APPENDIX - FOR ONLINE PUBLICATION ONLY

Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US*

By Olivier Deschenes and Michael Greenstone

A. DATA APPENDIX

I. Hadley 3 Model

We downloaded the Hadley Climate Model 3 (Hadley 3) data from the British Atmospheric Data Centre (http://badc.nerc.ac.uk/home/), which provides a wealth of atmospheric data for scientists and researchers. Hadley Centre data appears on BADC thanks to the Climate Impacts LINK Project, a distributor of archived climate model output to researchers. The future climate predictions generated by the Hadley Centre were initially prepared for the International Panel on Climate Change's (IPCC) Special Report on Emissions Scenarios (SRES). Daily climate predictions generated by the Hadley 3 model are available for all years from 1990 through 2099 and for several climate variables – we downloaded the predicted maximum and minimum temperatures and precipitation levels for each day during the years 1990-2099.

The Hadley 3 grid spans the entire globe; latitude points are separated by 2.5°, and longitude points are separated by 3.75°. We use the 89 gridpoints that fall on land in the contiguous United States to develop climate predictions for each county in the United States. Following the procedure used to create a complete daily temperature record for each US county between 1968 and 2002, we use inverse-distance weighted average to assign grid point predictions to counties. All grid points located in a pre-specified radius of a county's centroid are used to impute the climate prediction, with measurements from grid points located further away from the centroid receiving less weight. A radius of 300 kilometers ensures that every county gets a valid Hadley 3 prediction for every day between 1990 and 2099.

II. National Center for Atmospheric Research's Community Climate System Model 3

We downloaded the NCAR Community Climate System Model (CCSM) 3 data from the World Climate Research Programme's Coupled Model Intercomparison Project's data portal (https://esg.llnl.gov:8443/index.jsp), which aims to organize a variety of past, present, and future climate data from models developed across the world for use by researchers. Daily climate predictions generated by the CCSM3 model are available for all future years from 2000 to 2099 and for several climate variables – we downloaded the predicted mean temperatures and precipitation levels for each day during the years 2000-2099.

The CCSM3 grid spans the entire globe; latitude and longitude points are both separated by 1.4°. We use the 416 gridpoints that fall on land in the contiguous United States to develop climate predictions for the contiguous United States. Again, we use inverse-distance weighted average to assign grid point predictions to counties. All grid points located in a pre-specified radius of a county's centroid are used to impute the climate prediction, with measurements from grid points located further away from the centroid receiving less weight. A radius of 200 kilometers ensures that every county gets a valid CCSM3 prediction for every day between 2010 and 2099.

III. Emissions Scenarios

We utilize predictions from the application of the A1FI scenario to the Hadley 3 model. This scenario assumes rapid economic growth (including convergence between rich and poor countries) and a continued heavy reliance on fossil fuels. Given the abundant supply of inexpensive coal and other fossil fuels, a switch to alternative sources is unlikely without greenhouse gas taxes or the equivalent, so this is a reasonable benchmark scenario. This scenario assumes the highest rate of greenhouse gas emissions, so it is a worst case.

We also present results from the application of the A2 scenario to the CCSM 3. This scenario assumes slower per capita income growth but larger population growth. Here, there is less trade among nations and the fuel mix is determined primarily by local resource availability. This scenario is characterized as emphasizing regionalism over globalization and economic development over environmentalism. It is "middle of the road" in terms of greenhouse gas emissions, but it is still a business as usual scenario, because it does not reflect policies to restrict emissions.

IV. EIA Energy Consumption Data

The consumption data is derived from several different reports and forms depending on energy source. Coal consumption data for most sectors comes from the EIA's Annual Coal Report; electric power sector coal use is the exception, coming instead from forms EIA-906 "Power Plant Report" and EIA-920 "Combined Heat and Power Plant Report". Natural gas consumption data comes from the EIA's Natural Gas Annual. Most petroleum data is the "product supplied" data found in EIA's Petroleum Supply Annual, with the exception again of electric power sector use, which is reported on EIA-906 and EIA-920. Solar, wind, geothermal, and most biomass energy use data is also reported on those forms. Residential and commercial use of biomass is reported on forms EIA-457 "Residential Energy

Consumption Survey" and "Commercial Buildings Energy Consumption Survey". Nuclear electric power and other electricity data comes from the EIA Electric Power Annual. Finally, system energy losses and interstate flow are estimated in the State Energy Data System.

V. Sample Constructions Rules for Weather Data

The weather data are drawn from the National Climatic Data Center (NCDC) Summary of the Day Data (File TD-3200). The key variables for our analysis are the daily maximum and minimum temperature as well as the total daily precipitation. To ensure the accuracy of the weather readings, we developed a rule to select the weather stations. Specifically, we dropped all weather stations at elevations above 7,000 feet because they were unlikely to reflect the weather experienced by the majority of the population within a county. Among the remaining stations, we considered a year's readings valid if the station provided measurements for all days in that year. The average annual number of stations with valid data in this period was 2,837 and a total of 6,969 stations met our sample selection rule for at least one year during the 1968-2002 period. The acceptable station-level data is then aggregated at the county level by taking an inverse-distance weighted average of all the valid measurements from stations that are located within a 200 km radius of each county's centroid. The valid measurements from acceptable stations are weighted by the inverse of their squared distance to the centroid so that more distant stations are given less weight. This procedure yields a balanced panel of 3,074 counties with acceptable weather data that accounts for 99% of all deaths in the US from 1968 to 2002.

Appendix Table 1: Average Annual Mortality Rates Per 100,000 Population, 1968-2002

Age Group:	All Causes of Death (1)	Cardiovascular Disease (2)	Neoplasms (3)	Respiratory Disease (4)	Motor-Vehicle Accidents (5)
Infants	1,161.8	23.7	4.6	51.7	---
1-44	114.0	14.3	15.3	2.8	1.3
45-64	864.4	332.7	278.7	38.2	13.3
65+	5,210.2	2,787.9	1,069.5	420.2	17.2
All Ages	877.1	403.0	191.3	59.5	5.6

Notes: Averages are calculated for a sample of 107,310 county-year observations. All entries are weighted averages, where the weight is population in the relevant demographic group in a county-year. The ICD-9 codes corresponding to the causes of deaths are defined as follows: Neoplasms = 140-239, Cardiovascular Disease = 390-459, Respiratory Disease = 460-519, Motor Vehicle Accidents = E810-E819.

Appendix Table 2: Estimates of the Impact of Extreme Temperatures on Annual Mortality Rate, By Cause of Death

	Impact on Annual Mortality Rate Per 100,000:			
	Days < 10° F	Days 10°-20° F	Days 80°-90° F	Days >90° F
All Causes of Death	**0.692***	**0.587***	**0.330**	**0.936***
	(0.323)	**(0.251)**	**(0.207)**	**(0.289)**
Cardiovascular Disease	0.770*	0.294*	0.129	0.395*
	(0.181)	(0.149)	(0.100)	(0.138)
Neoplasms	-0.023	0.090	-0.032	0.017
	(0.085)	(0.068)	(0.043)	(0.076)
Respiratory	-0.100	-0.066	0.009	0.077
	(0.058)	(0.041)	(0.025)	(0.051)
MVA	-0.043*	-0.032*	-0.014*	-0.021
	(0.014)	(0.012)	(0.006)	(0.012)

Notes: The estimates are from fixed-effect regressions by age group weighted by the population count in the relevant age group in a county-year (see equation 5). The age-group specific estimates are combined into an "all-age" estimate by taking a weighted sum of the age-specific estimates, where the weight is the average population in each age group. For each group there are 107,310 county-year observations over the 1968-2002 period. The dependent variable is the annual all-cause and cause-specific mortality rate in the relevant age group in a county-year. Temperature exposure is modeled with 9 temperature-day bins defined as the number of days in a given temperature category in a county-year. The estimates on the lowest two (coldest) and highest two (hottest) bins are reported. Other control variables include a set of 11 indicator variables capturing the full distribution of annual precipitations. Standard errors are clustered at the county-by-age group level. The ICD-9 codes corresponding to the causes of deaths are defined as follows: Neoplasms = 140-239, Cardiovascular Disease = 390-459, Respiratory Disease = 460-519, Motor Vehicle Accidents = E810-E819.