Convergence in Adaptation to Climate Change:
Evidence from High Temperatures and Mortality, 1900-2004

Online Appendix

Alan Barreca,
Tulane University, IZA, and NBER

Karen Clay
Carnegie Mellon University and NBER

Olivier Deschenes*
University of California, Santa Barbara, IZA and NBER

Michael Greenstone
University of Chicago and NBER

Joseph Shapiro
Yale University and NBER

January 2015

Table of Contents: **Page:**

---

* Deschenes (corresponding author): Department of Economics, 2127 North Hall, University of California, Santa Barbara, CA 93106, IZA, and NBER (email: olivier@econ.ucsb.edu).

We thank our discussant Joshua Graff-Zivin and Benjamin Olken for their comments and IZA for financial support.

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 (MCOD) 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 in not available in the public domain MCOD 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 2 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).

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.

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, 1900-2004

	Time Period:				
	1900-04	**1900-39**	**1940-59**	**1960-79**	**1980-04**
A. Estimated Effect of Number of Days Above 90° F	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
States Below Median in Long-Term Average Annual D90	0.0885***	0.1305***	0.1545***	0.0961	0.0226
	(0.0168)	(0.0215)	(0.0382)	(0.0641)	(0.0171)
States Above Median in Long-Term Average Annual D90	0.0099**	0.0245***	0.0146***	0.0067***	0.0026*
	(0.0033)	(0.0042)	(0.0034)	(0.0016)	(0.0010)
Difference	-0.0785***	-0.1060***	-0.1399***	-0.0893	-0.0200
	(0.0171)	(0.0208)	(0.0386)	(0.0647)	(0.0175)
B. Predicted Annual Deaths due to Days Above 90 °F					
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.

Appendix Table 2: 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

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