

Will Adaptation to Climate Change be Slow and Costly? Evidence from High Temperatures and Mortality, 1900-2004*

Alan Barreca, Karen Clay, Olivier Deschenes, Michael Greenstone and
Joseph S. Shapiro

January 2015

Abstract

This paper builds on Barreca et al.'s (2013) finding that over the course of the 20th century the proliferation of residential air conditioning led to a remarkable decline in mortality due to extreme temperature days in the United States. Using panel data on monthly mortality rates of U.S. states and daily temperature variables for over a century (1900-2004) it explores the regional evolution in this relationship and documents two key findings. First, the impact of extreme heat on mortality is notably smaller in states that more frequently experience extreme heat. Second, the difference in the heat-mortality relationship between hot and cold states declined over the period 1900-2004, though it persisted through 2004. For example, the effect of hot days on mortality in cool states over the years 1980-2004, a period when residential air conditioning was widely available, is almost identical to the effect of hot days on mortality in hot states over the years 1900-1939, a period when air conditioning was not available for homes. Continuing differences in the mortality consequences of hot days suggests that health motivated adaptation to climate change may be slow and costly around the world.

JEL Classification: H4, I1, Q4, Q5

Keywords: adaptation, climate change, air conditioning, compensatory behavior, convergence, mortality

* Barreca: Department of Economics, 206 Tilton Hall, Tulane University, New Orleans, LA 70118, IZA, and NBER. Clay: Heinz College, 4800 Forbes Avenue, Carnegie Mellon University, Pittsburgh, PA 15213, and NBER. Deschenes (corresponding author): Department of Economics, 2127 North Hall, University of California, Santa Barbara, CA 93106, IZA, and NBER (email: olivier@econ.ucsb.edu). Greenstone: University of Chicago, Department of Economics, 1126 E. 59th Street, Chicago, Illinois 60637, and NBER. Shapiro: Yale University, Department of Economics, 37 Hillhouse Ave, New Haven, CT 06511, and NBER. We thank our discussant Joshua Graff-Zivin and Benjamin Olken for their comments and IZA for financial support.

A growing literature attempts to predict the impacts of climate change on agronomic, economic, and health outcomes. A key challenge in this literature is identifying the effect of climatic variables on various outcomes independently of the effects of many confounding variables that are correlated with both climate and with economic outcomes.

Most recent studies therefore derive identification from weather variability, i.e., from yearly or monthly fluctuations in weather in a given location. This panel variation permits identification of weather effects by implicitly controlling for time-invariant factors that differ across location and that are potentially correlated with climate. However, by relying on high-frequency weather variation, these studies ignore long-term adaptation to changes in climatic variables that evolve over decades or more. To the extent that more adaptation becomes available in the future, either through greater diffusion of existing technologies or through development of new technologies, predicted impacts of climate change based on short-term weather variation will likely overstate the true long-term impact of climate change. On the other hand, intensification effects may lead to an understatement of the true long-term impact of climate change (Dell, Jones, and Olken 2014). Thus more empirical studies of adaptation are needed, although a few studies now exist (Barreca et al. 2013, Boustan et al. 2012, Burke and Emerick 2013).

This paper presents a hybrid approach that combines high-frequency temperature variation with long-run climate differences to analyze two questions. First, how important is adaptation in the cross-section? For example, does extreme heat increase mortality more in areas that less often experience extreme heat? Second, do cross-sectional adaptation rates converge? For example, does the vulnerability of cold locations to extreme heat converge toward the

vulnerability of hot locations to extreme heat?

Using panel data on monthly mortality rates of U.S. states and daily temperature variables for over a century (1900-2004), we document two findings. First, the impact of extreme heat on mortality is notably smaller in states that more frequently experience extreme heat. This finding stands in contrast to previous research that provides conflicting findings on whether regions with hotter climates experience smaller effects from hot days (Dell, Jones, and Olken 2014). Second, the difference in the heat-mortality relationship between hot and cold states declined over the period 1900-2004, though it persisted through 2004.

I. Data Sources and Summary Statistics

The data for this paper are taken from Barreca et al. (2013). This section briefly describes these data and reports summary statistics. See the Online Appendix and Barreca et al. (2013) for further details.

Mortality Rate Data. The main outcome analyzed is the total mortality rate (all-age, all-cause mortality rate) for the period 1900-2004 at the state-year-month level. The total monthly mortality rate has changed substantially over time: it declined from an average of 129.4 (per 100,000 population) in the first decade of the 20th century to an average of 71.7 in the last decade. Further, there is important seasonality in the total monthly mortality rate: the average for January over 1900-2004 is 91.6 while it is only 77.1 in July. Finally, there are notable geographical differences in seasonality of monthly mortality. For example, the average July mortality rate in California is 65.9 while it is 78.9 in Illinois.

If these differences over year, month, and space are correlated with climate, associations between temperature and mortality may not represent the causal effect of temperature on mortality. We address this issue in the regression models below by leveraging the presumably

exogenous shocks to the monthly temperature distribution within states over time while controlling for other temporal and geographical factors (Deschenes and Greenstone 2007).

Temperature Data. The independent variable of central interest for this paper is a measure of extreme temperature, defined as the number of days in a state-year-month where the daily average temperature is greater than 90° F (henceforth labelled $D90$). For the continental United States as a whole, there are 0.7 such days per year on average over 1900-2004. In the analysis below we study the heterogeneity of the mortality response to $D90$ by considering how the effect of $D90$ varies across the 10 deciles of its long-term distribution, defined as the average annual number of $D90$ per state over 1900-2004.¹ Historical exposure to high temperatures varies greatly across the deciles: state in the lowest decile (CO, CT, MA, ME, MT, NH, OR, RI, VT, WA, WY) experience one high temperature day every 50 years on average, states in the fifth decile experience one every 11 years, and states in the highest decile experience (AZ, KS, NV, OK, TX) more than 1.2 high temperature day per year on average.

II. Econometric Strategy

The results are based on the fitting variants of the following regression equation:

$$(1) \log(Y_{sym}) = \sum_{d=1}^{10} \beta_d D90_{sym} \times 1(Decile_{sy}^{D90} = d) + X_{sym} \phi + \alpha_{sm} + \gamma_{ym} + \varepsilon_{sym}$$

where $\log(Y_{sym})$ is the log of the monthly mortality rate in state s , year y , and month m , and X_{sym} is a vector of control variables.² The equation also includes a full set of state-by-month fixed effects (α_{sm}) and year-month fixed effects (γ_{ym}). The state-by-month fixed effects control for all

¹ In practice, we create population-weighted deciles. Since there are 49 states in the sample, and since the variable $D90$ has a skewed distribution with many values close to zero, the distribution of states and population across deciles is lumpy. Some deciles have far fewer than 5 states or more than 10% of the national population. The shares of population across the ten deciles of the long-term distribution of $D90$ are: 0.10, 0.11, 0.10, 0.15, 0.05, 0.09, 0.10, 0.17, 0.02, and 0.10.

² The vector of control variables, X_{sym} , includes population share living in urban areas, and the share of the state population in one of four age categories: infants (0-1), 1-44, 45-64 and 65+ years old. The vector also includes indicators for unusually high or low amounts of precipitation in the current state-year-month. In the interest of space, we do not report the estimated coefficients for the variables included in X_{sym} .

unobserved state-by-month invariant determinants of the mortality rate. As a result, fixed differences across states in health care quality or the level of health capital will not confound the *D90* variables. The year-by-month fixed effects control for all factors that have contributed to changes in mortality rates over time that are common across state (e.g., the introduction of Medicare and Medicaid). The specification also includes a quadratic time trend that is allowed to vary at the state-month level to control for smooth changes in the state-by-month mortality rate across years.

The variable of central interest for this paper is *D90*, the number of days where the average daily temperature exceeds 90° F. Its coefficient β is interpreted as the effect on the log mortality rate of one day where the average temperature is greater than 90°F relative to a day where it is smaller than 90° F.³ In order to test for differences in long-term adaptation, we estimate the parameter β separately across the ten deciles of the long-term distribution (e.g., 1900-2004) of *D90*. Since equation (1) includes state-month fixed effects, the coefficient β is identified from state-month specific deviations in temperature distribution about the state-month long term averages, after adjustment for the rich set of covariates, year-month national shocks, and state-level factors that evolves according to a quadratic time trend. Due to the unpredictability of realizations of the temperature distribution across months, we presume that this variation is orthogonal to unobserved determinants of mortality rates.

Finally, we address the possibility that temperature shocks may lead to inter-temporal mortality displacement, or “harvesting,” of persons who are ill and already near death. To this end, the *D90* variables in equation (1) enter for the current month and for the previous month,

³ Estimates of the impact of temperature above 90°F on log mortality rates are very similar when all other temperature bins are included as in Barreca et al. (2013) (coefficient = 0.0098, std error = 0.0029) or as here where a single variable with the number of days with average temperature above 90°F is used to predict log mortality rates (coefficient = 0.0101, std error = 0.0033).

and we obtain the cumulative dynamic estimate by summing the estimated coefficients for each of the two months. See Barreca et al. (2013) for more details.

III. Results

Figure 1 shows how heat-related mortality varies across the cross-section of climates observed in the United States. It reports the estimated effect of *D90* days on mortality across the ten population-weighted deciles of the long-term distribution of *D90*. The deciles are calculated across the 49 states in the sample. This provides a simple test of whether the population in the 10 deciles have adapted differentially to the frequency of extreme temperature shocks. It is a compelling test because it holds national-level institutions and available technologies constant. These adaptations can occur through physical acclimatization, adoptions of costly technologies (e.g., air-conditioning, see Barreca et al. 2013), and changes in behavior (e.g., time spent indoors, see Graff-Zivin and Neidell 2014). All of these adaptations require resources, be they money or time, and return comfort, changes in mortality risks, and potentially other payoffs.

The results in Figure 1 point to a clear relationship between the effect of high temperature shocks and their historical frequency. The estimated effect of *D90* on log mortality rates in the states in the lowest decile is 0.30 (std error = 0.13), so that a single temperature day greater than 90° F increases monthly mortality rates by roughly 31% relative to all temperatures below 90° F. The estimated effect for the highest decile is strikingly smaller—it is two percent as large, at 0.0064 or 0.68% (std error = 0.0023). The evidence is consistent with long-term adaptation, i.e., states that experience the highest risks of heat-related mortality (i.e., deciles 8-10) have adapted, whereas lower exposure states (i.e., deciles 1-3) have not adapted to the same extent. One identification concern in our approach is that states with warmer climates may differ in unobserved ways from other states and these differences may themselves affect the weather-

mortality relationship directly. To the extent that such confounding variables do not covary smoothly with climate, it is encouraging that the gradient of the weather-mortality relationship with climate is largely monotone and roughly linear in Figure 1.

Over the period 1900-2004, the mortality effect of high temperatures has declined substantially for all states. As shown in Barreca et al. (2013), 1960 is a natural dividing line to study changes in the temperature-mortality relationship, since it marks the near-completion of residential electrification and the beginning of the era of residential air conditioning (AC) in the U.S. homes. Notably, the adoption of residential AC was initiated earlier and reached higher levels in states with higher historical frequency of high temperature days. In 1960 (just a few years after the beginning of the diffusion of residential AC in the U.S.), the residential AC ownership rate in the states in the lowest decile of the long-term distribution of $D90$ was 4%, compared to 27% for the states in the highest decile. By 1980, the same residential AC ownership rates were 24% and 77%, respectively.

Figure 2 explores how the differential effect of $D90$ on log mortality changed before versus after 1960 across the cross-section of climate. To this end, we re-estimated the coefficients on $D90$ in equation (1) by decile, separately by in the pre- and post-1960 periods. The evidence is consistent with convergence in adaptation, where the largest reductions in heat-related mortality have occurred in states located in the lowest deciles of the $D90$ distribution. For example in the second decile, the heat-related mortality declined by 32% after 1960. In deciles 3-5, the reduction in heat-related mortality amounts to about 10%, relative to the pre-1960 levels. Despite this important mortality risks due to high temperatures, adaptation remains incomplete since across most deciles, days with temperatures above 90° F continue to significantly predict increases in log mortality rate in the post 1960 period.

Table 1 more finely examines the timing of the change in the effect of *D90* on mortality in states where the long-term average annual days with temperature greater than 90°F is above (deciles 6-10) and below the median (deciles 1-5). Estimates in Panel A are regression coefficient estimates that can be compared to those reported in Figure 1.

Over the course of the century, the mortality effect of *D90* relative to the rest of the distribution is about 9% in low exposure states (below median) and about 1% in the high exposure states (above median). Confirming the results in Figure 2, it is apparent that the effect of hot days on mortality is highly dependent on the part of the country where the day occurs. Comparing columns (2) and (5), we observe that for both the below median and above median states, the effect of high temperatures on log mortality has declined by close to 90% over the period 1900-2004.⁴

A striking finding in Table 1 is that despite the important decline in heat-related mortality across all states, the large difference between high/low historical exposure states shown in column (1) remains in all time periods: columns (2) – (5) show that the effect of *D90* on mortality in the below median states is always roughly 10 times larger than the corresponding effect in above median states. This difference spans very different periods where the cost of self-protection varied greatly.

IV. Conclusions and Future Directions

A few lessons emerge from this paper's results. First, there have been remarkable reductions in heat-related mortality across all climates in the United States during the 20th century due to various forms of adaptation. Second, hotter parts of the country have been better protected against the mortality impacts of high temperatures in all subperiods of the 20th century. So even as the level effect of hot days on mortality has declined substantially as technology has

⁴ Appendix Table 1 reports qualitatively similar estimates over 1940-2004 when we control for per capita income.

advanced, the relative effect of a hot day has remained more dangerous in states above the median of the frequency of these days versus below it. This is perhaps best illustrated by observing that the mortality coefficient (0.0226) in the below median states for 1980-2004 period when residential AC was not a novel technology is almost identical to the same coefficient for the above median states during the 1900-1939 period (0.0245) when cooling technologies were much more expensive and AC was not available for homes. Third, the evidence is broadly consistent with the idea that self-protection is costly and communities/states and individuals are continually making judgments about its costs versus benefits. Investments in cooling technologies in relatively low *D90* exposure areas may not be worth the benefit of reduction risks of heat-related mortality. Put another way, adaptation/self-protection is subject to a cost-benefit test and communities and individuals will undertake more of it when the price goes down. Further just as zero pollution would fail a cost-benefit test, it appears that zero heat-related mortality may fail one, although the levels are getting very low in hot states.

What do this paper's results mean for climate change? At least with respect to the mortality consequences of higher temperatures, the paper paints a mixed picture. Throughout the 20th century, there have always been opportunities for individuals to protect themselves against high temperatures and they have taken advantage of those opportunities whenever the benefits were sufficiently high. At the same time, continuing differences in the mortality consequences of hot days suggest that self-protection is costly and underscore that climate change will be costly, even if its mortality impacts are mitigated by the adoption of technologies that protect people's health.

The analysis requires several important caveats. This approach in this paper cannot identify the mechanisms through which hot and cold regions have adapted to extreme heat;

Barreca et al. (2013) look into three important mechanisms. We leave additional details for future research. Additionally, the analysis of mortality in this paper hints at the potential importance of adaptation for other domains of climate change such as agriculture and sea level rise, and additional analysis of those topics would help assess the breadth of our conclusions.

REFERENCES:

Barreca, A., K. Clay, O. Deschenes, M. Greenstone and J. S. Shapiro (2013). "Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the 20th Century". NBER Working Paper No. 18692.

Boustan, L. M. Kahn, and P. Rhode. (2012): "Moving to Higher Ground: Migration Response to Natural Disasters in the Early Twentieth Century." *American Economic Review*, 102(3): 238-44.

Burke M., and K. Emerick (2013): "Adaptation to Climate Change: Evidence from US Agriculture," Mimeographed.

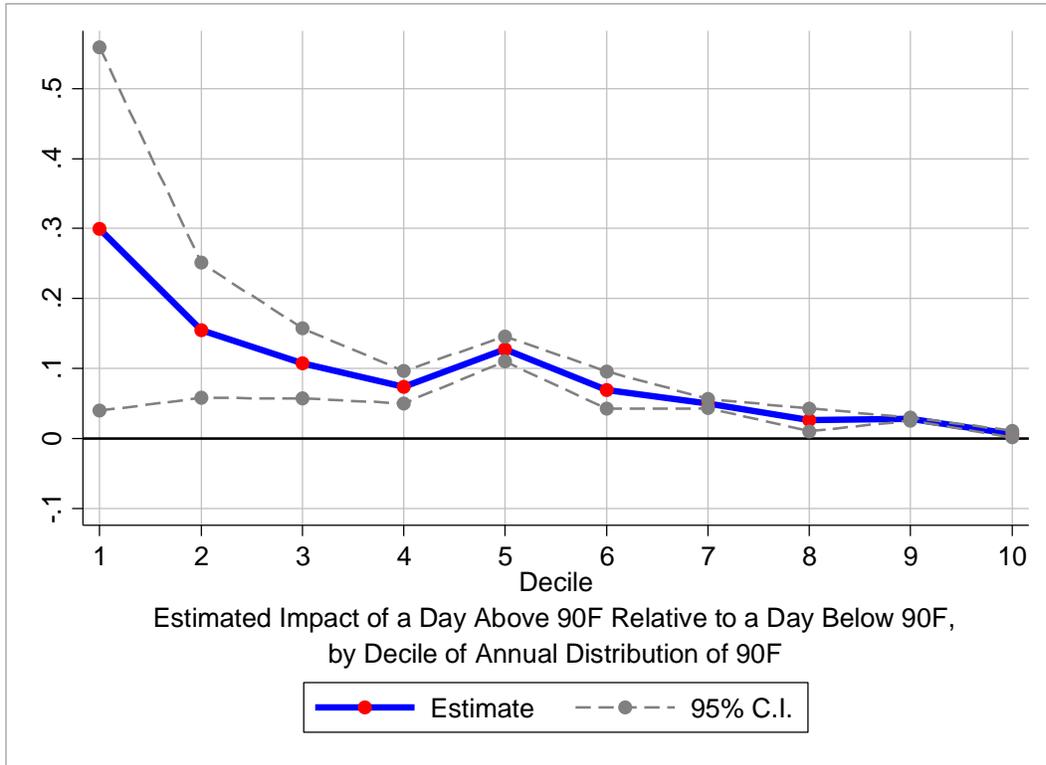
Dell, M., B. Jones, and B. Olken (2014): "What Do We Learn from the Weather? The New Climate-Economy Literature," *Journal of Economic Literature*, 52(3): 740-798

Deschenes, O. and M. Greenstone. (2007): "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather." *American Economic Review*, 97(1): 354-385.

Deschenes, O. and M. Greenstone (2011). "Climate Change, Mortality, and Adaptation: Evidence From Annual Fluctuations in Weather in the US." *American Economic Journal: Applied Economics* 3(4): 152-185

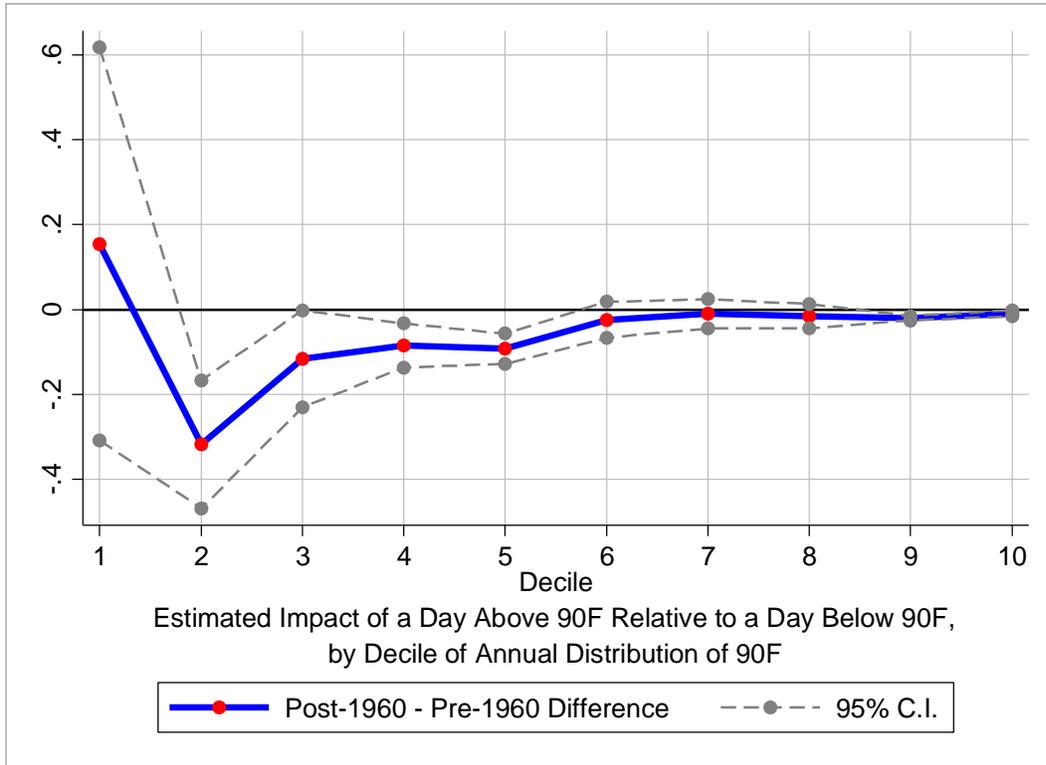
Graff Zivin, J., and M., Neidell (2014): "Temperature and the allocation of time: Implications for climate change." *Journal of Labor Economics*, 32(1): 1 - 26.

Figure 1. Estimated Effect of Days with Temperature Greater Than 90°F, by Decile of its Long-Term Distribution, 1900-2004



Notes: The dependent variable is the log monthly mortality rate. Cumulative dynamic effects for temperature exposure windows of 2 months are reported. Regressions are weighted by state population and include age group specific population shares interacted with month, urban population share interacted with month, year×month fixed effects, state×month fixed effects, state×month quadratic time trends. Standard errors clustered on state.

Figure 2. Change in Estimated Effect of Days with Temperature Greater Than 90°F, by Decile of its Long-Term Distribution, Post 1960– Pre 1960 Difference



Notes: The dependent variable is the log monthly mortality rate. Cumulative dynamic effects for temperature exposure windows of 2 months are reported. Regressions are weighted by state population and include age group specific population shares interacted with month, urban population share interacted with month, year×month fixed effects, state×month fixed effects, state×month quadratic time trends. Standard errors clustered on state.

Table 1: Estimated Effect of Days with Temperature Greater Than 90°F, by Median of the Long-Term Distribution of Annual Days with Temperature Greater Than 90°F, 1900-2004

	<u>Time Period:</u>				
	<u>1900-04</u>	<u>1900-39</u>	<u>1940-59</u>	<u>1960-79</u>	<u>1980-04</u>
<u>A. Estimated Effect of Number of Days Above 90° F</u>	(1)	(2)	(3)	(4)	(5)
States Below Median in Long-Term Average Annual D90	0.0885*** (0.0168)	0.1305*** (0.0215)	0.1545*** (0.0382)	0.0961 (0.0641)	0.0226 (0.0171)
States Above Median in Long-Term Average Annual D90	0.0099** (0.0033)	0.0245*** (0.0042)	0.0146*** (0.0034)	0.0067*** (0.0016)	0.0026* (0.0010)
Difference	-0.0785*** (0.0171)	-0.1060*** (0.0208)	-0.1399*** (0.0386)	-0.0893 (0.0647)	-0.0200 (0.0175)
<u>B. Predicted Annual Deaths due to Days Above 90 °F</u>					
States Below Median	269.4	409.1	357.5	115.8	112.0
States Below Median, at 2004 Population	446.4	892.5	668.7	141.8	126.6
State Above Median	952.1	738.8	924.4	533.3	627.5
State Above Median at 2004 Population	1,935.5	2,590.2	3,191.7	1,350.0	821.3

Notes: The dependent variable is the log monthly mortality rate. Cumulative dynamic effects for temperature exposure windows of 2 months are reported. Regressions are weighted by state population and include age group specific population shares interacted with month, urban population share interacted with month, year×month fixed effects, state×month fixed effects, state×month quadratic time trends. Standard errors clustered on state.

This Appendix discusses the data sources, sample construction and presents additional results.

A1. Description of Data

i. Vital Statistics Data

Vital Statistics Data. The data used to construct mortality rates at the state-year-month level for the 1900-2004 period come from multiple sources. For the years prior to 1959, state-year-month death counts were digitized from of the Mortality Statistics of the United States annual volumes. From 1959 to 2004, the mortality data come from the Multiple Cause of Death (MCOB) files. These data have information on state and month of death for the universe of deaths in the United States. However, geographic information on state of residence is not available in the public domain MCOB files starting after 2004, which explains why we limit our sample to the years up to 2004.

We combine the mortality counts with estimated population to derive a monthly mortality rate (per 100,000 population). Population counts are obtained from two sources. For the pre-1968 period, we linearly interpolate population estimates using the decennial Census (Ruggles et al. 2010). For the years 1969 through 2004, we use state-year population estimates from the National Cancer Institute (2008). State-year population totals are also used to weight the regressions.

ii. Weather Data

The weather data are drawn from the National Climatic Data Center (NCDC) Global Historical Climatology Network-Daily (GHCN-Daily), which is an integrated database of daily climate summaries from land surface stations that are subjected to a common set of quality assurance checks. According to NCDC, GHCN-Daily contains the most complete collection of U.S. daily climate summaries available. The key variables for the analysis are the daily maximum and minimum temperature as well as the total daily precipitation.

To construct the monthly measures of weather from the daily records, we select weather stations that have no missing records in any given year. On average between 1900 and 2004 there are 1,800 weather stations in any given year that satisfy this requirement, with around 400 stations in the early 1900s and around 2,000 stations by 2000. The station-level data is then aggregated to the county level by taking an inverse-distance weighted average of all the measurements from the selected stations that are located within a fixed 300km radius of each county's centroid. The weight given to the measurements from a weather station is inversely proportional to the squared distance to the county centroid, so that closer stations are given more weight. Finally, since the mortality data are at the state-year-month level, the county-level variables are aggregated to the state-year-month level by taking a population-weighted average over all counties in a state, where the weight is the county-year population. This ensures that the state-level temperature exposure measure correspond to population exposure, which reduces measurement error and attenuation bias.

From these data, we construct the *D90* variable, which measures the number of days in a given state-year-month where daily average temperature is above 90°F. Daily average temperature is constructed by taking the simple average of the daily minimum and maximum temperatures. The GHCN-Daily data is also used to construct the control variables $LOWP_{sym}$ and $HIGHP_{sym}$, representing unusually high or low amounts of precipitation in the current state-year-month. These are defined as indicators for realized monthly precipitation that is less than the 25th ($LOWP_{sym}$) or more than the 75th ($HIGHP_{sym}$) percentiles of the 1900-2004 average monthly precipitation in a given state-month. In the interest of space, we do not report the estimated coefficients associated with these variables

iii. Other Data

In the specification of Appendix Table 1 we also control for per capita income at the state-year level. Per capita income is only available for 1929 onwards (Bureau of Economic Analysis 2012).

Appendix References:

Bureau of Economic Analysis. (2012). Personal Income Summary. Retrieved May 22, 2012, from <http://www.bea.gov/iTable/iTable.cfm?ReqID=70&step=1&isuri=1&acrdn=4>

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series: Version 5.0. Minneapolis: University of Minnesota, 2010.

National Cancer Institute, U.S. Population Estimates 1968-2005. Downloaded March 2008 from [/http://seer.cancer.gov/popdata/download.htmlS](http://seer.cancer.gov/popdata/download.htmlS).

Appendix Table 1: Estimated Effect of Days with Temperature Greater Than 90°F, by Median of the Long-Term Distribution of Annual Days with Temperature Greater Than 90°F, Controlling for Per Capita Income, 1940-2004

	<u>Time Period:</u>				
	<u>1940-04</u>	<u>1900-39</u>	<u>1940-59</u>	<u>1960-79</u>	<u>1980-04</u>
<u>A. Estimated Effect of Number of Days Above 90° F</u>	<u>(1)</u>	<u>(2)</u>	<u>(3)</u>	<u>(4)</u>	<u>(5)</u>
States Below Median in Long-Term Average Annual D90	0.0622*** (0.0112)	---	0.1454*** (0.0179)	0.0218 (0.0560)	0.0406** (0.0141)
States Above Median in Long-Term Average Annual D90	0.0061** (0.0019)	---	0.0127* (0.0050)	0.0072*** (0.0019)	0.0033*** (0.0009)
Difference	-0.0561*** (0.0115)	---	-0.1327*** (0.0183)	-0.0146 (0.0565)	-0.0373* (0.0143)

Notes: The dependent variable is the log monthly mortality rate. Cumulative dynamic effects for temperature exposure windows of 2 months are reported. Regressions are weighted by state population and include age group specific population shares interacted with month, urban population share interacted with month, year×month fixed effects, state×month fixed effects, state×month quadratic time trends. Standard errors clustered on state.