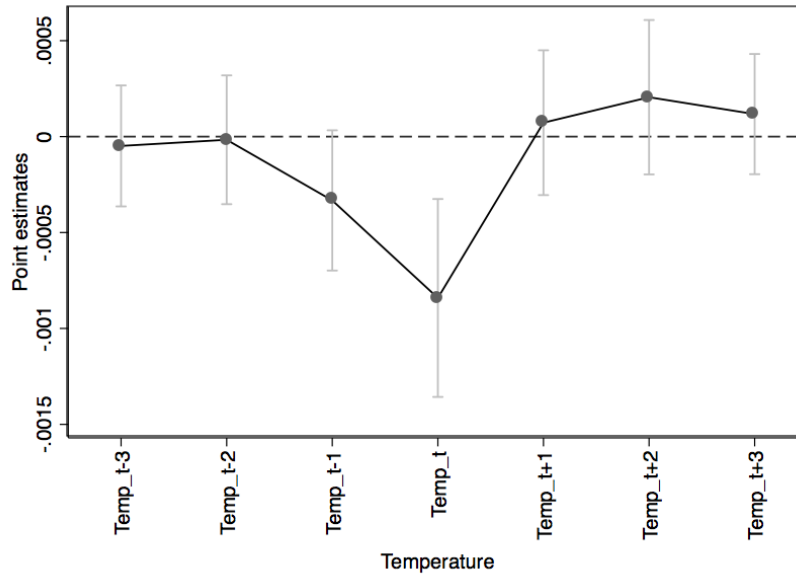
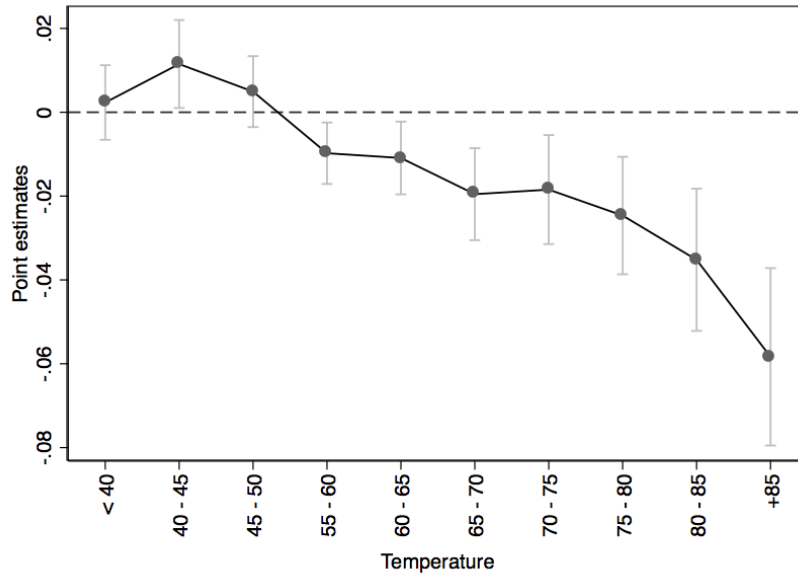


Figure 5: Timing of exposure: 6 AM - 4 PM



This figure plots the coefficients for the temperature. Gray lines show the 90 percent confidence interval based on standard errors clustered on city-week. The dependent variable is favourable judgement. The regression includes controls for temperature, sky cover, dew point, air pressure, wind speed, precipitation, ozone, carbon monoxide, $PM_{2.5}$, year-month, city, day of week, judge name, type of application and nationality of applicant dummies.

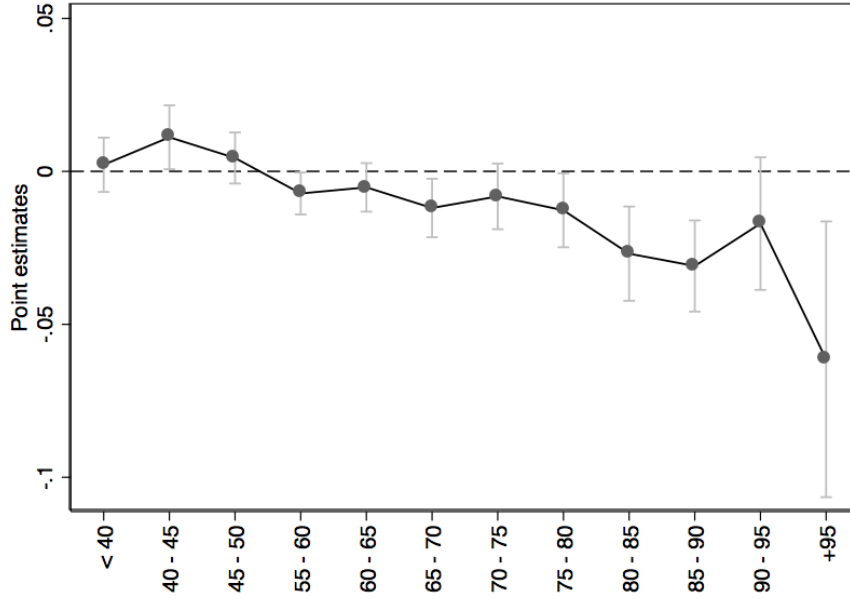
Figure 6: Non-linear estimates: Temperature, 6 AM - 4 PM



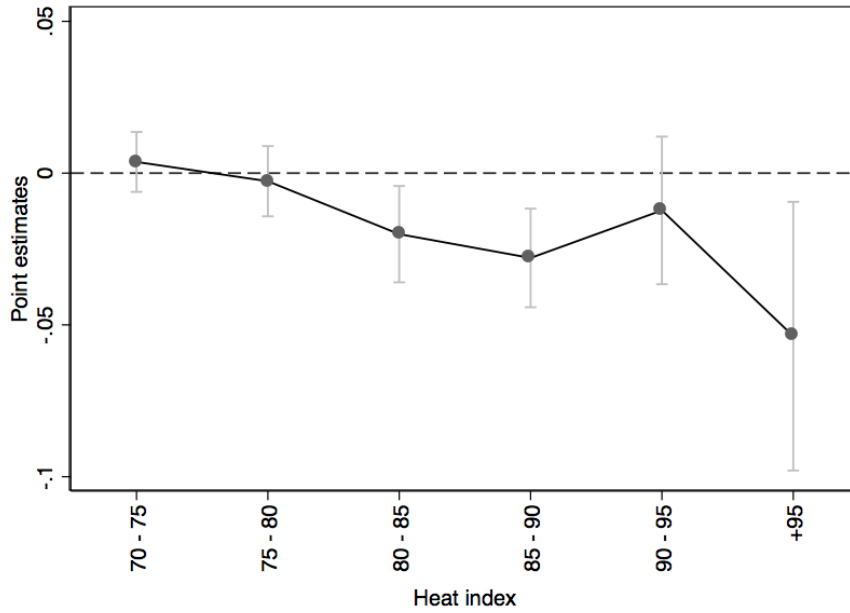
Note: This figure plots the coefficients for the temperature indicator variables. Gray lines show the 90 percent confidence interval based on standard errors clustered on city-week. The dependent variable is favourable judgement. The regression includes controls for sky cover, dew point (5 degree F indicators), air pressure, wind speed, precipitation, ozone, carbon monoxide, $PM_{2.5}$, year-month, city, day of week, judge name, type of application and nationality of applicant dummies.

Figure 7: Non-linear estimates: Heat index, 6 AM - 4 PM

(a) Whole sample

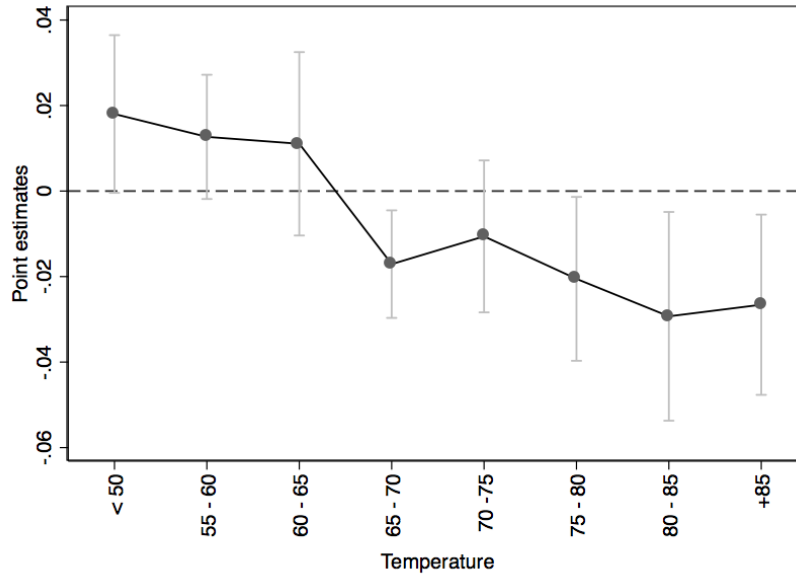


(b) HI > 65



Note: These figures plot the coefficients for the heat index indicator variables, Panel (a) depicts the whole sample and Panel (b) shows the subsample of HI > 65. Gray lines show the 90 percent confidence interval based on standard errors clustered on city-week. The dependent variable is favourable judgement. The regression includes controls for sky cover, air pressure, wind speed, precipitation, ozone, carbon monoxide, $PM_{2.5}$, year-month, city, day of week, judge name, type of application and nationality of applicant dummies.

Figure 8: Non-linear estimates: Parole, temperature, calendar day



Note: This figure plots the coefficients for temperature. Gray lines show the 90 percent confidence interval based on standard errors clustered on institution-year. The dependent variable is favourable judgement. The regression includes controls for sky cover, dew point (5 degree F indicators), air pressure, wind speed, precipitation, ozone, carbon monoxide, $PM_{2.5}$, year-month, institution, day of week, inmate name, judge name and type of application dummies.