ADAPTATION AND THE MORTALITY EFFECTS OF TEMPERATURE ACROSS
U.S. CLIMATE REGIONS

Garth Heutel
Nolan H. Miller
David Molitor

Adaptation and the Mortality Effects of Temperature Across U.S. Climate Regions
Garth Heutel, Nolan H. Miller, and David Molitor
NBER Working Paper No. 23271
March 2017, Revised July 2018
JEL No. I18,J14,Q54

ABSTRACT

Using 20 years of data on Medicare beneficiaries, we predict the end-of-century mortality effects of climate change among the U.S. elderly, accounting at the ZIP code level for both adaptation and regional heterogeneity in the temperature-mortality relationship. We find that this relationship varies systematically with current climate, reflecting both beneficial adaptation to frequently-experienced temperatures and harmful de-adaptation to infrequently experienced temperatures. We model adaptation to future climate using the estimated cross-sectional heterogeneity in temperature effects and combine this with projected end-of-century changes in local temperature distributions to make mortality predictions. Incorporating regional heterogeneity into these predictions reverses the conventional wisdom on the regional distribution of climate change effects: cold places bear more, not less, of the burden. Incorporating adaptation yields mortality effects of climate change that are much lower than those estimated without incorporating adaptation.

Garth Heutel
436 Andrew Young School
Department of Economics
Georgia State University
PO Box 3992
Atlanta, GA 30302-3992
and NBER
gheutel@gsu.edu

Nolan H. Miller
College of Business
University of Illinois
4033 BIF
515 East Gregory Drive
Champaign, IL 61820
and NBER
nmiller@illinois.edu

David Molitor
University of Illinois at Urbana-Champaign
340 Wohlers Hall
1206 S. Sixth Street
Champaign, IL 61820
and NBER
dmolitor@illinois.edu

I. INTRODUCTION

The broad consensus among climate scientists is that global temperatures are rising, driven by human activities that generate greenhouse gases (IPCC, 2014). The prospect of continued warming over the coming decades has concentrated attention on understanding the impact of climate change on human well-being, especially mortality, and on developing strategies to moderate its harmful effects. Optimal policy decisions regarding costly strategic responses rely on accurate estimates of these effects and the extent to which adaptation can offset them. Because both climate change and many adaptation initiatives take place at local levels, it is important to accurately estimate the economic damage of climate change both nationally and regionally.

A central empirical challenge in estimating the effects of climate change is that climate is a long-run distribution of weather outcomes and does not vary exogenously in most settings. To address this challenge, one common strategy is to estimate economic damages as a function of weather and then calculate climate damages using shifts in the future weather distribution predicted by climate models (Deschênes and Greenstone, 2011). In the most comprehensive application of this approach, Hsiang et al. (2017) estimate that excess mortality will account for about 70 percent of end-of-century climate damages in the United States but find that northern, cooler regions will generally bear lower mortality costs or even benefit from climate change relative to warmer regions. However, these climate damage estimates and their qualitative implications rely on the strong assumption that the relationship between temperature and mortality is uniform across regions and constant over time. That is, they fail to allow for regional heterogeneity in the current temperature-mortality relationship or adaptation to future changes in climate.

In this paper, we estimate local and national end-of-century impacts of climate change on U.S. elderly mortality that permit both regional heterogeneity and future adaptation. Our analysis uses a novel combination of national Medicare administrative data and weather monitor readings from the National Oceanic and Atmospheric Administration (NOAA) at the daily, ZIP code level for the period 1992 – 2011. The result is a dataset with unprecedented temporal and geographic scope and granularity, allowing us to explore the effects of temperature in ways that previously were not possible. We begin by first estimating for each U.S. ZIP code the mortality effects of both hot and cold temperatures and how these effects vary across regions based on their current climate. To predict the mortality effects of climate change, we combine the estimated climate-specific treatment effects with ZIP code-specific predictions of end-of-century temperature increases. We model adaptation to future climate based on how temperature effects currently vary across climate regions. Thus, if the future climate in Minneapolis is predicted to be similar to Chicago's current climate, our approach assumes that Minneapolis will adapt to future climate just as Chicago has adapted to their current climate. To demonstrate the importance of incorporating geographic heterogeneity

and future adaptation into climate damage predictions, we benchmark our results to predicted damages under the common assumptions that the temperature-mortality relationship neither varies across regions nor changes over time.

We present two main sets of results. The first is a non-parametric analysis aimed at establishing the importance of accounting for regional heterogeneity in mortality predictions. We build on the panel approach of Deschênes and Greenstone (2011) and exploit exogenous variation in daily temperature across years in a given location (ZIP code) to identify the causal effects of temperature on elderly mortality. We estimate the potentially non-linear effects of daily temperature in 5°F increments from <10°F to 95+°F and allow these effects to vary flexibly based on whether the ZIP code is in the coolest, middle, or warmest third of ZIP codes based on historical climate.

We find substantial heterogeneity in temperature effects across the climate terciles. The effects of hot days are much larger in cool regions than in warm regions. In the coolest third of ZIP codes, where hot days (85-90°F) occur on average about once every five years, a hot day (85-90°F) increases mortality by 3.3 deaths per 100,000 Medicare beneficiaries relative to a moderate day (65-70°F). In contrast, in the hottest regions, which average over 15 hot days per year, a hot day (85-90°F) results in 0.1 fewer deaths per 100,000 than a moderate day (65-70°F). Conversely, cold days are much less harmful in cool regions than in warm regions. We interpret this geographic heterogeneity as evidence of historical adaptation. As one potential channel for this adaptation, we find that nearly all of the cross-sectional differences in heat-related deaths can be explained by differential residential air conditioning (AC) adoption.[1]

Our analysis reveals that while regions have become adapted to temperatures they face frequently, neither warm nor cool regions are well-adapted to temperatures they encounter infrequently. This pattern is consistent with the fact that many adaptive technologies, such as heating, air conditioning, or snow-removal, protect specifically against heat or cold, but not both. Because of budget constraints, individuals and local governments will tend to invest in adaptive infrastructure that protects against the most likely weather events. This suggests that regions with moderate climates that experience both very hot and very cold days may have the most difficulty mitigating the harmful effects of extreme temperatures. Moreover, these patterns indicate that attempts to account for adaptation to hot weather under a warming climate also must account for the potential for regions to simultaneously de-adapt to cold weather.

After establishing non-parametrically the importance of regional heterogeneity and adaptation in mortality predictions, we incorporate these factors into our second set of results, which comprise the contribution of the paper. Here, we predict the mortality impact of projected end-of-century (2080-2100)

---

[1] This result complements the finding of Barreca et al. (2016) that AC adoption has been a major contributor to the decline in heat-related mortality over time by suggesting that remaining gaps in AC penetration continue to drive heat-related mortality.

climate change using a tractable model of the temperature-mortality relationship that accounts for current heterogeneity and allows for future adaptation. We estimate future climate damages for each ZIP code by combining localized temperature-mortality estimates with localized predictions of how daily temperature distributions will shift in the future, as derived from all 21 climate models for which daily scenarios are produced and distributed as part of the Coupled Model Intercomparison Project Phase 5 (CMIP5). To allow temperature-mortality effects to vary continuously by temperature and local climate, we estimate the daily temperature-mortality relationship as a semi-parametric, smooth function of both daily average temperature and the region's climate. As with our non-parametric approach, this function captures how temperature effects currently vary across climate regions. Importantly, because the estimated mortality effects of temperature in this calculation are a function of climate, we can model adaptation by basing each ZIP code's future mortality curve on its predicted future climate. This approach captures the mortality effects of climate change under adaptation to the extent that regions will adapt to a given future climate just as regions that currently experience that climate have historically adapted.[2]

We find that accounting for climate heterogeneity and adaptation when predicting end-of-century mortality impacts due to climate change substantially influences the sign, magnitude and geographic distribution of predicted damages relative to existing methods. Our preferred estimates, which incorporate climate heterogeneity and future adaptation, suggest an overall decrease in U.S. elderly mortality of approximately 0.62% by the end of the century, compared to a predicted increase in mortality of 0.87% when we use standard methods that ignore heterogeneity and adaptation. Since we do not model the costs of adaptation or the non-mortality effects of climate change on welfare, this finding does not imply that climate change will improve elderly well-being. Nonetheless, this finding quantifies the scope for adaptation to future climate change, and finds that when allowing for adaptation, average mortality effects are lower but cool regions still face mortality increases.

Underlying the overall decrease in expected mortality is important regional heterogeneity. We find that the coolest regions suffer the most from climate change, with a predicted increase in mortality of 1.37% by the end of century. This is due to the fact that these regions are currently poorly adapted to dealing with heat. Although future adaptation will mitigate the future effects of heat somewhat, the fact that these regions will have to deal with both extreme heat and cold in the future prevents adaptation from mitigating the mortality impact of climate change as much as in the warmer regions. In contrast, we predict the hottest regions will experience a 2.75 percent decrease in elderly mortality by the end of the century, due to these regions being well adapted already to dealing with heat and to the decreased number of extremely cold days they will experience in the future, which allows them to further concentrate on dealing with heat. Our

---

[2] Although we do not present these results, our parameterization of the temperature-mortality relationship would allow us to compute mortality predictions under various partial adaptation scenarios.

finding that the mortality impact of climate change will be worst in the coolest regions and best in the hottest ones is in contrast to the findings of Houser et al. (2014) and in Hsiang et al. (2017). However, these papers both assume a single, homogeneous temperature-mortality relationship that applies across the entire U.S. and do not account for future adaptation and find that the mortality effects of warming will harm the hottest areas the most.

This paper contributes to a growing literature that explores adaptation to climate change in the context of the mortality effect of temperatures. As surveyed by Hondula et al. (2015), much of this work is in public health, environmental health, and epidemiology. Among the approaches they catalog, the one most similar to ours is the "analog city" approach, where a city's future temperature-mortality relationship is assumed to resemble that of a city that currently experiences that city's future climate. The success of such a method depends on identifying an appropriate analog. The survey by Hsiang (2016) refers to a generalized version of this approach as "time-series variation with stratification," in which weather effects may vary according to historical frequency of weather events or other regional characteristics that may drive or result from adaptation. Our study is in this vein, although the temporal granularity and spatial coverage of our data allow us to fully estimate the temperature-mortality relationship and its dependence on climate at a nationally representative level. Methodologically, the studies most closely related to ours are Butler and Huybers (2013) and Auffhammer (2017), which use a similar approach to consider regional adaptation in maize production and in energy use, respectively. In a contemporaneous working paper, Portnykh (2017) considers adaptation in mortality using Russian data.[3]

The main goal of this paper is to generate end-of-century mortality predictions that take into account past and future adaptation. To do so requires estimating the entire temperature-mortality relationship as a function of climate. While some existing studies have analyzed part of the relationship between temperature, mortality and climate, ours is the first to do so for the full range of temperatures in a parameterized way that can be easily fed into climate-change models to generate predictions. For example, Deschênes and Greenstone (2001) group states into Census Divisions. While they find heterogeneity across regions, they are unable to find evidence that this heterogeneity is driven by adaptation to current climate, i.e., that usually-hot areas are less sensitive to heat and usually-cold areas are less sensitive to cold. Barreca et al. (2016) conduct a heterogeneity analysis based on groups of contiguous states with broadly similar

---

[3] Kahn (2015) discusses the urban economics literature on climate change adaptation. Other papers explore adaptation in areas other than the temperature-mortality link. For example, Graff Zivin and Neidell (2014) study the effect of rising temperatures on time use, Dell et al. (2012) study the effect of temperature on growth, Burke and Emerick (2016) study the effect of temperature on agricultural output, and Fried and Sheldon (2016) and Hsiang and Narita (2012) study the effect of extreme weather events like tornados on economic damages. Kahn (2016) summarizes the literature on firm and household adaptive responses to climate change.

climates. They reject the hypothesis that the mortality impact of a hot day is the same across all regions, and the qualitative pattern suggests that the mortality effect of a hot day is smaller in historically warmer areas.[4]

The study closest in spirit to our analysis is Barreca et al. (2015), which divides states into deciles based on the average number of days per year that the temperature is greater than 90°F. If finds that the mortality effect of a day with temperature above 90°F relative to a typical day in that area with temperature below 90°F is larger in the typically-cooler deciles.[5] However, our work expands on that of Barreca et al. (2015; 2016) in two important ways. First, we conduct our analysis at the ZIP code level, which allows us to more precisely identify and parameterize the relationship between temperature effects and climate. Second, and more critically, our analysis characterizes the relationship between temperature and the entire temperature-mortality curve. Adaptation to frequently-experienced temperatures is generally accompanied by offsetting effects from infrequently-experience temperatures. These de-adaptations are as important, or even more important, than adaptations in determining the predicted mortality impact of climate change. Thus, focusing only on the impact of hot days misses an important aspect of how the temperature-mortality relationship varies with climate.

Finally, our paper also addresses econometric issues that may arise when estimating fixed effects regressions in the presence of heterogeneous effects. As highlighted by Angrist and Kreuger (1999) and Gibbons and Suarez Serrato (2017), fixed-effects regressions average climate-specific effects using weights based on both the sample frequency with which the climate region experiences the treatment and the conditional variance of the treatment (Angrist and Krueger, 1999; Gibbons and Suárez Serrato, 2017). Thus, and fixed effects regressions do not consistently recover population average treatment effects in the presence of heterogeneous effects. Our approach addresses this issue by estimating damages for each region before aggregating to the national level.

The remainder of this paper proceeds as follows. Section II describes our data. In section III, we estimate climate-specific temperature-mortality relationships and the extent to which air conditioning explains this relationship. Section IV makes predictions of long-run climate-change-induced mortality, with and without climate-based regional heterogeneity and with and without adaptation. Section V concludes.

---

[4] In an epidemiological study of 11 U.S. cities, Curriero et al. (2002) find a qualitative pattern of larger effects of cold temperatures in the southern cities and larger effects of heat in northern cities, although the study relies on functional form restrictions and does not report whether observed differences are statistically significant. Deschênes (2014) reviews the empirical literature on temperature, human health, and adaptation.

[5] One potential criticism of this study is that a typical day with temperature below 90°F is warmer in historically warmer regions, and thus the size of the temperature shock associated with a >90°F day is smaller. Although this could account for their results, our findings confirm the general pattern identified in that paper.

II. Data

II.A. Data Description

Our analysis combines, for the first time, two primary data sources: weather monitor readings from the National Oceanic and Atmospheric Administration's National Centers for Environmental Information (NCEI) and Medicare administrative data, both covering the period 1992 – 2011 at the daily level. We briefly describe the data used for our analysis, and we provide more detailed data descriptions and summary statistics in Appendix A.

The NCEI climatological data include daily monitor-level measures of high and low temperatures for many monitors throughout the nation. The dataset contains 6,786 monitors for 1992 and 9,017 monitors for 2011. The data are quality controlled by NCEI, and any observations that do not meet quality standards are recoded as missing. The average, per-monitor number of days per year in which a non-missing value of temperature is present is about 330-340. Because the geographic level of the Medicare data observations is the ZIP code, we first average the daily high- and low-temperature values for each monitor and then aggregate to ZIP code measures using inverse distance weighting.[6] In addition to the daily temperature data, we use NOAA's historical Climate Normals data, which include each monitor location's average number of cooling degree days (CDD) over the three-decade period 1981-2010. These long-run climate measures are used to categorize ZIP codes as hot or cold. CDD are based on daily temperatures and are designed to reflect the energy needed to cool a building to a base temperature, typically 65°F. For example, one full day at a temperature of 75°F would represent 10 CDD, while a day with temperatures below the base temperature would represent 0 CDD. A region's annual CDD is the sum of daily CDD values across all days in the year. More details on the weather monitor data are described in Appendix Table A1.

Our measures of mortality come from the Medicare Master Beneficiary Summary File from 1992 – 2011, which is an annual directory providing enrollment data on 100 percent of individuals eligible for Medicare in that calendar year. The directory includes beneficiary date of death derived from Social Security Administration records, ZIP code of residence, date of birth, sex, and whether the beneficiary is enrolled in managed care (Medicare Advantage).

---

[6] This is the same method used by Currie and Niedell (2005), and also by Beatty and Shimshack (2014) for aggregating pollution monitor data. Specifically, using data from the U.S. Census Gazetteer files on coordinates of geographic centroids of ZIP codes, each ZIP code is assigned a temperature reading by averaging readings from all temperature monitors located within 20 miles of the ZIP code centroid, weighting the monitors by the inverse of the location to the centroid. ZIP codes are not assigned a temperature reading if they are not within 20 miles of a weather monitor. This affects fewer than 2 percent of ZIP codes, and an even smaller fraction of the population, since regions without a nearby weather monitor tend to be sparsely populated. On average, each ZIP code is matched to 4 monitors and assigned the weighted average temperature from those 4 monitors.

Finally, we use data from the nationally representative Residential Energy Consumption Survey (RECS) to derive estimates of ZIP code-level air conditioning use. The RECS is administered roughly every four years by the U.S. Energy Information Administration and records information on housing characteristics, household appliances (including air conditioning), geographic information, and household demographics. We use data from RECS survey years 1993, 1997, 2001, 2005, and 2009 to provide balanced coverage of our study period 1992-2011.

II.B. Summary Statistics

The primary unit of analysis for our study is a ZIP code-day. There are roughly 32,000 ZIP codes in our sample, yielding over 230 million ZIP code-day observations over our sample period (1992 – 2011). The U.S. exhibits substantial climate variation, with average annual temperature ranging from less than 40°F in the northern Midwest to more than 70°F in the South. Figure 1 summarizes the distribution of temperature over this period across each of 19 temperature bins ranging in 5°F increments from <10°F to >95°F. The grey shaded region presents the distribution of daily average temperature for the U.S. as a whole, while the blue, grey and orange curves present the distributions for the coolest, middle and warmest thirds of U.S. ZIP codes (weighted by population). In addition to differences in average temperatures, the climate terciles exhibit significant heterogeneity, especially in the tails. Very hot days occur frequently in the warmest ZIP codes by very rarely in the coolest. The opposite is true for very cold days.

Table 1 presents summary statistics of ZIP-day observations binned by average daily temperature. Columns 2 – 4 present the daily median, low and high temperatures within each bin. One feature worth noting is that a day with a high average temperature often features an extremely high daily high temperature and also a high daily low temperature. For example, days in the 90-95°F bin have a median average temperature of around 92°F, but a high temperature around 105°F. Further, the daily low is almost 80°F, meaning that even the night offers little respite from the heat. A similar pattern emerges at the low end, where days with low average temperatures often feature very cold low temperatures and high temperatures that remain well below freezing.[7] Column 5 presents the average daily mortality rate for the 65+ population during our sample. Unconditionally, the average daily mortality rate in the 65+ population over this time period was 13.3 deaths per 100,000 beneficiaries. However, mortality rates were systematically lower on warmer ZIP-days, with the lowest mortality rate of 11.8 deaths per 100,000 occurring after days in the hottest temperature bin where average temperatures were at least 95°F. A naïve interpretation of this pattern

---

[7] Appendix Figure A2 shows the number of days per beneficiary exposed to 90°F or higher, for each year in the sample, along with a three-year moving average and a linear trend. The spike in 2011 corresponds to the summer heat wave that year. The linear trend finding of approximately one additional 90+ day every 20 years is roughly consistent with predictions of the effects of climate change, for example see http://www.climatecentral.org/gallery/graphics/new_york_july_days_over_90_degrees.

is that replacing cooler days with very hot days reduces mortality. Yet this conclusion could be flawed for two important reasons. The first is that hot days tend to occur during the summer, potentially confounding the temperature effect with seasonality. The second is that not all regions are equally likely to experience hot days, and the population of individuals residing in regions where this occurs most often may differ systematically from cooler regions. The richness of our data provides an exceptional opportunity to address these potential confounders by controlling very flexibly for both location and seasonality, as described below.

III. HETEROGENEITY IN THE MORTALITY EFFECTS OF TEMPERATURE

As a first step toward our main goal of predicting the mortality effects of climate change, we estimate the mortality effects of temperature across regions. We begin our analysis by non-parametrically estimating the current relationship between temperature and mortality, accounting for heterogeneity based on current climate. We allow for regional heterogeneity by estimating the model separately for the coolest, middle, and warmest third of U.S. ZIP codes.[8] Later, we will employ a semi-parametric approach and estimate the temperature-mortality relationship at the ZIP code level. We begin with the non-parametric, tercile-level approach to have a benchmark against which to compare the semi-parametric estimates to ensure that the results are not overly-determined by functional form restrictions. Furthermore, the tercile-based analysis provides a straightforward illustration of the importance of accounting for heterogeneity.

We find significant heterogeneity in the effects of temperature across these climate regions with patterns consistent with adaptation to local climate. Supporting the interpretation that this heterogeneity is evidence of adaptation rather than immutable geographic factors correlated with current climate, we show that differential air conditioning adoption across climate regions can account for regional differences in the mortality impact of hot days. Before moving on to our climate change predictions, this section concludes by showing how failure to incorporate heterogeneity can lead to errors of both sign and magnitude in estimating aggregate mortality effects.

III.A. Empirical Strategy

For a first characterization demonstration of how the mortality effects of temperature vary across climate regions, we partition ZIP codes into three climate regions based on whether the long-run average cooling degree days (CDD) in each ZIP is in the coolest, middle, or warmest third of the population-weighted

---

[8] For comparison, the Appendix describes estimates that assume a homogeneous response across the country.

distribution of CDD (Appendix Figure A7 maps CDD across U.S. ZIP codes).[9] To allow for the greatest flexibility, we then non-parametrically estimate the temperature-mortality relationship for each tercile.

Following the approach of Deschênes and Greenstone (2011), we use year-over-year variation in daily temperature to identify the causal effect of temperature on mortality. Our analysis uses daily observations of mortality and temperature at the ZIP code level. Our primary outcome of interest, $mortality_{zd}$, is defined as the number of deaths per 100,000 beneficiaries in ZIP code z within a specified window of 3 days after the reference day d to capture possible lags in mortality effects as well as near-term mortality displacement (harvesting).[10] To investigate the potential for more long-term harvesting, in Appendix C we conduct extensive robustness checks, including considering windows longer than 3 days. We do not observe harvesting at either very hot or very cold temperatures when extending the mortality window beyond 3 days, and therefore we focus our primary analysis to 3-day window outcomes. Our estimating equation is:

$$
\begin{aligned}
mortality_{zd} = &\sum_{b\in B\backslash\{55-60\}} \beta_b^{cool} tempbin_{zd}^b \times \mathbf{1}(ZIP\ z\ in\ coolest\ third\ of\ regions) \\
&+ \sum_{b\in B\backslash\{65-70\}} \beta_b^{mid} tempbin_{zd}^b \times \mathbf{1}(ZIP\ z\ in\ middle\ third\ of\ regions) \\
&+ \sum_{b\in B\backslash\{75-80\}} \beta_b^{hot} tempbin_{zd}^b \times \mathbf{1}(ZIP\ z\ in\ warmest\ third\ of\ regions) \\
&+ ZipDay_{zd} + L_{zd} + StYr_{zd} + \varepsilon_{zd},
\end{aligned}
\tag{1}
$$

where an observation is a ZIP code z on day d. Thus, while the unit of observation is the ZIP code, we constrain all ZIP codes within the same climate tercile to have the same temperature-mortality relationship.

The primary independent variables of interest in Equation (1) are temperature indicators $tempbin_{zd}^b$ defined by which of 19 temperature bins $B = \{< 10, 10-15, \ldots, 90-95, 95+\}$ the average temperature in ZIP code z falls on day d. These temperature indicators allow for arbitrary nonlinearities in the relationship between temperature and mortality. Because Equation (1) includes ZIP code fixed effects, the coefficients on the set of temperature indicators for each climate region are only identified up to a common constant. Given that the choice of the omitted temperature category for each climate region is arbitrary, we omit the bin at which mortality is minimized for that climate region.[11] As a result, the

---

[9] The coolest tercile contains ZIP codes with less than 783 CDD. The middle tercile contains ZIP codes with 783 – 1416 annual CDD, and the warmest tercile contains ZIP codes with more than 1416 CDD.

[10] For example, the 3-day mortality rate assigned to a ZIP code for January 1, 2000, is the number of deaths on January 1, 2, and 3, divided by the elderly population, and multiplied by 100,000.

[11] An alternative approach would be to omit the same temperature category (e.g. 65 – 70 F) for all climate regions. However, the results from that regression can be obtained from Equation (1) as the difference $\beta_b^c - \beta_{\{65-70\}}^c$.

coefficients β_b^c describe for a ZIP code in climate region c the average difference in mortality from a day with average temperature in bin b relative to day with a temperature that minimizes mortality in region c.

We identify the effects of temperature on mortality by isolating year-over-year variations in temperature and mortality, controlling for both geography and seasonality using ZIP code by day of year fixed effects, $ZipDay_{zd}$. This control strategy accounts for the fact that mortality follows a seasonal pattern, with deaths being highest in winter and lowest in summer, and that this pattern varies by location. To account for serial correlation in daily temperature and potentially lagged mortality effects, L_{zd} includes three fully interacted sets of 5-degree average temperature bins for the preceding 2 and 6 days as well as the subsequent 2 days. Finally, state-by-year fixed effects $StYr_{zd}$ are included to control for arbitrary annual shocks that may vary by state, such as changes to Medicare or Medicaid policy. All regressions are weighted by the ZIP code Medicare population, and standard errors are clustered at the county level.

III.B. Results

Figure 2, Panel A depicts results from estimating Equation (1) with 3-day mortality as the outcome. For clarity, we show coefficient estimates and confidence intervals for the hottest and coldest terciles, while for the middle tercile we show only the coefficient estimates.[12] For comparison, we also include, in Figure 2, Panel B, the results of estimating Equation (1) under the assumption of homogeneous temperature effects. Details are provided in the Appendix.

The heterogeneity in temperature effects is a strong indicator of adaptation to the current climate. The hottest ZIP code tercile, depicted with the red line, has its lowest mortality at 75-80°F, while the coldest ZIP code tercile, depicted with the blue line, has its lowest mortality at 55-60°F. More strikingly, as temperatures increase above 75°F, the colder regions feature a stark increase in mortality, while warmer regions have much more modest effects. For example, an 85-90°F day increases the mortality rate in the coldest decile by more than 3 deaths per 100,000, while it has nearly no effect in the hottest decile. On the other hand, the increase in mortality in the hot region as temperature decreases is much larger than in the cold or medium regions. A 10-15°F day has no effect on mortality in the coldest tercile but increases deaths in the hottest tercile by 2 deaths per 100,000 population. Overall, the coldest regions see almost no mortality impact from even the coldest days.

---

[12] Estimates are reported and indicated with a marker for temperature bins comprising at least 0.05 percent of daily ZIP code observations (about two days every ten years, on average) in that climate region. The dashed curves are 5th order polynomials fit to the reported estimates for each tercile and fit to all temperature bins.

We attribute this heterogeneity in the impact of heat/cold to technological, behavioral, and/or biological differences across the regions, i.e., adaptation.[13] It is important to note that this adaptation has two aspects: regions are good at dealing with temperatures they commonly experience, and they are bad at dealing with temperatures that they do not. Thus, a cold region is good at dealing with cold but bad at dealing with heat, while the opposite is true for a hot region. This suggests that increasing temperatures may be particularly bad for cold regions, since they will experience fewer cold days, which they are adapted to deal with, and more hot days, which they are not. It further suggests that adaptation to heat may come at the expense of de-adaptation to cold, and vice versa.

In our analysis, we distinguish in a statistically precise and economically meaningful way between temperature effects in the hottest and coldest terciles over the entire range of temperatures except for the 65-70°F and 70-75°F bins, where mortality effects are small regardless of climate. Relative to previous studies (Deschênes and Greenstone, 2011; Barreca et al., 2015, 2016), our results are more precise for two reasons. First, the Medicare data allow us to examine temperature effects at a much finer geographic level and over finer temperature bins. Second, we are able to partition the ZIP codes based on climate rather than state, and removing within-partition heterogeneity in climate increases the precision of our estimates. These results inform our preferred approach in the next section, where we smoothly parameterize the dependence of temperature effects on local climate and use this relationship to predict the impact of adaptation on temperature-related mortality.

Comparing Panels A and B of Figure 2 provides insight into the possible biases introduced by aggregating nonlinear temperature effects across disparate climates using fixed-effects regression. Part of the bias is based on the sample frequency of the treatment in different regions, and is relatively straightforward to understand. Under the assumption of homogeneous effects commonly employed in the literature (Deschênes and Greenstone, 2011; Barreca et al., 2015, 2016; Hsiang et al., 2017), the impact of hot days is primarily identified based off of places that experience many hot days, i.e. warm places. Similarly, the impact of cold days is primarily identified off cool places. Because of adaptation, places that frequently experience a given temperature have relatively low mortality from that temperature. The result is that the homogeneous effects depicted in Panel B tend to be biased toward the lower envelope of the effects in Panel B. The other part of the bias in fixed-effects estimation arises from the fact that the average treatment effect estimate is based on the conditional variance of the treatment, which in the case of multiple-regression with many fixed effects can be highly complex and difficult to characterize.

---

[13] To be precise, differences in the marginal effects, i.e., the slopes of the curve, are attributable to adaptation. The difference in average mortality across hot and cold regions is not depicted in these diagrams. That is, although the slopes between the points in the curves are well-identified, the height of the curve depends on the values of the fixed effects from the regression, and the particular time and place at which the fixed effects are evaluated.

To take one example of how properly accounting for treatment heterogeneity matters, homogeneous effects would imply that replacing a 40-45°F day with an 85-90°F day would have no net mortality effect. Using the climate-specific effects in Panel B, however, mortality would increase by 3 deaths per 100,000 individuals in the coldest tercile and decrease by 1 death per 100,000 individuals in the warmest tercile. The discrepancy arises in part because homogeneous effects underweight both the benefit to losing a cold day in the warmest tercile (which doesn't usually experience cold days) and the harm of gaining a hot day in the coldest tercile (which doesn't usually experience hot days).

III.C. Air Conditioning as Adaptation

There are two possible categories of explanations for regional heterogeneity in the temperature-mortality relationship. The first is that the heterogeneity is substantially due to changes in human physiology and behavior, i.e., adaptation. The second is that the regional differences we observe are due to characteristics that are correlated with current climate but not the results of human choices or physiology. This distinction is important for interpretation because if regional differences are caused by factors that are immutable in this sense, then even though Chicago in the future may face the climate that Dallas does now, we should not expect the Chicago of the future to be as good at dealing with heat as Dallas currently is. Thus, understanding the extent to which *current* heterogeneity is due to adaptation is important for understanding the extent to which *future* adaptation may mitigate the impact of climate change.

There are numerous ways in which people and communities adapt to their current climate, including biological acclimatization, infrastructure investments, architectural styles, etc. It is beyond the scope of this paper to identify each component of regional adaptation or to estimate how much regional heterogeneity is explained by each method of adaptation. However, in this section we examine more closely one method of adaptation—residential air conditioning (AC)—and the extent to which AC adoption can moderate the relationship between temperature and mortality. Strikingly, we find that differential AC adoption can explain near all of the differences in heat-related mortality that we observe across regions, but it does not explain differential cold-related mortality. We interpret this as evidence that regional differences are substantially due to behavioral differences, and thus that it is reasonable to expect that regions could continue to adapt to their changing climate in the future. It is important to note, however, that this is a statement about technological feasibility. While regions *could* adapt their behavior, infrastructure, etc., to address future climate, we do not assess the cost of doing so. It may be that such changes, though technologically feasible, are prohibitively expensive to implement.

Because we do not observe AC adoption at the ZIP code level, we impute a value for the ZIP code penetration rate by fitting a machine learning (LASSO) model of AC adoption based on housing unit characteristics including housing stock age, geography, and climate and using data from the Residential

Energy Consumption Survey (RECS), the American Community Survey (ACS), and the 2010 Census. Full details of the data and imputation procedure are presented in Appendix Section A.3. To address the potential concern that including local climate to predict local AC penetration may generate an artificial relationship between the mortality effects of temperature and AC adoption rates, we show for robustness in Appendix Figure A8 that results are very similar when climate variables are excluded from the imputation step. Empirically, AC adoption depends strongly on climate. As a result, omitting climate from the set of AC predictors introduces the possibility of systematic bias in the imputed AC measure that is correlated with climate. Because of this, our preferred imputation procedure includes local climate in the set of predictors.

To estimate how temperature effects vary by AC penetration, we interact the continuous value of AC penetration (between 0 and 1) with the temperature bins in Equation (1). We also include in this regression Census Region by temperature bin fixed effects. Figure 3, Panel A, presents these regression results for 3-day mortality outcomes. The curve represents the regression coefficients on the interaction between the AC penetration variable and the specified temperature bin. The interaction effects are significantly negative for temperatures above 70°F. For example, the coefficient on the interaction between AC penetration and the 80-85°F temperature bin is -6.1. This implies each 10 percentage point increase in AC penetration corresponds to 0.61 fewer deaths per 100,000 individuals from an 80-85°F day. At hotter temperatures, the air conditioning effect is even stronger: for example, the coefficient on the 90-95°F day bin implies about 2 fewer deaths per 100,000 for each 10 percentage point increase in AC penetration. By contrast, for temperatures below about 65°F, the interactions between air conditioning and temperature are all close to zero, indicating that the mortality effect of cold days is independent of the level of AC penetration.

Based on these estimates, we calculate how much of the difference in heat sensitivity across regions can be explained by differences in air conditioning penetration. Over the sample, AC penetration was 93.2 percent in the warmest ZIP codes, 82.1 percent in the middle third of ZIP codes, and 63.2 percent in the coolest. Panels B and C of Figure 3 present counterfactual simulations where we compute what the mortality curve would have been if the warmest ZIP codes had the AC penetration rates of the other two terciles, respectively. Thus, Panel B estimates what the temperature-mortality relationship would be in the warmest third of ZIP codes if the AC penetration rate in those ZIP codes were 30 percentage points lower than it actually is, as is the case in the coolest ZIP codes. Strikingly, the counterfactual warmest ZIP code curve closely tracks the actual curve for the coolest ZIP codes at temperatures above 65-70°F, suggesting that differences in AC penetration can explain nearly all of the differences in heat-related mortality between the warmest and coolest regions. At the same time, differences in AC penetration explain little of the difference between the regions on cold days, consistent with air conditioning as the mechanism driving the regional differences in heat-related mortality only. Panel C repeats this exercise, this time comparing a

13

counterfactual where the warmest ZIP codes are adjusted for the approximately 11 percentage point difference in AC between the warmest and middle terciles. Once again, the counterfactual warmest tercile aligns almost perfectly with the middle tercile on hot days, again suggesting that differences in AC penetration account for a substantial portion of the estimated difference between the regions.

These results extend the finding from Barreca et al. (2016) that increases in AC adoption over time can explain a substantial share of the reduction in heat-related mortality in the U.S. over the past century.[14] Our results show that in addition to explaining the reduction in heat-related mortality in the time-series, AC adoption can explain nearly all the cross-sectional differences in heat-related deaths observed across U.S. climate regions. These counterfactual exercises provide strong evidence that the heterogeneity in temperature effects across climate regions is driven primarily by adaptation, and in particular AC, rather than by immutable regional characteristics. Thus, we might expect places to engage in this and other types of adaptive behavior in response to warming.

These results come with an important caveat: even if temperature shocks are exogenous, the decision to adopt AC is not and may depend on many factors, such as income or local climate. As a result, differences in mortality by AC penetration may reflect the effects of any of these factors that correlate with AC adoption and may not necessarily reflect the causal effect of AC adoption on heat-related mortality risk. However, many confounding factors like climate and income may be expected to affect temperature effects on both hot and cold days. Our finding that AC adoption only explains differential heat-related mortality but not cold-related mortality alleviates the concern that our AC effects simply reflect typical differences in mortality effects across climate regions or by levels of income.

III.D. The Importance of Regional Heterogeneity in Climate Change Impacts

Climate-related heterogeneity can have important implications for predicting how mortality will change as a region's climate changes. Although we will investigate this issue in greater detail in the next section using detailed, science-based climate change predictions, the simple exercise here can help to fix ideas. In Figure 4, we present counterfactual simulations that qualitatively explore the importance of regional heterogeneity in the temperature-mortality relationship. The counterfactual simulations considered here are very simple: we replace a given climate region's current distribution of temperature with the current distribution of one of the other climate regions and compute expected mortality in two ways. First, we use the homogeneous, national temperature-mortality relationship depicted in Figure 2 Panel A, and second, we use the climate-specific temperature-mortality relationships depicted in Figure 2 Panel B.

---

[14] Additionally, our results complement those of observational studies in the public health literature (e.g., Rogot et al.,1992), which show that air-conditioning reduces the mortality impact of hot days.

Figure 4 summarizes the mortality impact of these scenarios for each of the three climate terciles. The blue bars give the impact on temperature-related mortality computed using the homogeneous temperature-mortality relationship, and the green bars give the impact on heat-related mortality using the climate-specific temperature-mortality relationship. In each case, we compute the effect relative to a baseline using the region's current climate. The leftmost third of Figure 4 represents the coolest tercile of ZIPs. Under homogeneous effects (blue bars), increasing the temperature distribution to that of the middle or warmest tercile (i.e., warming) reduces mortality. For example, replacing the temperature of the coolest tercile with that of the warmest tercile decreases deaths by 0.4 percent.[15] However, using the climate-specific relationship (green bars) yields the opposite prediction: warming increases mortality, by 3.3 percent. The same occurs for the middle tercile, where moving to the temperature distribution of the warmest tercile reduces mortality under homogeneous effects (by 0.2 percent) but increases mortality under climate-specific effects (by 0.4 percent). Lastly, for the hottest tercile in the rightmost third of panel B, replacing its temperature with that of either the coldest or the middle tercile will increase deaths, under either the homogeneous or the climate-specific effects. But, the magnitude of the increase in deaths is much larger under the climate-specific effects. For example, replacing it with the coolest third will increase deaths by 0.4 percent under homogeneous effects but by 3.1 percent under climate-specific effects.

The simulations under the assumption of homogeneous effects suggest that warming would reduce mortality in cold places and increase mortality in hot places. However, under climate-specific effects, a change in a region's temperature always increases mortality, whether that change involves warming or cooling. This evidence further supports that regions exhibit significant adaptation to their current climate but also demonstrates that failing to account for heterogeneity when predicting climate impacts can lead to errors of both sign and magnitude.

IV. PREDICTING CLIMATE CHANGE-INDUCED MORTALITY AND ADAPTATION

A primary contribution of this paper is to develop estimates of the end-of-century mortality impact of climate change. Because heterogeneity in both the temperature-mortality relationship and the change in the temperature distribution can affect the predictions, we begin with a general framework for thinking about the problem. We then discuss how we derive model-based predictions of ZIP code-level changes in the temperature distribution and empirically estimate the temperature-mortality relationship. Finally, we present predictions of the change in mortality by the end of the century (2080-2100) under three sets of assumptions: homogeneous temperature effects with no future adaptation, current-climate heterogeneity

---

[15] By construction, replacing the coolest third's temperatures with that of the coolest third yields no change in mortality, so the bars in those columns are zero.

with no future adaptation, and current-climate heterogeneity with future adaptation. Although heterogeneity in current effects could be interpreted as arising from past adaptation, the importance of incorporating heterogeneous effects in climate change impact estimates does not depend on whether this heterogeneity results from adaptation.

IV.A. General Framework.

A major component of integrated assessment models (IAMs) of climate change is choosing a damage function to model the economic costs of climate change. The component of the damage function that arises from temperature-related mortality necessarily involves judgments about the extent to which adaptation will modify the temperature-mortality relationship. To fix ideas, let $m_z^p(t)$ denote the mortality effect of a day with average temperature in bin t in ZIP code z in period p. We will consider both the current period, *p=current*, and the future, *p=future*. Let $g_z^p(t)$ be the number of days per year in which the temperature falls in bin t in period p. Current annual mortality (CAM_z) is therefore:

$$CAM_z = \sum_t m_z^{current}(t)\, g_z^{current}(t).$$

We are interested in the change in excess mortality due to climate change. Future annual mortality (FAM_z) can differ from CAM_z for two reasons. First, climate change will alter the temperature distribution in location z, with the new distribution given by $g_z^{future}(t)$. Second, the temperature-mortality relationship may change between the current and future periods. There are many reasons why this might occur, including adaptation (e.g., AC adoption, changes in building practices, etc.) or environmental changes (e.g., changes in the prevalence vector-borne illnesses, changes in the agricultural sector, sea-level changes, etc.). Let $m_z^{future}(t)$ denote the future mortality effect of temperature bin t in location z. In this case, the change in excess mortality would incorporate both the change in the temperature distribution and the change in the temperature-mortality relationship:

$$FAM_z - CAM_z = \sum_t m_z^{future}(t)\, g_z^{future}(t) - \sum_t m_z^{current}(t) g_z^{current}(t). \qquad (2)$$

IV.B. Empirical Implementation

Computing the change in excess mortality involves four functions: current and future temperature distributions and current and future temperature-mortality relationships. The current temperature distribution is observed in the data. Future temperature distributions will be taken from climate change

projection models, described in section IV.B.1. The current temperature-mortality relationship is estimated from the data, though for this analysis we estimate this relationship at the ZIP code level as described in section IV.B.2. Finally, the future temperature-mortality relationship allowing for adaptation must be estimated/projected, as described in section IV.B.3.

IV.B.1 Predictions of future temperature distributions. We use predictions of the end-of-century (2080 – 2100) temperature distribution for each ZIP code using daily projections derived from all 21 climate models for which daily scenarios are produced and distributed as part of the Coupled Model Intercomparison Project Phase 5 (CMIP5).[16] Daily downscaled projections for each of these models come from the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) dataset. The NEX-GDDP data include daily minimum and maximum temperature predictions on a 25km by 25km grid (0.25 degree spatial resolution). We consider climate predictions made under the Representative Concentration Pathway (RCP) 8.5 "business as usual" scenario, where emissions continue to rise throughout the 21st century (Meinshausen et al., 2011).[17] Predicted daily average temperatures (the average of the predicted daily high and low temperatures) are assigned to ZIP codes using the climate model grid point nearest to the ZIP centroid. We then use the predicted daily average temperatures over the period 2080 – 2100 to construct the predicted daily temperature distribution at 1°F bins for each ZIP code during that period. This process yields 21 end-of-century climate (i.e. daily temperature distribution) predictions for each ZIP code.

Rather than presenting separate results for each of the 21 models, we average over all of the models to produce a single "meta" prediction of the temperature distribution for each ZIP code. We refer to this average model as the meta-model and the average predicted temperature distribution as the meta-prediction or the meta-distribution. The techniques we use apply equally well to the output of any single model. Further, since averaging is a linear operator, in addition to being the impact on the "average" model's predicted distribution, our results can also be interpreted as the average impact across all of the models' predicted distributions. Appendix Figure A9 shows the predicted change in U.S. average annual temperature under RCP 8.5 over the rest of the century for each of the 21 models we study and for the meta-model. The various models predict an end-of-century change in average temperature ranging from about 5°F to 11.5°F, while the meta-model predicts around 8.5°F of warming. Appendix Figure A10 maps the projected change in temperature under the RCP 8.5 meta-model. Although predicted warming tends to be higher in areas that are currently cooler, comparing Appendix Figures A9 and A7 shows that there is significant variation in predicted warming even among regions that currently have quite similar climates.

---

[16] See Auffhammer et al. (2013) for a discussion of the use of climate models in economic analysis.
[17] The NEX-GDDP data also contain climate predictions under other emissions scenarios, such as RCP 4.5, a mid-range projection under which emissions peak around 2040 and then decline. It is straightforward to modify our method to accommodate alternative models or scenarios.

IV.B.2 Estimating ZIP code-specific mortality relationships. While the non-parametric approach identifies adaptation to current climate, its usefulness for studying the possibility of future adaptation is limited for a number of reasons. First, even within climate terciles, there is a wide range of current climates. Second, ZIP codes that currently reside in the same tercile will often be predicted to experience very different changes in climate over the remainder of the century. Finally, while some changes in the temperature distribution or adaptation might be accurately captured by modeling ZIP codes as moving among climate terciles, many will not be. This is especially true of the currently-hottest ZIP codes, since climate change will likely move their temperature distribution out of the range of current experience.

With sufficient observations for each ZIP code at each temperature, we could estimate the temperature-mortality relationship non-parametrically for each ZIP code in the same way that we estimated it non-parametrically at the climate tercile level using Equation (1). In practice, however, there are not enough observations for each ZIP code to estimate this relationship precisely. Instead, we estimate the daily temperature-mortality relationship as a semi-parametric, smooth function $f(t, CDD)$ that depends on both daily average temperature and climate, as captured by the long-run average CDD (or CDD normal) of the ZIP code.

The regression equation used to estimate this semi-parametric function of temperature and climate is identical to Equation (1), except the temperature and climate indicators are replaced by this smooth function $f(t, CDD)$ of temperature and climate, yielding the estimating equation:

$$mortality_{zd} = f(t_{zd}, CDD_z) + ZipDay_{zd} + L_{zd} + StYr_{zd} + \varepsilon_{zd}. \tag{3}$$

We define $f(t, CDD)$ to be a linear spline in temperature, with knot points at 10-degree increments from 30 to 90°F, fully interacted with a spline in log CDD, with knot points at the 33rd and 66th percentiles of the current distribution of ZIP-level CDD normals, the same cutoff points used to define the climate terciles. Since $f(t, CDD)$ is identified up to a constant, we choose 65°F as the reference temperature. Specifically, if F_{CDD} is the CDF of the current CDD Normal distribution, then

$$f(t, CDD) = \beta_0(t - 65) + \sum_{k=3}^{9} \beta_k \mathbf{1}(t > 10k)(t - 65)$$

$$+ \sum_{p=0}^{2} \log CDD \cdot \mathbf{1}(CDD > F_{CDD}(33p)) \left[\beta_0^p(t - 65) + \sum_{k=3}^{9} \beta_k^p \mathbf{1}(t > 10k)(t - 65) \right].$$

Appendix Figure A11 summarizes the results of estimating Equation (3) by plotting the temperature-mortality relationship for two cold ZIP codes (Fargo, ND, and Minneapolis, MN); one moderate ZIP code (Chicago, IL); and two hot ZIP codes (Dallas, TX, and Miami, FL) using the estimated temperature effects

18

$f(t, CDD)$ evaluated at each ZIP code's current CDD normal. As with the non-parametric tercile-based regressions presented in Figure 2, Panel B, the coldest places suffer the most from hot days, and the hottest places suffer the most from cold days. To provide a more complete summary of how our choice of $f(t, CDD)$ fits the variation in the temperature mortality relationship across climates, we compute fitted values of $f(t, CDD)$ for each ZIP code using realized average daily temperatures over the sample period 1992-2011 and current CDD normal. We then re-estimate Equation (3) where the fitted mortality values are the outcome and where the only controls are the temperature and climate tercile indicators. Appendix Figure A12 compares these results to the non-parametric morality estimates reported in Figure 2, with the parametric estimates broadly lining up with the non-parametric estimates in each climate tercile.

Finally, we can relate the estimate of $f(t, CDD)$ to the mortality effect $m_z^p(t)$ of a day with average temperature t in ZIP code z in period p, introduced in our general framework above. Note that $f(t, CDD)$ is identified up to a constant, determined by mortality on a day with the reference temperature (65°F). Let $M65_z^{period}$ represent mortality on a 65-degree day in ZIP code z and a given *period*. Then

$$m_z^{period}(t) = M65_z^{period} + f\big(t, CDD_z^{period}\big).$$

IV.B.3. Adaptation Predictions. We estimate the change in mortality between the current period (1992 – 2011) and the end of the century (2080 – 2100) using the meta-predictions of climate change under the RCP 8.5 emissions scenario, described above. We construct these estimates under three different sets of assumptions about how the mortality effects of temperature vary spatially and over time to investigate the importance of accounting for regional heterogeneity and adaptation.

IV.B.3.a. Homogeneous current effects with no future adaptation. Our first set of predictions relies on two simplifications commonly made when predicting mortality effects of climate change. The first has been to use a regression average mortality estimate, $m^{period}(t)$, rather than region-specific estimates, $m_z^{period}(t)$. As discussed above, a key limitation with this approach is that if the mortality effect of temperature varies across climate regions, predicted damages from climate change based on average temperature effects could lead to biased local predictions and inefficient strategic policy responses for all regions.

The second simplification has been to estimate health damages under a "business as usual" assumption of no adaptation to future climate change, i.e. to define $m_z^{future}(t)$ to be equal to $m_z^{current}(t)$. This approach quantifies the mortality benefits from costly strategic climate strategies relative to a world in which no strategic action is taken. Importantly, this approach overstates the mortality costs of climate change if the marginal cost of mortality adaptation (such as through behavior change or technology adoption) is less than the cost of averted mortality. We implement this empirically by estimating a version

of Equation (3) where $f(t, CDD)$ does not depend on CDD to get a single temperature-mortality relationship $m(t)^{current} = M65 + f(t)$, and then using that relationship for all ZIP codes in both the current and future periods. Under these conditions, the change in mortality for each ZIP code z from Equation (2) is given by

$$FAM_z - CAM_z = \sum_t f(t) \left[g_z^{future}(t) - g_z^{current}(t) \right].$$

Note that while we use the same mortality function $f(t)$ for all ZIP codes, each ZIP's mortality change is computed with respect to its own predicted future temperature distribution. These assumptions comprise the case of *homogeneous current effects with no future adaptation.*

IV.B.3.b. Current-climate heterogeneity with no future adaptation. For our second set of predictions, we allow for each individual ZIP code to have its own temperature-mortality relationship, $m_z^{current}(t)$, by estimating Equation (3) where $f(t, CDD)$ is permitted to depend on CDD. Thus, any two ZIP codes with the same current CDD will have the same estimated temperature-mortality curve. We continue the no-adaptation, "business-as-usual" assumption of using this mortality relationship to capture both current and future conditions. Now, the change in mortality for each ZIP code z from Equation (2) is given by

$$FAM_z - CAM_z = \sum_t f(t, CDD_z^{current}) \left[g_z^{future}(t) - g_z^{current}(t) \right].$$

The key difference between this computation and the homogeneous business as usual prediction is that here we account for regional climate-specific heterogeneity by computing the expected mortality change with respect to each ZIP code's temperature-mortality relationship. These assumptions comprise the *current-climate heterogeneity with no future adaptation* case.

IV.B.3.c. Current-climate heterogeneity with future adaptation. In our third and final set of predictions, we account for both climate-specific heterogeneity in the temperature-mortality relationship and adaptation over time. We operationalize this by evaluating the future mortality effects of temperature in ZIP code z under the predicted future climate in that ZIP code. Specifically,

$$m_z^{future}(t) = M65_z^{future} + f(t, CDD_z^{future}),$$

where CDD_z^{future} is the average end-of-century CDD for ZIP code z predicted by the meta-model.

Intuitively, this approach assumes that if Chicago's climate changes so that its end-of-century CDD is equal to Houston's current CDD, then Chicago's end-of-century temperature-mortality relationship will

be the same as Houston's is today, up to the constant term $M65_z^{future}$. This approach is similar in spirit to the "analog cities approach" that has been used in the epidemiology literature, except that we match based on the full distribution of local climate, as captured by CDD, across the entire United States.

Two features of this approach to modeling adaptation warrant mention. First, it assumes that adaptation is complete in the sense that if future-Chicago has Houston's current climate, Chicago will look like Houston does today. Thus, we assume that Chicago fully adopts Houston's current technology. This need not be the case, both due to the cost of adaptation and the fact that some characteristics of current-Chicago may be immutable. An alternative, which this approach could accommodate, would be to assume that future-Chicago's temperature-mortality relationship is a weighted average of current-Chicago and current-Houston. The second key feature of this approach is that it ignores the possibility of technological progress. To the extent that future technology improvements moderate the temperature-mortality relationship, we will not capture these effects. Nevertheless, this approach does provide a useful illustration of the potential scope for adaptation in addressing temperature-related mortality due to climate change.

The final input necessary to compute $m_z^{future}(t)$ is $M65_z^{future}$, the future mortality on a day in the reference category, a 65-degree day. For our computations, we assume that mortality at 65 degrees does not change over time, $M65_z^{current} = M65_z^{future}$. Our justification for this assumption is that when the temperature is near 65 degrees, individuals typically do not choose to air condition their homes, and use of heat, if any, is minimal. This assumption is appropriate if regional differences in mortality on 65 degree days, after adjusting for seasonal and other fixed effects, reflect baseline differences in mortality across ZIP codes that are not affected by changes in climate.[18] Under this assumption, the end-of-century change in mortality for each ZIP code z from Equation (2) is given by

$$FAM_z - CAM_z = \sum_t f\left(t, CDD_z^{future}\right) g_z^{future}(t) - \sum_t f\left(t, CDD_z^{current}\right) g_z^{current}(t).$$

This method comprises the *current-climate heterogeneity with future adaptation* case.

---

[18] As pointed out by Schlenker et al. (2013), the treatment of this reference point can have important implications for making predictions under adaptation. We argue that in the context of human mortality, assuming that the outcome on 65 degree days does not vary systematically with adaptation to colder or warmer climates is more likely to be valid than in other contexts, such as agricultural production. In addition, the approach we outline here could be extended to accommodate alternative assumptions about current and future mortality at the reference category, although doing so would come at the cost of more-restrictive assumptions about which geographic-specific effects are immutable, and which are a function of climate. For example, one could assume that $M65_z^{future}$ is equal to $M65_{\hat{z}}^{current}$ for ZIP codes $\hat{z}$ that currently experience ZIP z's future CDD, but this assumption will fail if ZIP codes $\hat{z}$ and z differ along non-climate characteristics that affect mortality.

IV.C. End-of-Century Mortality Prediction Results

The results of our prediction exercise for each of the three cases described above are presented in Figures 5 and 6. In the leftmost scatter plot of Panel A of Figure 5, each blue dot represents the current climate (as measured by CDD Normal on the horizontal axis) and percentage change in predicted annual mortality by the end of the century (2080 – 2100, vertical axis) for a single ZIP code under the assumptions of homogeneous current effects and no adaptation. The large squares represent averages by CDD bin. Under homogeneous effects with no adaptation, we see that the largest mortality increases are expected to occur in the hottest regions, with modest or even negative effects in the coldest regions. This is not surprising, since under homogeneous effects making cold days warmer reduces mortality, while making hot days even hotter increases mortality.

The prediction that hot regions are expected to suffer more under climate change agrees with previous simulations based on homogeneous effects (Houser et al., 2014, Hsiang et al., 2017). However, as our previous results suggest, cold regions tend to be better at dealing with cold and worse at dealing with heat, while hot regions tend to be better at dealing with heat but worse at dealing with cold. The result is that the homogeneous temperature effects assumption tends to understate mortality in the cold regions and overstate mortality in the hot regions. This conclusion is supported by the middle scatter plot in Panel A of Figure 5, where we compute end-of-century mortality using each ZIP code's climate-specific temperature-mortality relationship still not allowing for adaptation (i.e., the current-climate heterogeneity with no future adaptation case). These results, depicted by the green dots in the center scatter plot (circles represent averages by CDD bin), exhibit the opposite relationship from the homogeneous effects plot.[19] Here we see that cold areas, which are expected to experience more hot days by the end of the century and are particularly bad at dealing with them, exhibit the highest expected mortality increase. Hot regions, which will experience more very hot days but are already quite good at dealing with extreme heat, will experience small, or even negative mortality changes.[20]

The rightmost scatter plot in Panel A of Figure 5 presents results for the current-climate heterogeneity with future adaption case. Two features of note emerge from this plot. First, relative to the middle panel with climate-specific current effects but no future adaptation, the predicted mortality change is generally lower. This is because hotter regions are generally better at dealing with hot days, so allowing

---

[19] The greater degree of vertical dispersion in the green dots relative to the blue, especially among the cooler regions, arises due to two reasons. First, ZIP codes with the same climate today can have different predicted future climates, including different fractions of very hot days. Second, because cooler regions are particularly bad at dealing with very hot days (i.e., the temperature-mortality curve is very steep for hot days, as in Figure 2, Panel B), small variation in the proportion of very hot days can induce very different mortality predictions.

[20] The possibility of negative mortality changes for hot regions is driven by the fact that they are particularly bad at dealing with cold, and they are far less likely to experience very cold days at the end of the century.

ZIP codes to adapt reduces the mortality impact of increasingly prevalent hot days. Second, the predicted mortality change relative to baseline is negative over a wide range of initial climates, extending from about 1000 CDD (e.g., current Chicago) on up. As illustrated in Figure 2, Panel A, as regions get hotter, they get better at dealing with heat but worse at dealing with cold. However, under most warming scenarios the likelihood of very cold days shrinks rapidly, allowing ZIP codes that adapt to warming to gain the benefits of better resilience to heat without suffering a large cost due to increased vulnerability to cold.[21]

Panel B of Figure 5 summarizes the predicted overall mortality impact of warming by the end of the century under the three cases, separating by the coolest, middle, and warmest third of ZIP codes, and also presenting the aggregate national level results. The blue bars show that, under homogeneous current effects with no future adaptation, mortality remains essentially constant in the coolest tercile and increases in both the middle and warmest terciles, with the largest effect coming in the warmest tercile. Overall, under the homogeneous current effects, no future adaptation assumption, end-of-century mortality would be expected to increase by 0.87 percent.

Allowing for current-climate heterogeneity but no future adaptation paints a very different picture. In this case, depicted by the green bars, we see much larger mortality increases in the coolest and middle third of ZIPs compared to homogeneous effects, but a smaller mortality increase in the largest third of ZIPs. This is because cold and moderate ZIP codes are worse at dealing with very hot days than the homogeneous effects curve suggests, while hot ZIP codes are better at it. The overall effect on end-of-century mortality under climate-specific effects is a mortality increase of 2.84 percent, over three times as large as the prediction under homogeneous effects.

Finally, the grey bars summarize the impact of allowing for future adaptation on top of current-climate heterogeneity. Here we see that, for each tercile and overall, the grey bar is substantially lower than the green bar, i.e., adaptation reduces predicted mortality. In each case, the magnitude of these differences is large, with the mortality effect shrinking by over 60 percent for the coolest third of ZIPs and actually becoming negative (relative to baseline) for the two other terciles and for the overall total.

The geographical impacts of our results are summarized in Figure 6, which maps the estimated mortality impact of climate change by ZIP code under the three cases that we simulate. Panel A, which assumes homogeneous temperature effects, shows that the states that are currently the hottest, in the Deep South and desert southwest, are the states that will suffer the highest mortality increases. Much of the coldest parts of the country, in the northeast, upper Midwest, and northwest, will actually see a decrease in

---

[21] The increased vertical dispersion of the green dots is not present under adaptation because when ZIP codes adapt to their future temperatures, very few days fall on steep parts of the temperature-mortality curve, which reduces the influence of small changes in the temperature distribution on average mortality.

mortality due to climate change. This geographic pattern mirrors the result found in Hsiang et al. (2017),[22] which also assumes homogeneous effects and no future adaptation. Panel B of Figure 6 which maps mortality predictions allowing for current-climate heterogeneity but not future adaptation, dramatically reverses this result. Here, the coldest places are those in which the mortality impacts are the highest. Only in some of the hottest places are mortality effects negative. Thus not only does accounting for heterogeneity change the overall mortality impact of climate change, it also markedly changes the geographic impact, essentially reversing which areas are expected to be hit hardest.

Finally, Panel C of Figure 6, incorporates future adaptation and shows that the mortality effects are decreased everywhere relative to the middle panel but the geographical pattern – more negative impacts for colder places than for hotter places – remains. The map shows that if ZIP codes in the future adapt to climate change in the sense that they adopt the technologies used by ZIP codes that currently experience their future climate, roughly the southern half of the country is estimated to experience either no increase or an actual decrease in temperature related in mortality. Thus, once again in contrast to the homogeneous current effects, no future adaptation results, these results suggest the greatest mortality impact of climate change, net of adaptation, is likely to be felt across the Northernmost U.S. states.

We find substantial scope for adaptation to mitigate, and even reverse, the potential mortality impact of climate change. This is analogous to the result in Butler and Huybers (2013), who find that losses to average maize yields drops from 14% to 6% after accounting for adaptation. However, it is worth stressing that our calculation is subject to a few caveats. We are not claiming that climate change will necessarily be a boon for elderly health and well-being. In one sense, the 0.62 percent decrease in annual deaths is a best-case scenario for adaptation, if we assume that our modeling of adaptation represents each ZIP fully adopting the technology of ZIPs that already experience those climates. Even if such adaptation is technologically possible, our simulations do not address the cost of adaptation or the transition path. However, even if it were possible for adaptation to substantially reduce mortality at a very low cost, it is important to note that mortality is only one input into elderly welfare. If adaptation to heat involves staying indoors and running the air conditioner, then a decrease in utility from outdoor activities may offset some or all of the mortality benefit of adaptation relative to the current situation.[23]

Another fundamental reason to be cautious about these simulations is that they assume that the temperature-mortality relationship—or, in the case under the adaptation assumption, the temperature-mortality-CDD relationship—will remain the same in the future as it is now. Over time, however,

---

[22] In particular, the map is nearly identical to their Figure 2, Panel B, though their mortality estimates are based on all-age mortality, not just elderly mortality, and are at the county rather than ZIP code level.
[23] Similarly, Schlenker et al. (2013) criticize the adaptation predictions of Butler and Huybers (2013) for failing to account for adaptation costs.

demographics may change or migration may take place in response to climate change. In addition, new technologies may make adaptation more complete or less costly, further lowering the mortality impact of climate change. On the other hand, generally warmer global temperatures may lead to changes in sea-levels, agriculture, vector-borne disease prevalence and other factors that will fundamentally alter the relationship between temperature and mortality and may make high temperatures even more dangerous.

V. CONCLUSION

This paper predicts the impact of climate change on U.S. elderly mortality and demonstrates the importance of accounting for regional heterogeneity in the relationship between temperature and mortality and adaptation to climate change. The exceptional temporal and geographic scope afforded by daily mortality data on the universe of elderly Medicare beneficiaries over two decades allows us to explore heterogeneity in temperature effects across local climates and make projections of future adaptation to a degree that has not been done before. This regional heterogeneity provides strong evidence that regions have adapted to their current climate, with residents of cooler climates being more vulnerable to extreme heat but less vulnerable to extreme cold than residents of hotter climates. Differential air conditioning adoption can explain essentially all the variation in heat-related mortality between cooler and warmer climates, suggesting that cool climates may have significant scope to adapt to future warming.

Using both simple counterfactual exercises and predictions of the actual change in the temperature distribution from 21 climate projection models, we demonstrate that failure to incorporate climate heterogeneity when predicting the effects of climate change can result in regional predictions that substantially understate the mortality impact of warming and mischaracterize which regions are likely to suffer the most. For example, in contrast to Hsiang et al. (2017), who ignore regional heterogeneity in the temperature-mortality relationship, we find that it is the coldest regions, rather than the warmest ones, that are likely to experience the greatest increase in mortality. And since, as Hsiang et al. (2017) show, mortality costs are expected to make up around 70 percent of the overall economic damage of climate change, our results imply that the geographic distribution of the overall cost of climate change is likely to change substantially as well. Moreover, we allow regions to adapt to future climate according to the degree of climate adaptation currently observed in the cross-section. This yields mortality impacts of climate change that are much lower than those calculated without allowing adaptation, and possibly even negative. Although we do not consider the cost of adaptation, our results show that there is significant ability to moderate the mortality impact of future warming even using technologies that are readily available today. Because our method for calibrating the adaptation to future climate changes is based on historical adaptation involving currently available adaptive technologies, the potential for future technological change to reduce the costs of adaptation may lower the mortality effect of climate change even further.

References

Angrist, Joshua D., and Alan B. Krueger. "Empirical strategies in labor economics." *Handbook of labor economics* 3 (1999): 1277-1366.

Auffhammer, Maximilian. "Climate Adaptive Response Estimation: Short and Long Run Impacts of Climate Change on Residential Electricity and Natural Gas Consumption Using Big Data." mimeo, University of California at Berkeley, 2017.

Auffhammer, Maximilian, Solomon M. Hsiang, Wolfram Schlenker, and Adam Sobel. "Using weather data and climate model output in economic analyses of climate change." *Review of Environmental Economics and Policy* (2013): ret016.

Baccini, Michela, Annibale Biggeri, Gabriele Accetta, Tom Kosatsky, Klea Katsouyanni, Antonis Analitis, H. Ross Anderson et al. "Heat effects on mortality in 15 European cities." *Epidemiology* 19, no. 5 (2008): 711-719.

Barreca, Alan, Karen Clay, Olivier Deschênes, Michael Greenstone, and Joseph S. Shapiro. "Adapting to Climate Change: The Remarkable Decline in the U.S. Temperature-Mortality Relationship over the 20th Century." *Journal of Political Economy*, 124 (2016), 105-159.

Barreca, Alan, Karen Clay, Olivier Deschênes, Michael Greenstone, and Joseph S. Shapiro. "Convergence in Adaptation to Climate Change: Evidence from High Temperatures and Mortality, 1900 – 2004." *American Economic Review: Papers and Proceedings* 105, no. 5 (2015): 247-251.

Beatty, Timothy KM, and Jay P. Shimshack. "Air pollution and children's respiratory health: A cohort analysis." *Journal of Environmental Economics and Management* 67, no. 1 (2014): 39-57.

Burke, Marshall and Kyle Emerick."Adaptation to Climate Change: Evidence from U.S. Agriculture." *American Economic Journal: Economic Policy*, 8(3) (2016): 106-40.

Butler, Ethan E., and Peter Huybers. "Adaptation of U.S. maize to temperature variations." *Nature Climate Change* 3, no. 1 (2013): 68-72.

Currie, Janet and Matthew Neidell. "Air Pollution And Infant Health: What Can We Learn From California's Recent Experience?" *Quarterly Journal of Economics*, 2005, v120(3,Aug), 1003-1030.

Curriero, Frank C., Karlyn S. Heiner, Jonathan M. Samet, Scott L. Zeger, Lisa Strug, and Jonathan A. Patz. "Temperature and mortality in 11 cities of the eastern United States." *American Journal of Epidemiology* 155, no. 1 (2002): 80-87.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. "Temperature shocks and economic growth: Evidence from the last half century." *American Economic Journal: Macroeconomics* 4, no. 3 (2012): 66-95.

Deschênes, Olivier. "Temperature, human health, and adaptation: A review of the empirical literature." *Energy Economics* 46 (2014): 606-619.

Deschênes, Olivier, and Michael Greenstone. "Climate change, mortality, and adaptation: evidence from annual fluctuations in weather in the U.S." *American Economic Journal: Applied Economics* 3, no. 4 (2011): 152-185.

Fried, Stephie, and Tamara Sheldon. "Climate Adaptation: Evidence From Extreme Weather." Working Paper (2016).

Gibbons, Charles E., and Juan Carlos Suárez Serrato. "Broken or Fixed Effects." NBER Working Paper #20342 (2017).

Graff Zivin, Joshua, and Matthew Neidell. "Temperature and the allocation of time: Implications for climate change." *Journal of Labor Economics* 32, no. 1 (2014): 1-26.
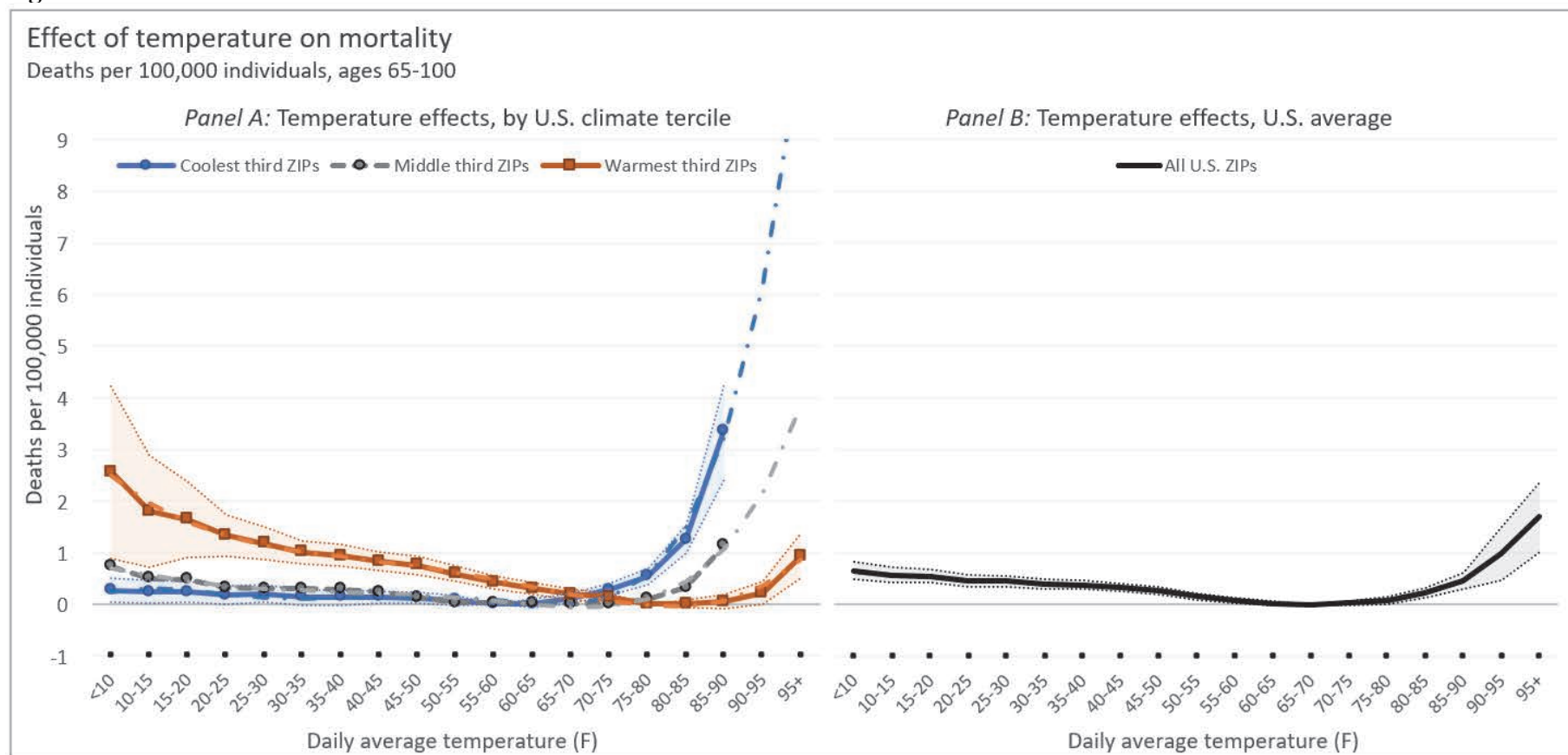
Houser, Trevor, Solomon Hsiang, Robert Kopp, Kate Larsen, Michael Delgado, Amir Jina, Michael Mastrandrea et al. "American climate prospectus." *Economic Risks in the United States*. Rhodium Group, LLC (2014).

Hondula, David M., Robert C. Balling Jr, Jennifer K. Vanos, and Matei Georgescu. "Rising Temperatures, Human Health, and the Role of Adaptation." *Current Climate Change Reports* (2015): 1-11.

Hsiang, Solomon. "Climate econometrics." *Annual Review of Resource Economics* 8 (2016): 43-75.

Hsiang, Solomon M., and Daiju Narita. "Adaptation to cyclone risk: Evidence from the global cross-section." *Climate Change Economics* 3, no. 02 (2012): 1250011.

Hsiang, Solomon, Robert Kopp, Amir Jina, James Rising, Michael Delgado, Shashank Mohan, D. J. Rasmussen et al. "Estimating economic damage from climate change in the United States." *Science* 356, no. 6345 (2017): 1362-1369.

Huynen, MMTE-Martens, Pim Martens, Dieneke Schram, Matty P. Weijenberg, and Anton E. Kunst. "The impact of heat waves and cold spells on mortality rates in the Dutch population." *Environmental Health Perspectives* 109, no. 5 (2001): 463.

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Pachauri, R.K and Meyer, L.A. (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Kahn, Matthew E., 2015. "Climate Change Adaptation: Lessons from Urban Economics," *Strategic Behavior and the Environment*, 5, no. 1 (2015): 1-30.

Kahn, Matthew E. "The Climate Change Adaptation Literature." *Review of Environmental Economics and Policy* (2016): rev023.

Meinshausen, Malte, Steven J. Smith, K. Calvin, John S. Daniel, M. L. T. Kainuma, J. F. Lamarque, K. Matsumoto et al. "The RCP greenhouse gas concentrations and their extensions from 1765 to 2300." *Climatic Change* 109, no. 1-2 (2011): 213-241.

Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw. "The impact of global warming on agriculture: a Ricardian analysis." *The American Economic Review* (1994): 753-771.

Portnykh, Margarita. "The Effect of Weather on Mortality in Russia: What if People Adapt?" Working Paper, 2017.

Ren, Cizao, Gail M. Williams, and Shilu Tong. "Does particulate matter modify the association between temperature and cardiorespiratory diseases?" *Environmental Health Perspectives* 114, no. 11 (2006): 1690.

Rogot, Eugene, Paul D. Sorlie, and Eric Backlund. "Air-conditioning and Mortality in Hot Weather." *American Journal of Epidemiology* 136, no. 1 (1992): 106-116.

Schlenker, Wolfram, Michael J. Roberts, and David B. Lobell. "U.S. maize adaptability." *Nature Climate Change* 3, no. 8 (2013): 690.

USGCRP. *Global Climate Change Impacts in the United States*. Thomas R. Karl, Jerry M. Melillo, and Thomas C. Peterson (eds.). United States Global Change Research Program. Cambridge University Press, 2009.

Figure 1

U.S. Daily Temperature Distribution
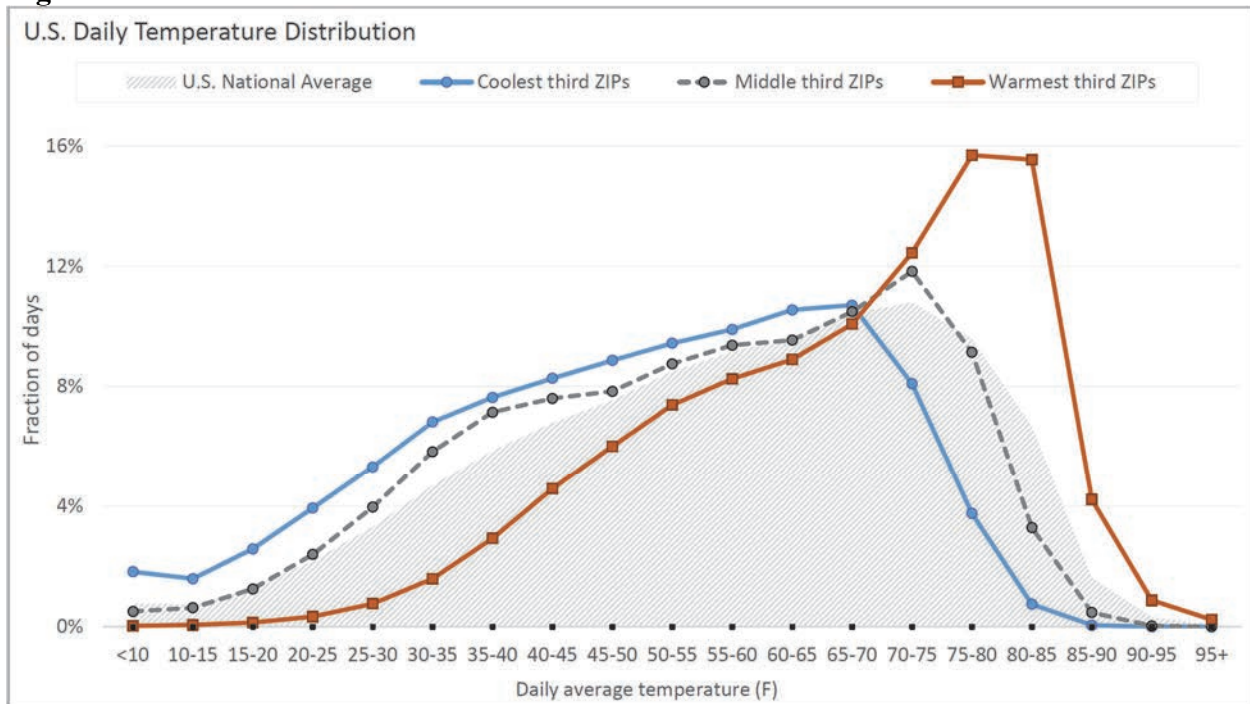
Notes: This figure summarizes the distribution of daily temperature in the United States over the period 1992-2011, based on climatological data from the National Oceanic Atmospheric Administration's National Centers for Environmental Information. The coolest, middle, and warmest third of ZIPs partition U.S. ZIP codes into population-weighted thirds based on 30-year average cooling degree days. Statistics are calculated over daily observations at the ZIP code level and are weighted by the elderly Medicare population in each ZIP. Each 5-degree bin contains all ZIP-day observations where the average temperature (calculated as the mean of the daily high and low temperatures) falls within the bin range.
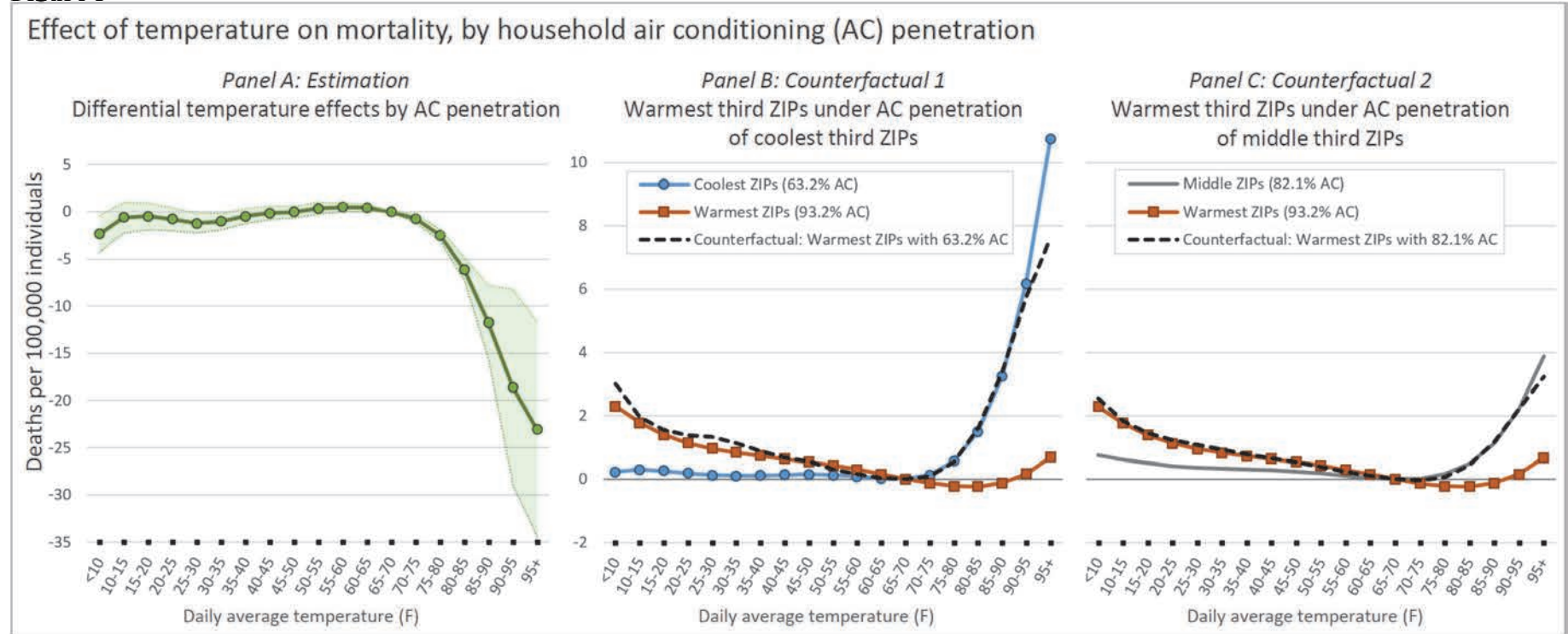
Figure 2

Effect of temperature on mortality
Deaths per 100,000 individuals, ages 65-100

Panel A: Temperature effects, by U.S. climate tercile — Coolest third ZIPs, Middle third ZIPs, Warmest third ZIPs

Panel B: Temperature effects, U.S. average — All U.S. ZIPs

Notes: Panel A plots estimates from a daily ZIP code level regression of 3-day mortality on daily average temperature where temperature bins are interacted with an indicator for whether the ZIP code is in the coolest, middle, or warmest third of ZIP codes based on 30-yar average cooling degree days (see Equation 1). The temperature estimates for each climate region reflect excess mortality from a given temperature day relative to a day in the mortality-minimizing temperature category for that region: 55°F-60°F, 65°F-70°F, and 75°F-80°F for the coolest, middle, and warmest regions, respectively. Estimates are reported and indicated with a marker for temperature bins comprising at least 0.05 percent of daily ZIP code observations (about two days every ten years, on average) in that climate region. The dashed curves are 5th order polynomials fit to the reported estimates for each tercile and fit to all temperature bins.

Panel B plots estimates from an analogous regression where the bin coefficients are assumed to be the same across all ZIP codes (see Appendix Equation A1). The estimate for each bin reflects excess mortality associated with replacing a day in the reference category (65°F-70°F) with a day in the specified bin.

Regressions in both panels include ZIP code by day of year fixed effects, state by year fixed effects, and flexible controls for temperature in the preceding 2 and 6 days and subsequent 2 days. Regressions are weighted by population and shaded regions report 95 percent confidence intervals based on standard errors clustered at the county level.

Figure 3

Effect of temperature on mortality, by household air conditioning (AC) penetration

Notes: Panel A reports estimates of how the mortality effect of each temperature bin varies with regional AC penetration. The estimates come from estimating a version of Equation (1) expanded to include daily average temperature bins interacted with ZIP code-level AC penetration as well separate temperature controls for each Census Region. The shaded region reports 95 percent confidence intervals based on standard errors clustered at the county level. *Panel B* and *Panel C* present the implied counterfactual mortality effects of temperature in the warmest third of ZIP codes if exposed to the lower AC penetration rates in the coolest and middle third of ZIP codes, respectively.

Figure 4

Predicted mortality effect of regional climate swaps

Mortality change in each climate tercile from replacing current climate with that of another tercile
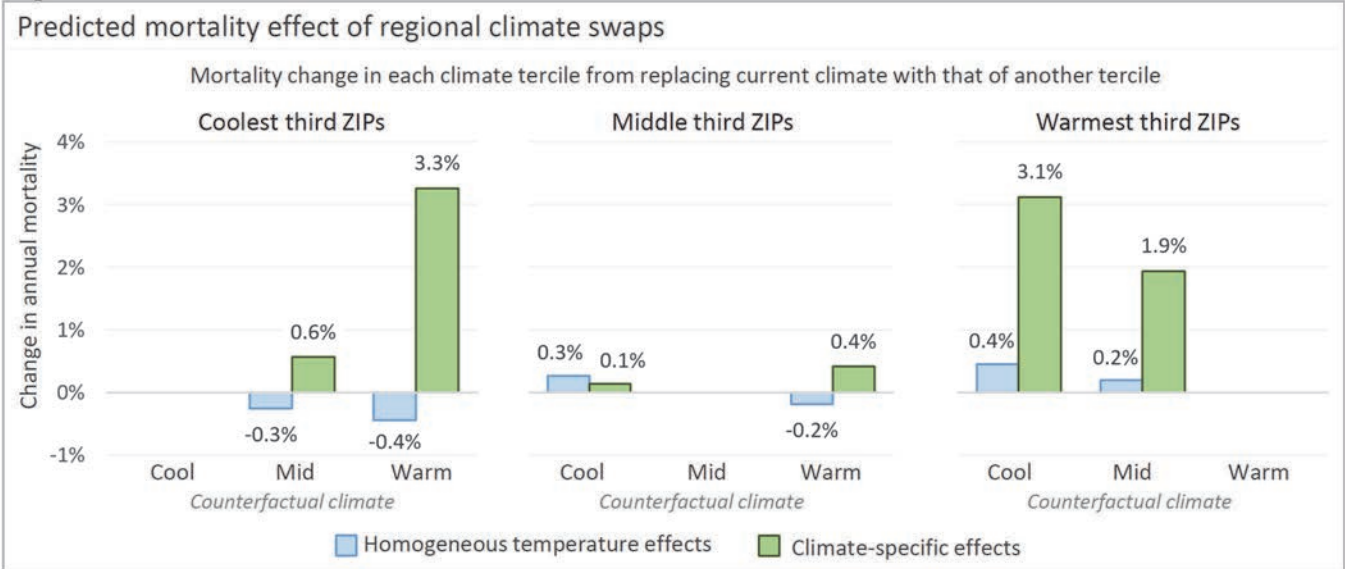
Notes: This figure summarizes business-as-usual mortality impacts from counterfactual scenarios in which each of these climate region's current distribution of temperature is replaced by the current distribution of one of the other two climate regions (see Figure 1).

Figure 5

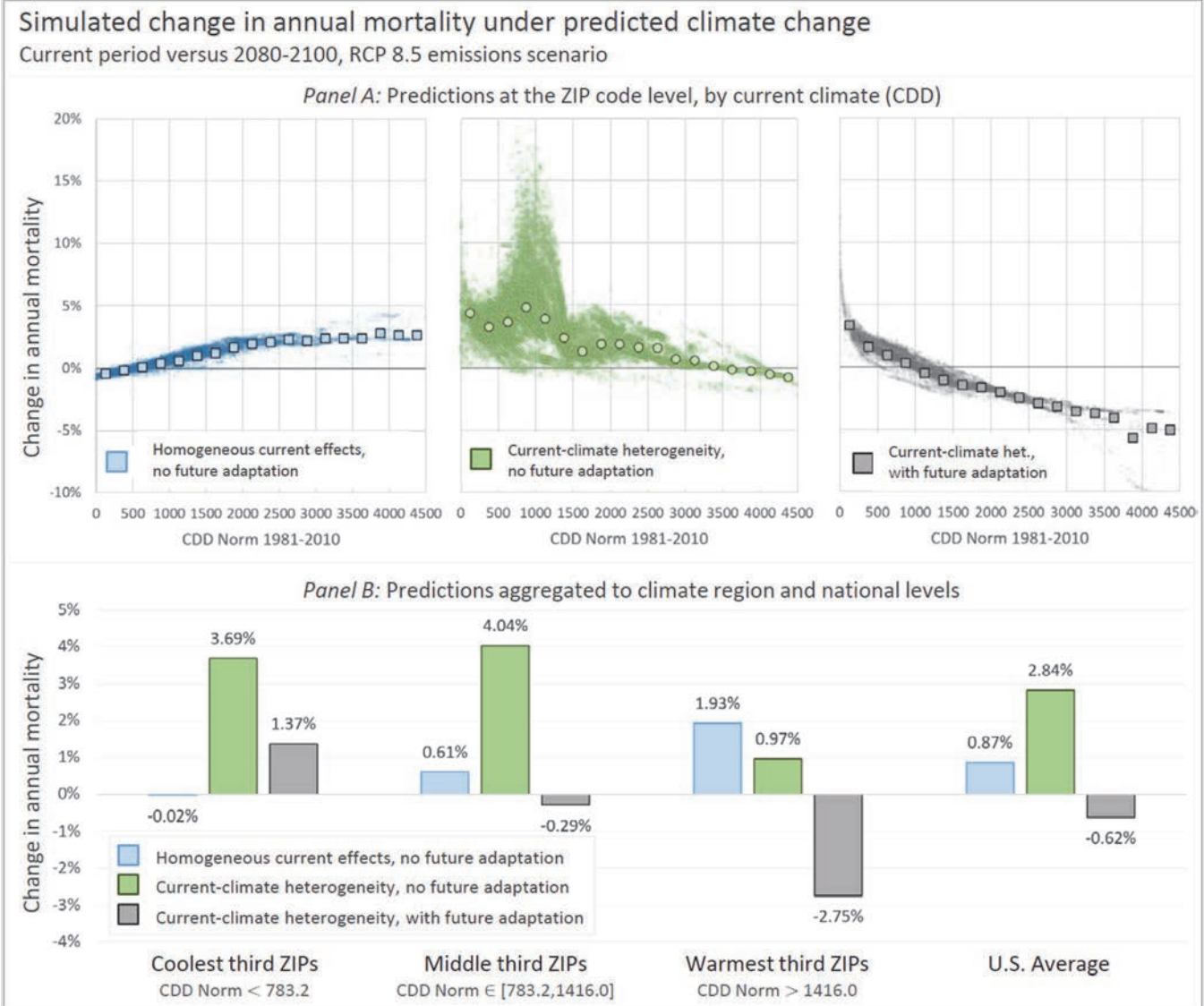Simulated change in annual mortality under predicted climate change
Current period versus 2080-2100, RCP 8.5 emissions scenario

Panel A: Predictions at the ZIP code level, by current climate (CDD)

Homogeneous current effects, no future adaptation

Current-climate heterogeneity, no future adaptation

Current-climate het., with future adaptation

CDD Norm 1981-2010

Change in annual mortality

Panel B: Predictions aggregated to climate region and national levels

Change in annual mortality

3.69%
1.37%
-0.02%

4.04%
0.61%
-0.29%

1.93%
0.97%
-2.75%

2.84%
0.87%
-0.62%

■ Homogeneous current effects, no future adaptation
■ Current-climate heterogeneity, no future adaptation
■ Current-climate heterogeneity, with future adaptation

Coolest third ZIPs
CDD Norm < 783.2

Middle third ZIPs
CDD Norm ∈ [783.2,1416.0]

Warmest third ZIPs
CDD Norm > 1416.0

U.S. Average

Notes: This figure summarizes results from predicting the mortality impact of end-of-century (2080-2100) climate change as projected by 21 climate models under the RCP 8.5 emissions scenario.

Figure 6

Predicted change in elderly annual mortality

A. Homogeneous current effects, no future adaptation

B. Current-climate heterogeneity, no future adaptation

C. Current-climate heterogeneity with future adaptation

-1.50   -0.75   0   0.75   1.50   3.00
Mortality change (%)
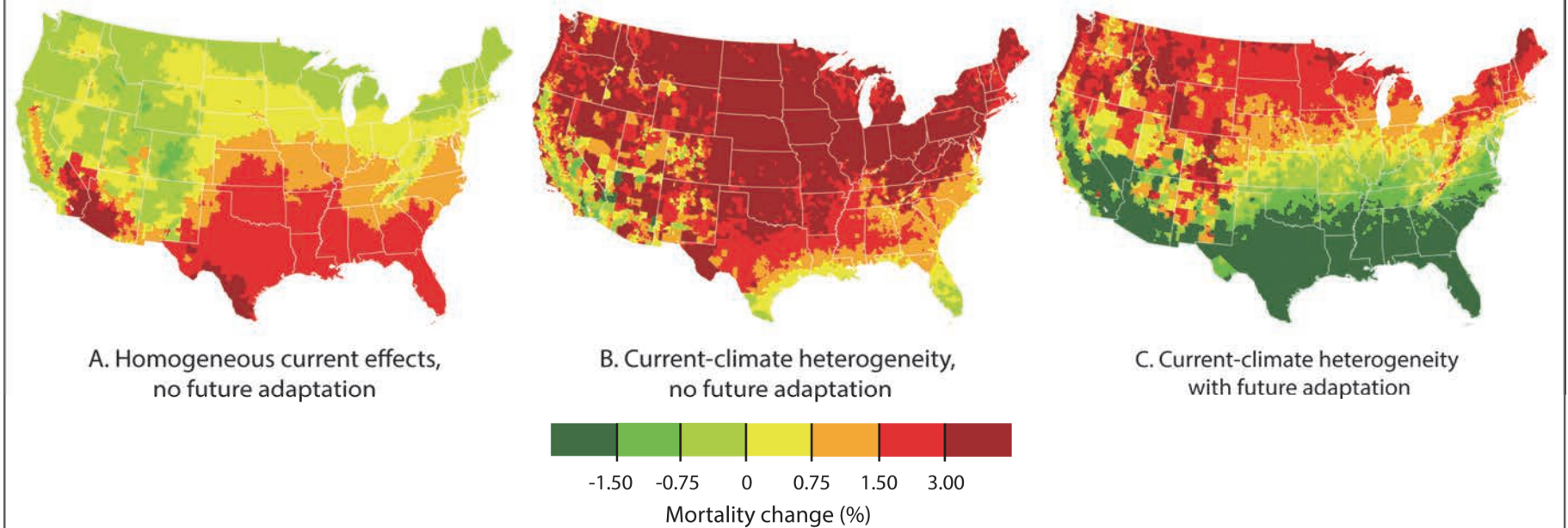
Notes: This figure summarizes results from predicting the mortality impact of end-of-century (2080-2100) climate change as projected by 21 climate models under the RCP 8.5 emissions scenario. Predictions are made at the ZIP code level.

Table 1

Daily average temperature bin (F)	Median average temperature	Median low temperature	Median high temperature	Daily mortality per 100,000
(1)	(2)	(3)	(4)	(5)
<10	5.3	-5.1	14.6	14.9
10-15	12.9	3.6	21.9	15.0
15-20	17.9	9.1	26.4	15.0
20-25	22.8	14.6	30.8	14.9
25-30	27.7	19.9	35.3	14.8
30-35	32.7	25.0	40.3	14.6
35-40	37.6	28.9	46.1	14.5
40-45	42.5	32.9	52.2	14.2
45-50	47.6	37.1	57.9	13.9
50-55	52.5	41.8	63.3	13.5
55-60	57.5	46.5	68.7	13.2
60-65	62.5	51.2	73.9	12.8
65-70	67.6	56.3	78.8	12.6
70-75	72.5	62.0	83.1	12.4
75-80	77.4	67.4	87.5	12.4
80-85	82.0	72.4	91.9	12.3
85-90	86.4	75.4	98.1	12.3
90-95	91.8	78.9	105.1	11.9
95+	96.6	83.8	110.0	11.8
Total	59.5	48.4	70.9	13.3

Notes: This table presents raw summary statistics across each of 19 temperature bins. Statistics are calculated over daily observations at the ZIP code level and are weighted by the elderly Medicare population in each ZIP code. Each temperature bin contains all ZIP-day observations where the average temperature (calculated as the mean of the daily high and low temperatures) falls within the bin range. Mortality rates are computed for the elderly (65+) Medicare population over the period 1992-2011.

Appendix (for Online Publication)

A Data Descriptions

A.1 Medicare Data

Our baseline sample consists of all Medicare beneficiaries age 65 – 100 and is derived from 100% Medicare enrollment information files for years 1992 – 2011.[1] These annual files include an observation for each beneficiary enrolled in Medicare for at least one day in that calendar year, whether enrolled in Original Medicare (fee-for-service) or Medicare Advantage. The enrollment files report a variety of demographic and enrollment variables, including unique beneficiary identifiers that link individuals over time; monthly indicators for Medicare eligibility; state, county, and ZIP code of residence based on the mailing address for official correspondence; and date of birth, date of death, and gender.

Medicare beneficiaries include the vast majority of elderly living in the United States. Panel A of Appendix Figure A1 compares the size of our baseline Medicare sample to Census estimates of the U.S. population age 65 and over. To aid comparison, we use Census estimates of the resident population on July 1 each year and limit the Medicare sample to beneficiaries residing in the 50 states and the District of Columbia and who turned 65 before July 1. Over the period 1992 – 2011, the Census estimates an average of 36.3 million elderly individuals each year, compared to 35.4 million elderly beneficiaries in Medicare. Thus, the Medicare sample covers over 97% of elderly living in the U.S., a share which remains roughly constant over the sample period.

The mortality variables used in our analysis are based on dates of death recorded in the Medicare enrollment files. Medicare's death data come primarily from the Social Security Administration but are augmented based on reviews triggered by hospitalization claims indicating patient death. The annual mortality rates in the Medicare data align closely to mortality rates based on national vital statistics death records and Census population estimates, as shown in Appendix Figure A1, Panel B. While all recorded deaths in the Medicare data are validated, some death *dates* in the data are not validated and are assigned the last date in the month of death. Because much of our analysis is performed at the daily level, we drop individuals who die at any point in the year

---

[1] The Research Data Assistance Center (ResDAC) provides a helpful overview of the Medicare enrollment information files at http://www.resdac.org/training/workshops/intro-medicare/media/3.

and who do not have a validated death date flag. This restriction affects fewer than 3% of the deaths in our sample, with the share of deaths with unvalidated dates diminishing over time.

A.2   Temperature Data

Appendix Table A1 provides summary statistics on the coverage of the National Centers for Environmental Information (NCEI) weather monitors across our sample period.[2] The number of monitors increases almost monotonically from 6,786 in 1992 to 9,017 in 2011. For all years, the number of non-missing days for both temperature observations that are used to create the average daily temperature (the daily maximum and minimum temperatures) is large, and most days of the year are covered. More than 98% of ZIP codes are covered by at least one monitor in all years (that is, for 98% of ZIP codes there is at least one monitor within 20 miles of the ZIP code centroid, though not necessarily within the ZIP code). The number of monitors per ZIP code averages from 3.47 to 4.86.

A.3   Air Conditioning Data

Our data on air conditioning are derived from the Residential Energy Consumption Survey (RECS) which is administered by the Energy Information Administration (EIA) and surveys a nationally representative sample of housing units in the US across 50 States and the District of Columbia. The RECS collects information on energy consumption and energy-related characteristics of housing units, including air conditioning. We use this information to estimate a model of air conditioning as a function of the location and characteristics of housing units. Specifically, we estimate a housing-unit level regression of whether the housing unit has air conditioning (AC) equipment as a flexible function of cooling degree days (a fourth-order polynomial) in the housing unit location, year built, number of rooms, urban location, mobile home status, and ownership status. We interact each of these housing characteristics with dummies for Census Region and Census Division. To avoid overfitting, we use LASSO penalized regression and select the penalty to minimize 10-fold cross-validated mean squared prediction error. We then apply this model to ZIP code level housing characteristics from the Census to compute predicted air conditioning for each U.S. ZIP code.

This process requires constructing a set of comparable housing characteristics in both RECS and Census data. Below, we describe the housing characteristics we use and how we define them in terms of original RECS and Census variables.

---

2 Available from http://www.ncdc.noaa.gov/. The NCEI was formerly the National Climatic Data Center (NCDC).

A.3.1 Residential Energy Consumption Survey Data

We use RECS data from survey years 1993, 1997, 2001, 2005, and 2009. We model whether a housing unit in the RECS data has air conditioning as a function of housing-level characteristics including the housing unit type, the geographic location of the housing unit, and the climate for the location of the unit. Below we describe the specific variables used in this model and how these variables are defined in terms of the original RECS variables from each year of the survey.

1. *Air conditioning (AC).* An indicator variable for whether a household reports to have air conditioning equipment at home. This corresponds to the RECS variable AIRCOND ("Do you have air conditioning equipment at home") in years 1993, 1997, 2001, and 2005. For year 2009, the AIRCOND survey question changed to whether air conditioning equipment is "used." To obtain a consistent measure of AC ownership, for 2009 we rely instead on the variable DNTAC ("No air conditioning equipment, or unused air conditioning equipment") which stratifies all households into three groups based on their report of AC ownership and usage status: (1) Have AC equipment but don't use it; (2) Have AC equipment and use it; (3) Don't have any AC equipment. We define a household unit to have AC equipment if it belongs to group (1) or (2).

2. *Year built.* A set of indicator variables (summing up to one for each household) identifying the decade when the housing unit was built (1939 or earlier, 1940 to 1949, …, 1990 to 1999, 2000 or later). These are based on the RECS variable YEARMADE ("Year home built") which reports the decadal interval when the home is built (mostly in early year) or the year the home is built. For each year, we group household units into decade-built according to the categorization used by the 2007-2011 5-year American Community Survey (ACS).

3. *Number of rooms.* A set of indicator variables (summing up to one for each household) identifying the number of rooms in the housing unit (1 room, 2 rooms, …, 8 rooms, 9 or more rooms). For years 1993, 1997, 2001, we count rooms by adding up RECS survey variables BEDROOMS ("Number of bedrooms") and OTHROOMS ("Number of other rooms"). We use the variable TOTROOMS ("Total number of rooms") which is available for year 2005 and 2009. We then assign number of rooms according to the categorization used by the 2007-2011 5-year ACS.

4. *Urban.* An indicator variable for whether the housing unit is in the urban area, i.e. non-rural area. This is based on the RECS variable UR (or URBRUR in early years). The RECS urban status information is obtained from interviewer observation (1993), household report (1997, 2001, 2005), and the Census Bureau's Urban/Rural geographic identifier (2009).

5. *Mobile*. An indicator variable for whether the housing unit is a mobile home. This is based on the RECS variable TYPEHUQ ("Type of home as report by respondent"). TYPEHUQ reports whether the housing unit is mobile, single-family detached, single-family attached, apartment in building with 2-4 units, or apartment in building with 5+ units.

6. *Own*. An indicator variable for whether the housing unit is owned. This is based on the RECS variable KOWNRENT ("Housing unit owned or rented"). KOWNRENT reports whether the housing unit is owned by someone in the household, rented, or occupied without payment.

7. *CDD65*. A continuous measure of cooling degree days (CDD) of the survey year with a base temperature of 65 Fahrenheit, i.e. the total number of degrees the daily temperature exceeds 65 Fahrenheit from January to December of the survey year. We draw this information from the RECS variable CDD65 (CD65 in earlier years). RECS creates this variable using temperature monitoring data from the National Climate Data Center's weather stations. A housing unit is first matched to the closest weather station, and then CDD is computed using the station's daily temperature data from January to December of the survey year. To mask the household location, a "random error" was added to the CDD. RECS does not provide further information regarding the structure of the random error.

8. *Census region*. A set of categorical variables (summing up to one for each household) identifying the Census region location of the housing unit. This corresponds to the RECS variable REGIONC ("Census Region") which includes Northeast, Midwest, South, and West Census Region.

9. *Census division*. A set of categorical variables (summing up to one for each household) identifying the Census division location of the housing unit. This corresponds to the RECS variable DIVISION ("Census Division"). For years 1993, 1997, 2001, and 2005, DIVISION divides housing units into New England (CT, MA, ME, NH, RI, VT), Middle Atlantic (NJ, NY, PA), East North Central (IL, IN, MI, OH, WI), West North Central (AR, LA, OK, TX), South Atlantic (DC, DE, FL, GA, MD, NC, SC, VA, WV), East South Central (AL, KY, MS, TN), West South Central (AR, LA, OK, TX), Mountain (AZ, CO, ID, MT, NM, NV, UT, WY), and Pacific Census Division (AK, CA, HI, OR, WA). In 2009, the Mountain Census Region is further divided into the Mountain North Sub-Division (CO, ID, MT, UT, WY) and the Mountain South Sub-Division (AZ, NM, NV). We combine the two groups to the single Mountain Division to keep the division identifier coherent across years.

10. *Sampling weights*. The sampling weight variable NWEIGHT provided by the RECS. This is the sampling weight for the observation in each survey year, which equals approximately to the inverse of the probability of selection into the sample. RECS applies a couple of adjustments to the weights, including (1) adjustments to interview nonresponse; (2) post-stratification for matching total energy consumptions by fuel types; (3) benchmarking which ensures that the total RECS weights add up to the ACS's total number of occupied housing units.

A.3.2 Census Data

We use ZIP code level housing characteristics from the 2007-2011 5-year American Community Survey (ACS) and the 2010 Census to construct a comparable set of housing characteristics as used for predicting AC using the LASSO model fit to the RECS data described above. We obtained Census data provided by the Minnesota Population Center (2016) through the National Historical Geographic Information System (NHGIS), and the variable names we refer to below are from the NHGIS.

1. *Year built*. Eight variables each measuring the fraction of housing units that were built 1939 or earlier, 1940 to 1949, …, 1990 to 1999, and 2000 or later. We create these variables from the ACS variable "Year Structure Built" which counts total number of housing units built in each of the decadal intervals listed above. We convert counts to fractions by dividing the total number of housing units in the ZIP code.

2. *Number of rooms*. Nine variables each measuring the fraction of housing units that have 1 room, 2 rooms, …, 8 rooms, and 9 or more rooms. These are based on the ACS variable "Rooms." We convert counts to fractions by dividing the total number of housing units in the ZIP code.

3. *Urban*. Fraction of urban population from 2010 Census data. We divide the urban population in the ZIP code by the total population.

4. *Mobile*. Fraction of housing units that are mobile home. We base this on the ACS variable "Units in Structure" which provides number of mobile homes in the ZIP code, and we divide counts by the total number of housing units in the ZIP code.

5. *Own*. Fraction of occupied housing units that are owned. We draw this information from the ACS variable "Tenure by Household Size." We divide number of owner occupied housing units by the total number of occupied housing units in the ZIP code.

6. *CDD65*. Cooling degree days for the ZIP code, obtained from NOAA Climate Normals over the period 1980-2010 using inverse distance weighting of all monitors within a 20-mile radius of the ZIP code centroid.

7. *Census region*. A set of categorical variables (summing up to one for each ZIP code) identifying the Census region of the ZIP code centroid. We use US Census Bureau's standard definition of Census regions.

8. *Census division*. A set of categorical variables (summing up to one for each ZIP code) identifying the Census division of the ZIP code centroid We use US Census Bureau's standard definition of Census divisions.

B Homogeneous Mortality Effects of Temperature

In this section, we estimate the mortality effects of temperature the same way as in section III of the paper, but ignoring the heterogeneity of mortality effects across regions. That is, we estimate one single temperature-mortality curve for the entire country. We do this to demonstrate the importance of heterogeneity in temperature effects when making mortality predictions. While we recognize this bias, beginning with an examination of the homogeneous effects case provides a useful benchmark and point of comparison with the literature.

B.1 Empirical Strategy

The estimating equation is the same as equation (1) in the text, except that we do not allow for different coefficients for the three climate terciles:

$$mortality_{zd} = \sum_{b \in B \backslash \{65-70\}} \beta_b tempbin_{zd}^b + ZipDay_{zd} + L_{zd} + StYr_{zd} + \varepsilon_{zd}, \tag{A1}$$

where an observation is a ZIP code z on day d.

The primary independent variables of interest in Equation (A1) are temperature indicators $tempbin_{zd}^b$ defined by which of 19 temperature bins $B = \{< 10, 10 - 15, \dots, 90 - 95, 95+\}$ the average temperature in ZIP code z falls on day d. The omitted temperature category consists of days with average temperature in the 65-70°F range. As a result, the coefficients β_b capture the difference in mortality from a typical day with average temperature in bin b relative to a typical 65-70°F day.

B.2 Results

Figure 2, Panel A, depicts the results of estimating Equation (A1) with 3-day mortality as the outcome. The black line plots the coefficients β_b, which capture the difference in mortality from

a typical day with average temperature in bin *b* relative to a day with average temperature in the 65-70°F range. The shaded region around the solid black line represent the 95 percent confidence interval for the point estimates. The estimates imply that both hot and cold temperatures lead to increases in mortality relative to more moderate temperatures. For example, substituting a typical 65-70°F day for a typical day in the highest temperature bin (95°F +) generates about 1.7 additional deaths per 100,000 people. Substituting instead for a cold day with average temperature below 10°F generates around 0.7 additional deaths per 100,000 people compared to a 65-70°F day. In Appendix Figure A3, we present results from several robustness specifications which vary the structure of temperature leads and lags as well as the level at which standard errors are clustered (ZIP code, county, state, or ZIP code-date two-way clustering), all of which yield results very similar to our baseline specification.

C Robustness: Mortality Displacement (Harvesting)

A natural concern in the case of temperature-related deaths is that the observed increase in deaths does not represent a substantial loss of life years. Rather, it represents a moving forward in time of deaths that would have occurred in a few days or weeks even in the absence of the weather event. Implicit in this view is the idea that the event attacks the frailest people in the population, so that those who die were likely to have had few life years remaining. The extent of harvesting has important implications for policymakers, since the cost in terms of life years lost is much lower if decedents due to extreme temperatures would have been expected to live only a short time if the weather shock had not occurred compared to the case where decedents are randomly selected from the population.

As noted in the survey by Oudin Åström et al. (2011), relatively few studies in the public health and epidemiology literatures have investigated the role of mortality displacement in heat-related mortality. What evidence exists tends to be mixed and dependent on the location of the study. Hajat et al. (2005) find mixed results, with evidence of short-term harvesting after heat waves in London but not in Dehli. In the Czech Republic, Kyselý (2004) finds a relative rise in total mortality of 13% during heat waves, but a net mortality effect of only 1%, attributing the rest to short-term mortality displacement. Toulemon and Barbieri (2008) find evidence of modest harvesting during the 2003 heat wave in France. To the extent that a clear pattern emerges, it is that excess mortality during hot days tends to be due to short term displacement, at least to some extent, while the excess mortality during cold periods tends to build over time (Braga et al, 2001; Huynen et al., 2001; Deschênes and Moretti, 2009; Deschênes and Greenstone, 2011).

We investigate the possibility of harvesting in two ways. First, we expand the after-event window to include not only the event day and the following two days (3-day mortality), but also the following six days (7-day mortality) and 27 days (28-day mortality). We also compare specifications that look only at deaths on the event day itself (1-day mortality). If the additional deaths are due to very near-term harvesting, then an increase in deaths immediately following a temperature event should be followed by a decrease in deaths shortly thereafter. In this case, as the post-event window becomes longer, the net mortality effect over that window would become smaller. On the other hand, if the mortality effect of a temperature event remains stable or increases as the post-event window expands, this suggests that the estimated mortality effect is not simply an artifact of very short-term harvesting.

Appendix Figure A4 depicts regression results for 3-, 7-, and 28-day mortality.[3] For the hottest days, we see that 3- and 7-day excess mortality is higher than 1-day, and 28-day excess mortality is higher still, with the point estimate of the 28-day mortality of a 95+ degree day (2.2 excess deaths per 100,000 individuals) being about 30% larger than the 3-day impact (1.7 excess deaths per 100,000 individuals), although these effects cannot be statistically distinguished at conventional levels. This "build up" effect suggests that excess mortality takes time to manifest itself after the initial event, contrary to the harvesting story.[4] For moderately hot temperatures (80-85°F, 85-90°F, 90-95°F), we see little difference between the effects for the 3-, 7-, and 28-day windows. This result suggests harvesting in the very near term, i.e., that deaths from moderately hot days on the event day are offset by reduced mortality over the next two days.

Expanding the event window after cold days shows evidence of "build up," where mortality increases as the event window expands. For 28-day mortality, the effects of cold weather are larger in magnitude than the effects of hot weather, with a day in the 10-15°F range increasing mortality by 2.9 deaths per 100,000 people relative to a 65-70°F day, an effect that is slightly larger than that of very hot day. Interestingly, during our sample, very cold days (<10°F) occur almost 37 times as often as the hottest (>95°F) days (1.47% vs. 0.04%). Ignoring climatic heterogeneity in the effects of temperature, this suggests that the overall mortality impact of increasing

---

[3] All four of the regressions in Figure 3 use an identical set of lead and lag controls, although it is a different set of controls than used in the regressions in Figure 2. Because the post-event window goes up to 28 days, we include bins for the average temperatures in: the previous six days, the previous 2 days, the subsequent 2 days, the subsequent 6 days, and the subsequent 27 days. We also include all interactions of these bins. We omit plotting the standard errors for clarity.

[4] It is possible that harvesting, which reduces future mortality relative to the present, and delayed effects, which increases excess mortality in the future, take place simultaneously. In this case, increasing excess mortality as the event window increases only suggests that the buildup effect is larger than the harvesting effect. Because of this issue, below we directly investigate the hypothesis that those who are killed by the heat event are relatively frail by directly investigating the characteristics of who dies following the heat event.

temperatures may be negative if the reduction in the likelihood of extreme cold more than offsets the increased likelihood of extreme heat.

Our second approach to the question of harvesting begins with the observation that the essence of the "harvesting" critique is that deaths due to high heat are concentrated in the frailest individuals and represent smaller welfare losses since those who are killed only lose a few days or weeks of life. This question is of central importance to policymakers, since it has the potential to significantly affect the value of mortality reductions due to efforts to moderate or adapt to climate change. The rich demographic data provided by the Medicare files allow us to directly investigate the question of whether those who die on hot or cold days are more frail than the population overall and to directly estimate the loss in life expectancy. To do this, we investigate how mortality effects differ across demographic groups, and we also calculate estimates of the effect of temperature on life-years lost, rather than on mortality counts, using demographic data to infer decedents' counterfactual life-years remaining.

Appendix Figure A5 depicts regression coefficients for 3-day mortality regressions, separately estimated by five age groups (65-69, 70-74, 75-79, 80-84, and 85-100). (Appendix Table A2 reports the fraction of our sample within each age group and the mortality rate across age groups.) The largest marginal effect of both hot and cold days is among the oldest age groups. For 85-100-year-olds, a 95+ degree day causes an additional 4 deaths per 100,000 relative to the excluded temperature bin, while that effect is just about half as much for the next oldest group (80-84) and more than five times as large as for the youngest group (65-69). A <10°F day causes an additional 2.5 deaths per 100,000 among 85-100-year-olds; for all other age groups the effect of this coldest temperature bin is about 0.5 deaths. Hot days exhibit a roughly monotonic relationship between age and the mortality coefficient, but cold days have a larger effect only for the oldest age group. Overall Appendix Figure A5 demonstrates that the mortality effects of temperature are likely to affect the oldest individuals more than the youngest (among the Medicare population). This relationship is not necessarily unexpected, however, since the oldest cohorts have the highest mortality rates anyway: a proportionate increase in all cohorts' mortality rates could lead to the pattern observed in Appendix Figure A5, at least for hot days.

Nevertheless, Appendix Figure A5 suggests that if mortality effects are concentrated in the oldest cohorts, then mortality counts or rates that weight all individuals equally may overstate the impact of extreme temperatures. In other words, the life-years lost (i.e., mortality cost) from extreme temperatures may be moderated by the fact that the oldest individuals (with the fewest counterfactual life-years remaining) are the most susceptible. We investigate this possibility by estimating our main regression equation where mortality is weighted by the decedent's expected life years remaining using estimates provided by the Social Security administration based on age

and gender.[5] We assign each individual alive at the start of the day a value of expected remaining life-years using this table, which in turn implies the number of expected life years lost for each death. This procedure is likely to generate an upper bound on the true effect of temperature on life-years lost, since those who are killed by extreme temperatures are likely to be less healthy than the average individual within an age-gender cell. We sum the total expected life-years lost per 100,000 individuals and use this sum as our outcome measure. Regression results are reported in Appendix Figure A6.

While Appendix Figure A5 demonstrates that the effect of temperature on mortality counts is highest among the oldest individuals, Appendix Figure A6 shows that the U-shaped relationship between temperature and life years lost is similar across age groups. This pattern would be unlikely to emerge if the marginal deaths due to temperature were concentrated solely among those with the shortest expected lifespans, since these individuals are given very little weight in the new regressions. Indeed, extrapolating from these estimates, we find that the marginal decedent due to extreme temperatures has an expected remaining lifespan that is quite similar to that of the typical person who dies on more temperate days (details available upon request). This observation supports the conclusion that the mortality effects we measure represent "true" marginal deaths associated with substantial loss of life years and not merely short-term mortality displacement.

---

[5] Available at http://www.ssa.gov/oact/STATS/table4c6.html.

References

Braga, Alfésio Luís Ferreira, Antonella Zanobetti, and Joel Schwartz. "The time course of weather-related deaths." *Epidemiology* 12, no. 6 (2001): 662-667.

Deschenes, Olivier, and Michael Greenstone. "Climate change, mortality, and adaptation: evidence from annual fluctuations in weather in the US." *American Economic Journal: Applied Economics* 3, no. 4 (2011): 152-185.

Deschenes, Olivier, and Enrico Moretti. "Extreme weather events, mortality, and migration." *The Review of Economics and Statistics* 91, no. 4 (2009): 659-681.

Hajat, Shakoor, Ben G. Armstrong, Nelson Gouveia, and Paul Wilkinson. "Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London." *Epidemiology* 16, no. 5 (2005): 613-620.

Huynen, MMTE-Martens, Pim Martens, Dieneke Schram, Matty P. Weijenberg, and Anton E. Kunst. "The impact of heat waves and cold spells on mortality rates in the Dutch population." *Environmental Health Perspectives* 109, no. 5 (2001): 463.

Kyselý, Jan. "Mortality and displaced mortality during heat waves in the Czech Republic." *International Journal of Biometeorology* 49, no. 2 (2004): 91-97.

Minnesota Population Center. National Historical Geographic Information System: Version 11.0 [Database]. Minneapolis: University of Minnesota. 2016. http://doi.org/10.18128/D050.V11.0.

Oudin Åström, Daniel, Forsberg Bertil, and Rocklöv Joacim. "Heat wave impact on morbidity and mortality in the elderly population: a review of recent studies." *Maturitas* 69, no. 2 (2011): 99-105.

Toulemon, Laurent, and Magali Barbieri. "The mortality impact of the August 2003 heat wave in France: investigating the 'harvesting' effect and other long-term

Appendix Figures and Tables

Appendix Figure A1
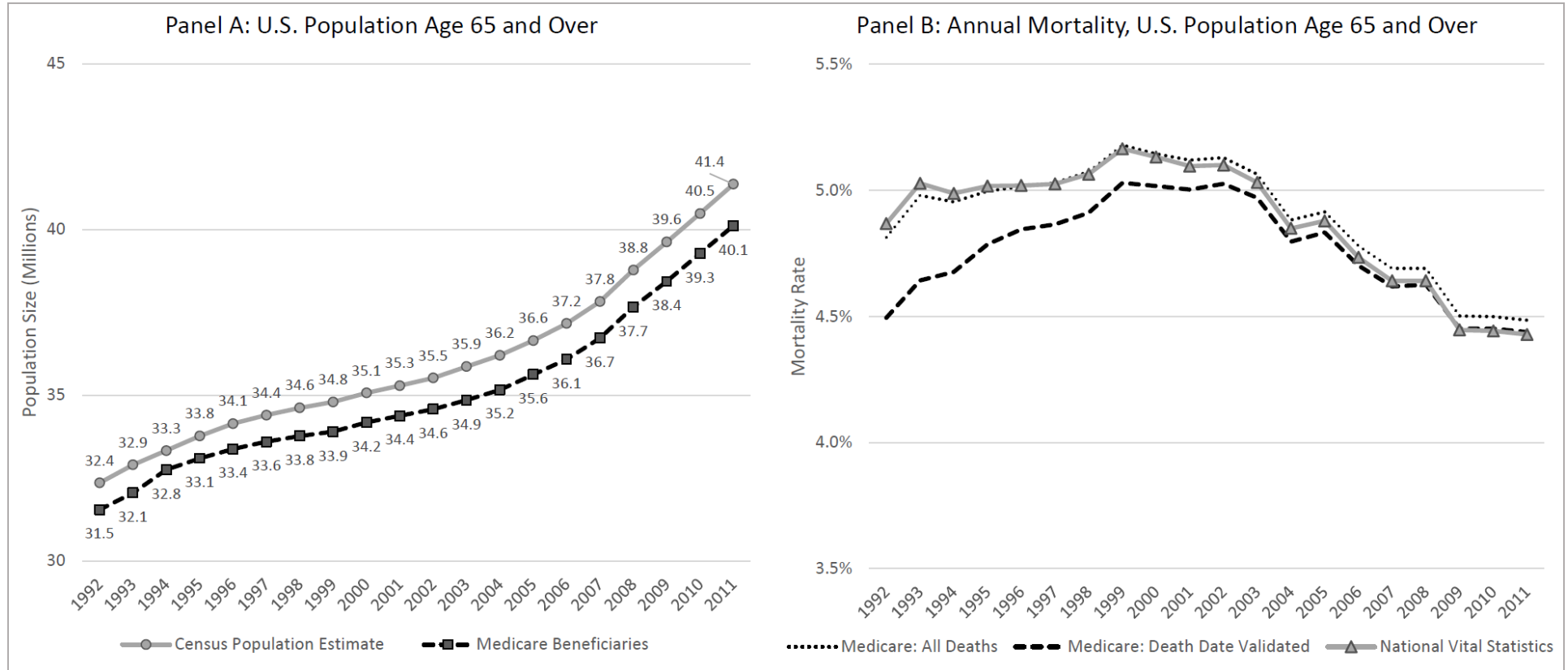
Notes: Figure compares Medicare and Census elderly population and mortality rates over the sample period 1992-2011.

    Left Panel: Census population estimates are derived from U.S. Census Bureau files. Estimates for 1992-2009 are intercensal estimates of the July 1 resident population age 65 and over; estimates for 2010-2011 are postcensal estimates of the July 1 resident population age 65 and over. Medicare beneficiaries for a given calendar year include all individuals age 65-100 in the corresponding annual Medicare enrollment file, limited to those who turned 65 before July 1 of the year and have a ZIP code of residence located in the 50 states or the District of Columbia.

    Right Panel: National vital statistics mortality data come from the Compressed Mortality File (CMF), which is produced by the National Center for Health Statistics and is based on death certificates filed in the 50 states and the District of Columbia. To obtain vital statistics mortality rates, we divide total CMF deaths among the 65 and over population in a given year by the Census population estimates shown in the *Left Panel*. The dashed lines report annual mortality rates based on death dates recorded in the Medicare annual enrollment files. The figure reports both the total mortality rate in the Medicare sample, as well as the mortality rate among the analytical sample used in the paper which excludes individuals who have a validated death that year but do not have a validated death *date* flag.

Appendix Figure A2

Days Exposed to 90+ (F) Average Temperature, per Medicare Beneficiary
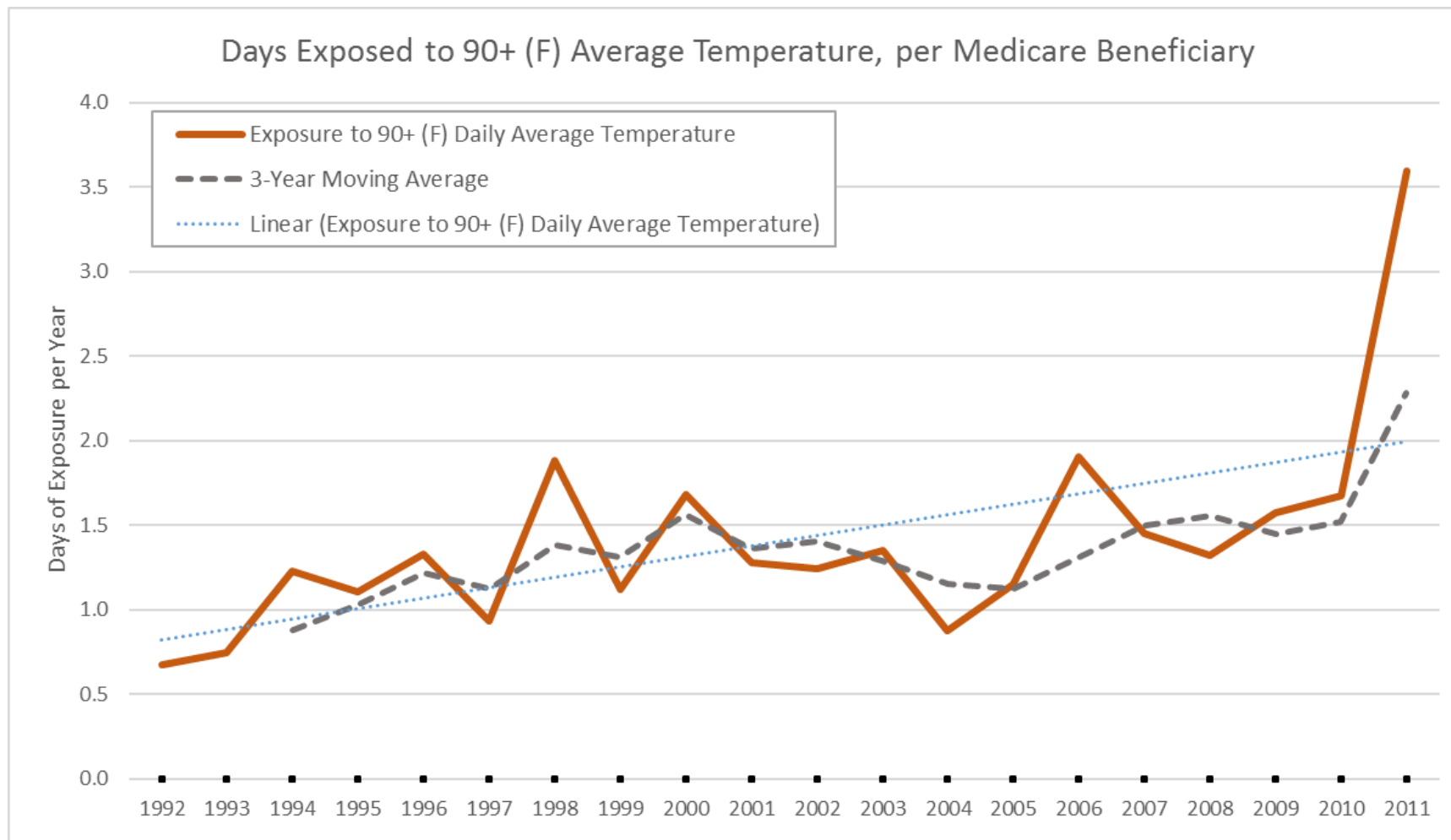
Notes: This graph plots in solid orange the number of days exposed to 90-degree or higher average temperature for a typical Medicare beneficiary, for each year in the sample. Annual exposure is computed by calculating the number of days each beneficiary is exposed to 90+ degree average temperature based on ZIP code of residence and then averaging across beneficiaries. The heavy-dashed gray line and lightly-dashed blue line describe the three-year moving average and a best fit linear trend of this time series, respectively.
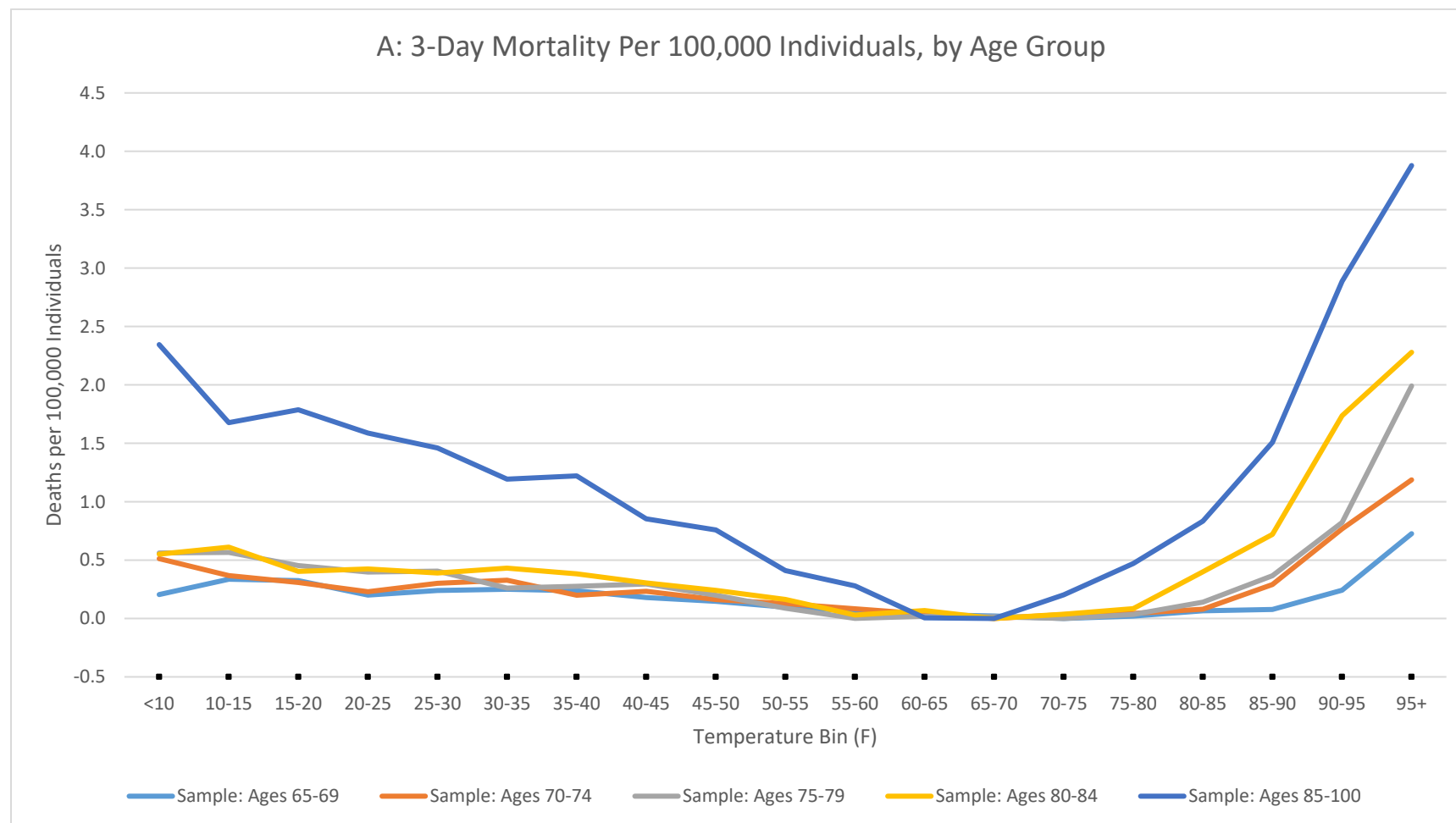
Appendix Figure A3
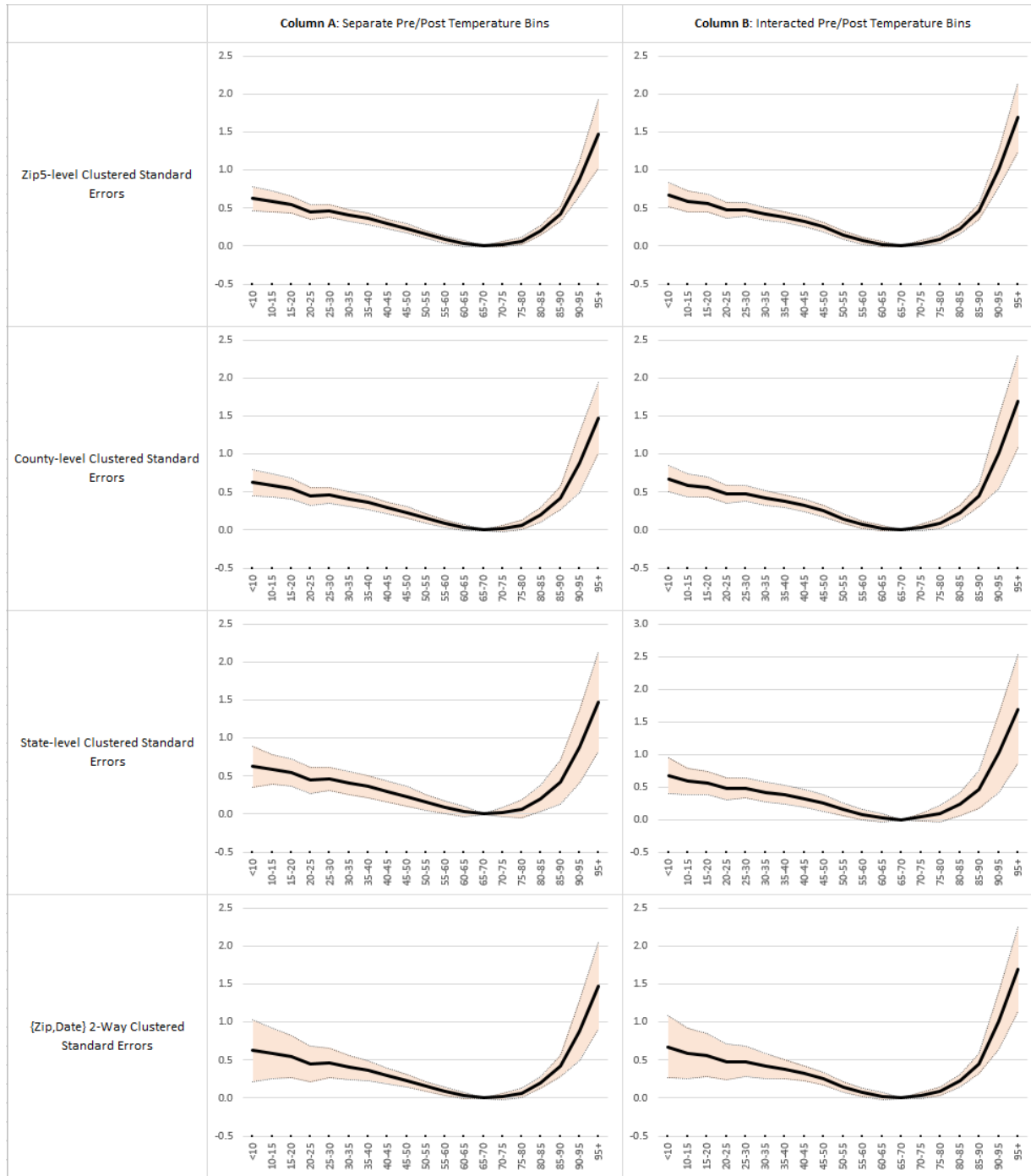
Notes: This figure presents coefficients from the same regression specification as shown in Figure 2 Panel A, but demonstrates the effects clustering standard errors at different levels and excluding interaction terms for the lead and lag temperature controls. Each figure plots estimates from a separate, daily ZIP code level regression of 3-day mortality on daily average temperature bins (see Equation A1). The estimate for each bin reflects excess mortality associated with replacing a day in the reference category (65°F-70°F) with a day in the specified bin. The shaded grey region represents the 95% confidence interval of the estimates. The orange line captures the marginal mortality effect of moving one temperature bin to the right, with the error bars showing the 95% confidence interval. All regressions include state-by-year and ZIP-code-by-day-of-year fixed effects. Regressions are weighted by population and standard errors are clustered at the county level.

Appendix Figure A4
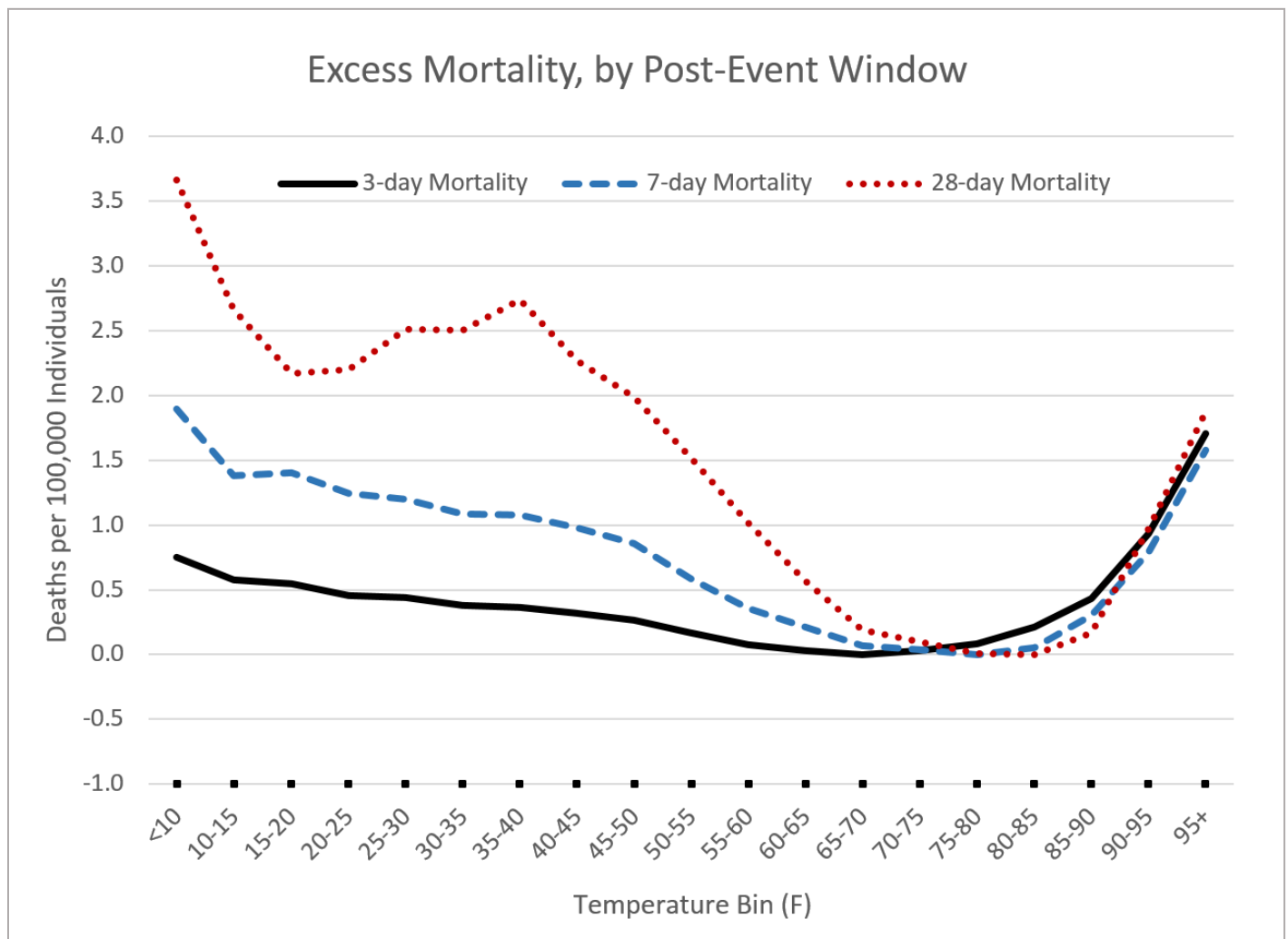
Notes: This figure presents coefficients from separate regressions for each of three outcome variables: 3-day, 7-day, and 28-day mortality. The regression specification is that of Equation (A1), but with temperature leads expanded to included average temperature in the 6 and 27 days following the event day.

Appendix Figure A5

A: 3-Day Mortality Per 100,000 Individuals, by Age Group

Legend: Sample: Ages 65-69 — Sample: Ages 70-74 — Sample: Ages 75-79 — Sample: Ages 80-84 — Sample: Ages 85-100

Notes: This figure plots estimates from a daily ZIP code level regression of 3-day mortality on daily average temperature bins (see Equation A1) separately for individuals in each of 5 age groups: 65-69, 70-74, 75-79, 80-84, and 85-100. The estimate for each bin reflects excess mortality associated with replacing a day in the reference category (65°F-70°F) with a day in the specified bin. All regressions include state-by-year and ZIP-code-by-day-of-year fixed effects and also controls flexibly for temperature leads/lags by fully interacting three sets of 5-degree average temperature bins for the preceding 2 and 6 days and subsequent 2 days. Regressions are weighted by population and standard errors are clustered at the county level.

Appendix Figure A6

C: 3-Day Life-Years Lost Per 100,000 Individuals, by Age Group
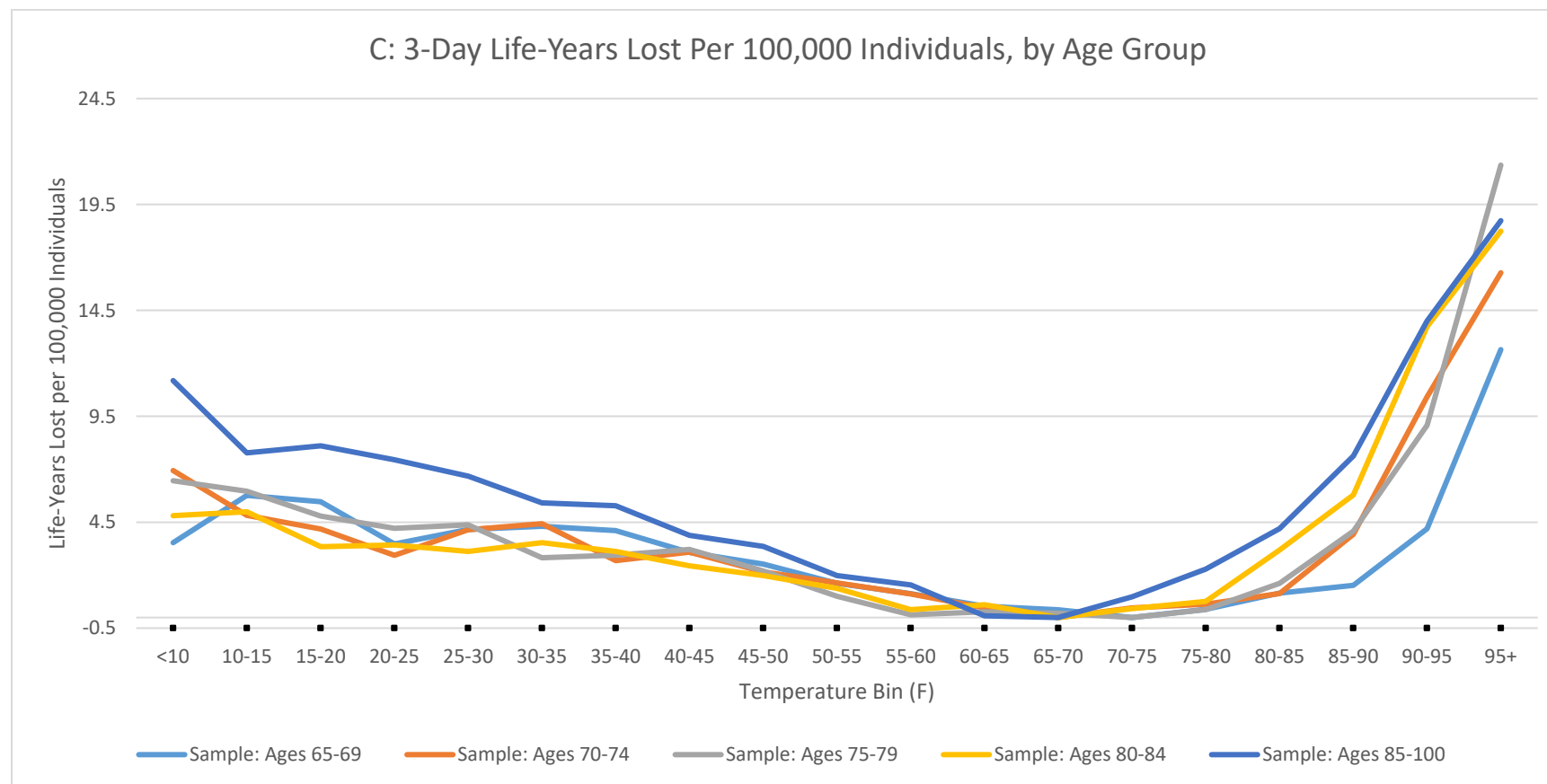
Notes: This figure plots estimates from a daily ZIP code level regression of life years lost (LYL) due to 3-day mortality on daily average temperature bins (see Equation A1) separately for individuals in each of 5 age groups: 65-69, 70-74, 75-79, 80-84, and 85-100. The estimate for each bin reflects additional loss of life years associated with replacing a day in the reference category (65°F-70°F) with a day in the specified bin. LYL is computed by assigning individuals their expected life years remaining based on age and gender using Social Security Administration life tables, and then summing over those alive on any ZIP-code-day to get the total number of remaining life years and then estimating equation (A1). The regression includes state-by-year and ZIP-code-by-day-of-year fixed effects and also controls flexibly for temperature leads/lags by fully interacting three sets of 5-degree average temperature bins for the preceding 2 and 6 days and subsequent 2 days. The regression is weighted by population and standard errors are clustered at the county level.
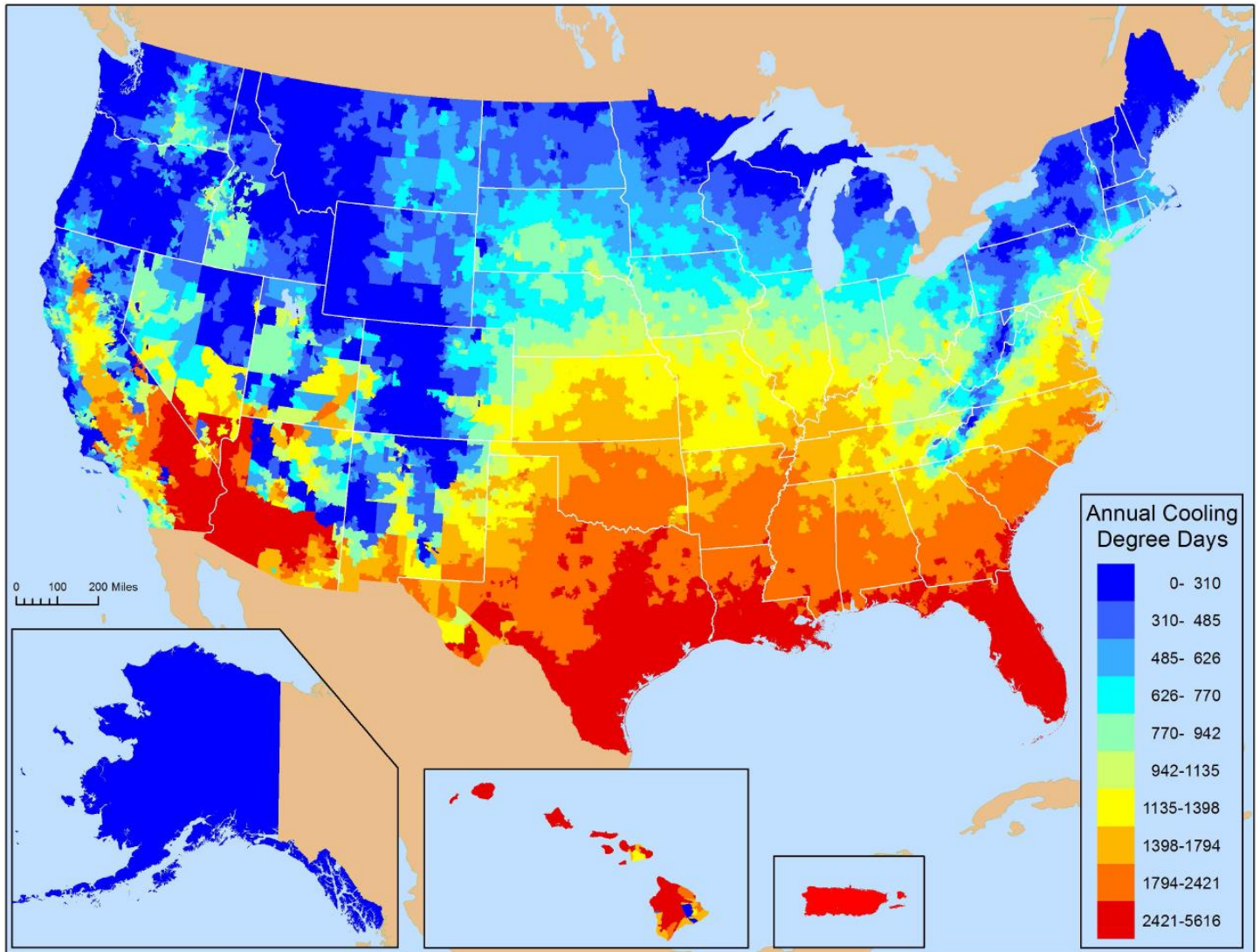
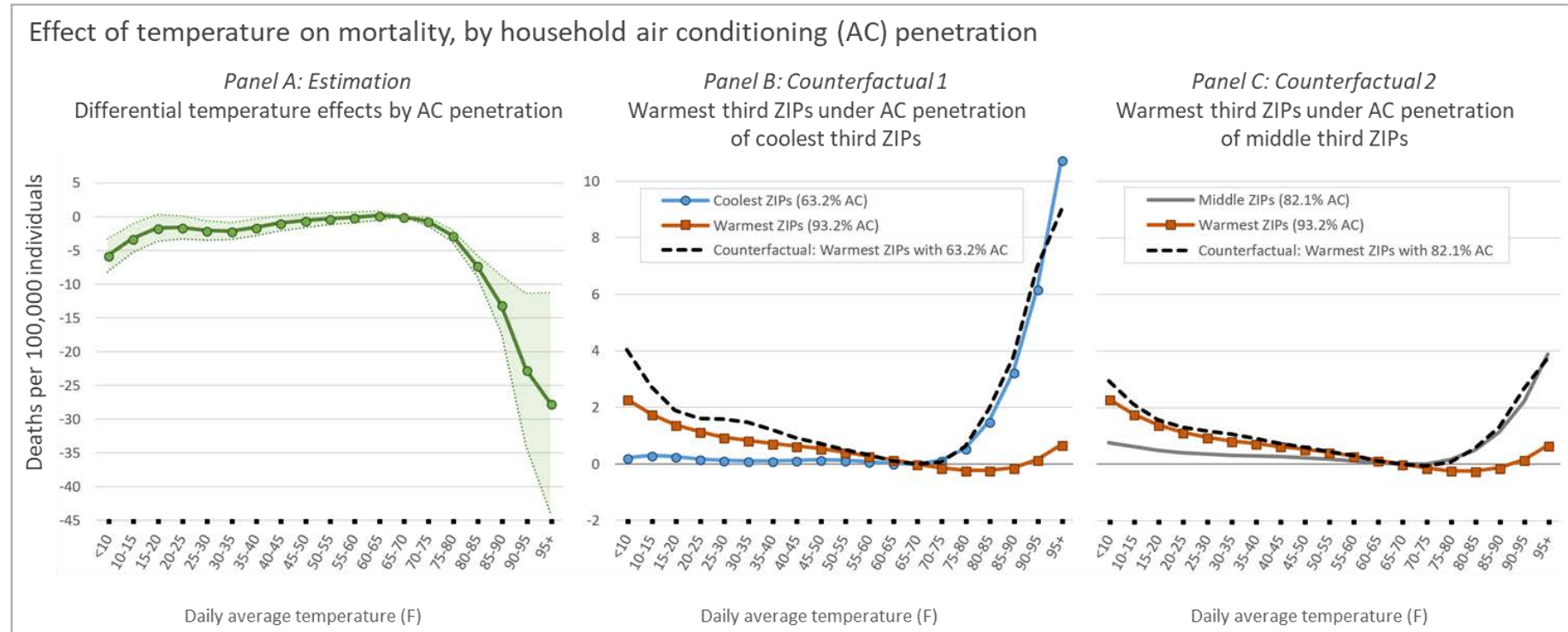**Appendix Figure A7**

Cooling degree days (1981-2010)

Notes: Map shows average cooling degree days over the 3-decade period 1981-2010, by U.S. ZIP code. Data based on NOAA Climate Normals for weather stations, aggregated to the ZIP code level using inverse distance weighting of all monitors within 20 miles of the ZIP code centroid.
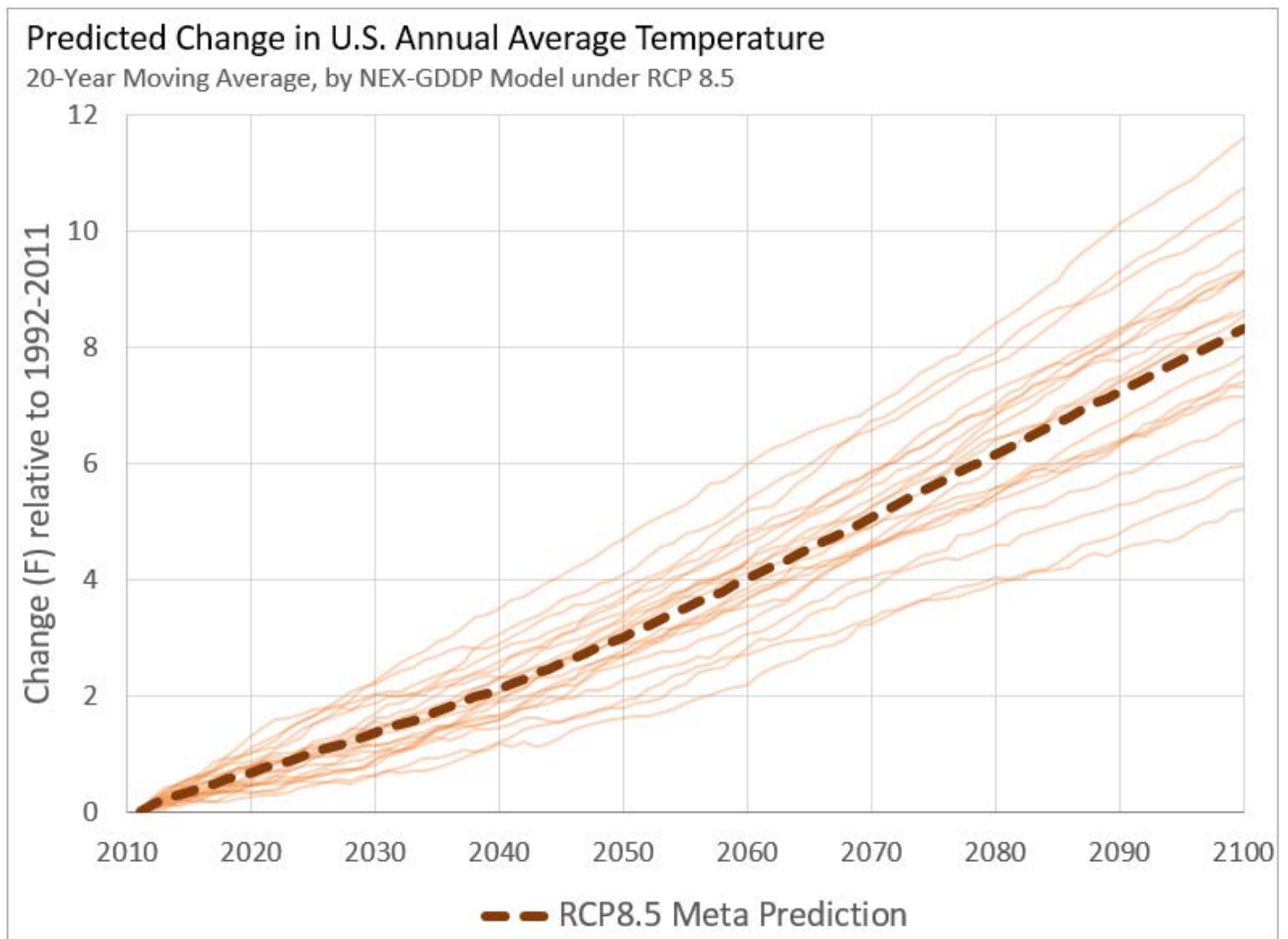
Effect of temperature on mortality, by household air conditioning (AC) penetration

Notes: This figure shows results from replicating the AC heterogeneity analysis presented in Figure 3, except that the AC penetration imputation procedure described in Appendix Section A.3 was modified to exclude any climate variables. *Panel A* reports estimates of how the mortality effect of each temperature bin varies with regional AC penetration. The estimates come from estimating a version of Equation (1) expanded to include daily average temperature bins interacted with ZIP code-level AC penetration as well separate temperature controls for each Census Region. The shaded region reports 95 percent confidence intervals based on standard errors clustered at the county level. *Panel B* and *Panel C* present the implied counterfactual mortality effects of temperature in the warmest third of ZIP codes if exposed to the lower AC penetration rates in the coolest and middle third of ZIP codes, respectively.

Appendix Figure A9

Predicted Change in U.S. Annual Average Temperature
20-Year Moving Average, by NEX-GDDP Model under RCP 8.5

— — RCP8.5 Meta Prediction

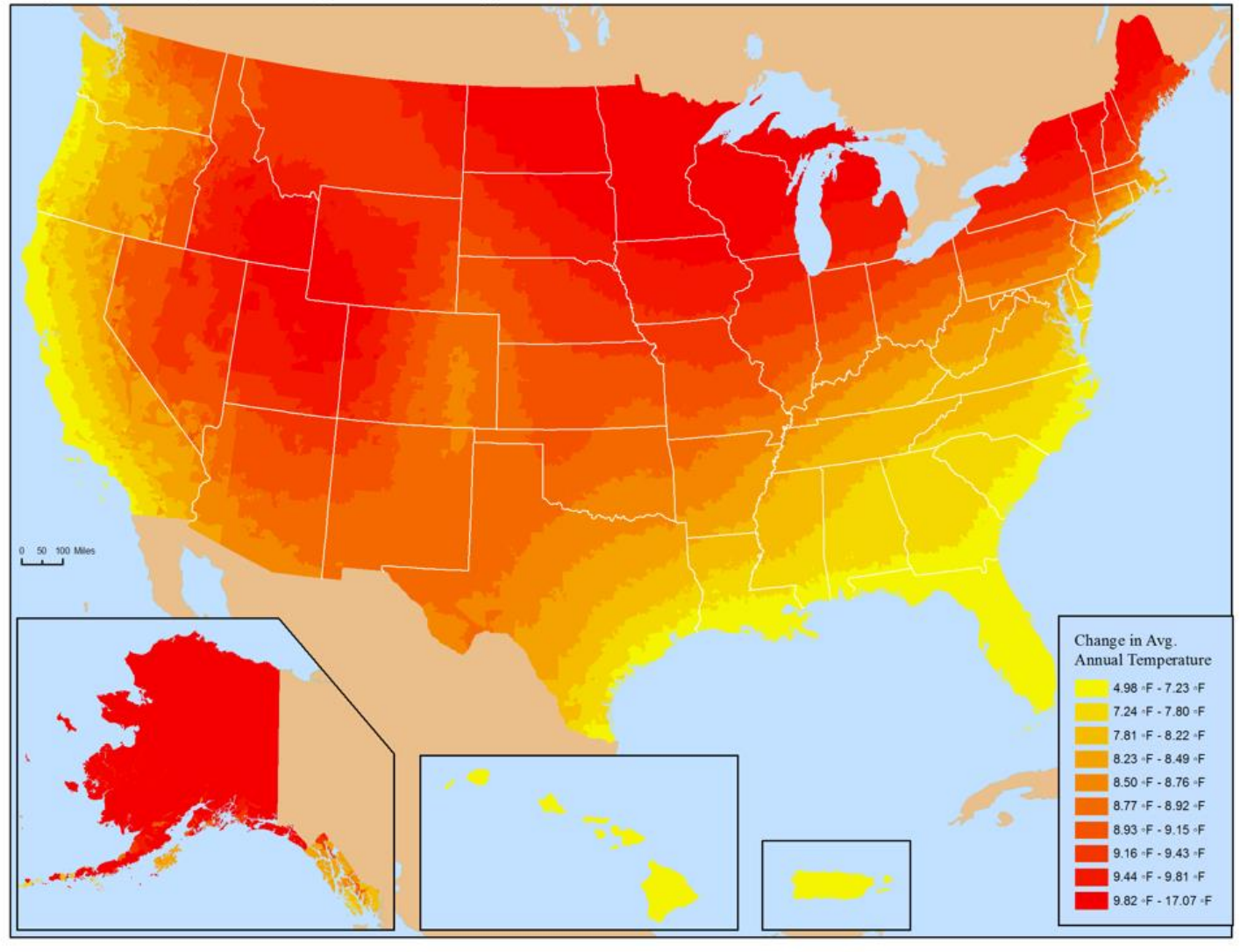*Notes:* This figure plots (thin brown lines) projected U.S. average annual temperature (20-year moving average) for each of the 21 climate models produced under CMIP5 for which daily weather scenarios are available. The thick dashed line plots projected annual average temperature under the "meta" model, which is defined by averaging across the temperature distributions of each of the 21 climate models.

Appendix Figure A10

Projected change in average annual temperature, current versus 2080-2100, under RCP8.5

Change in Avg. Annual Temperature

	4.98 °F - 7.23 °F
	7.24 °F - 7.80 °F
	7.81 °F - 8.22 °F
	8.23 °F - 8.49 °F
	8.50 °F - 8.76 °F
	8.77 °F - 8.92 °F
	8.93 °F - 9.15 °F
	9.16 °F - 9.43 °F
	9.44 °F - 9.81 °F
	9.82 °F - 17.07 °F

Notes: The figure shows, for each U.S. ZIP code, the predicted change in annual average temperature at the end of the century (2080 – 2100) relative to the current period (1992 – 2011) under the RCP 8.5 high emissions scenario. Predicted changes are derived by averaging the predictions of all 21 climate models produced under CMIP5 for which daily weather scenarios are available (see Appendix Figure A9).

Appendix Figure A11

Temperature-mortality relationship for select ZIP codes, based on climate normals
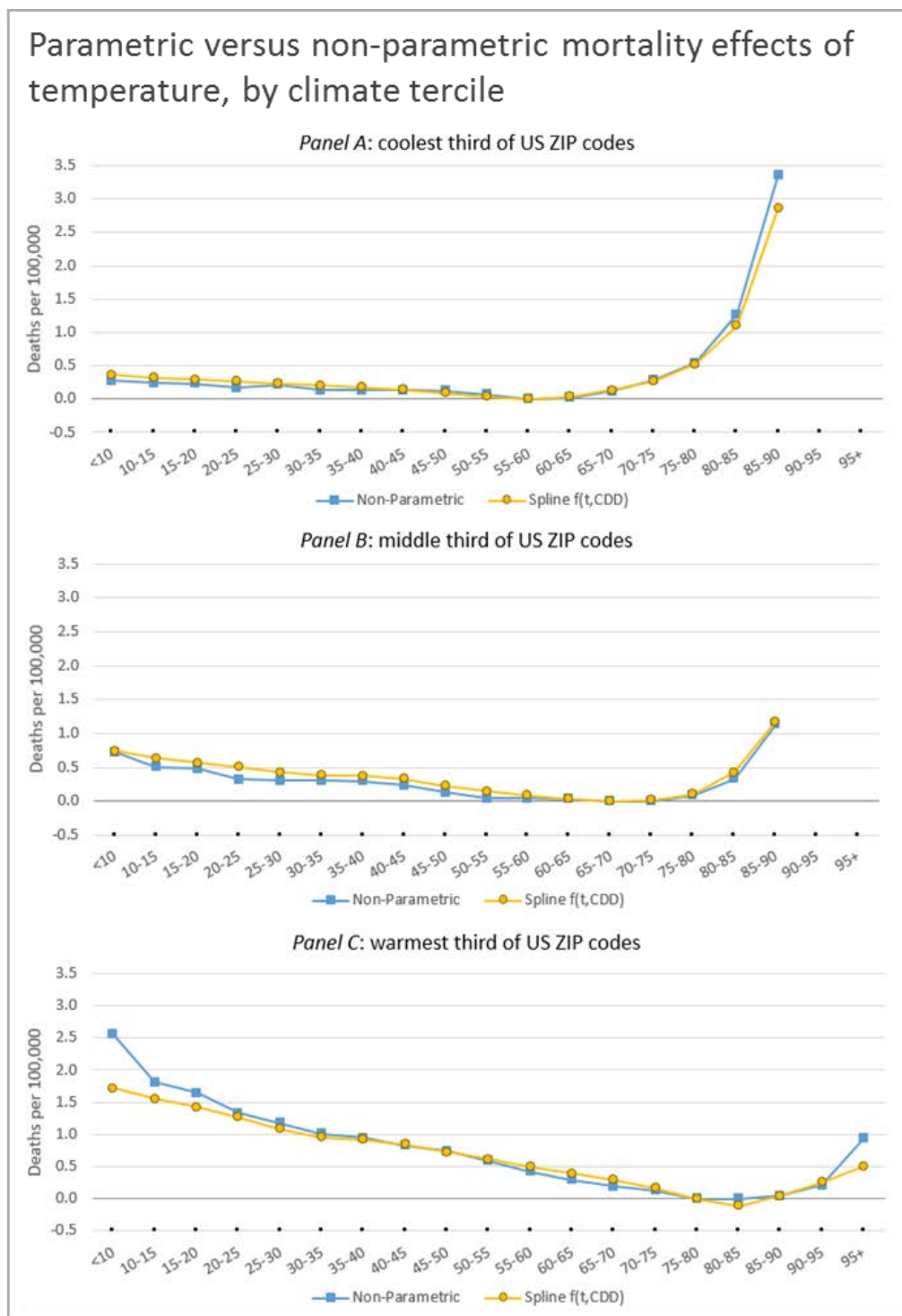
Deaths per 100,000, relative to 65° (F)

Temperature (F)

- - - - Fargo, ND (CDD=564) ——— Mpls/St. Paul MN (CDD=757) ——— Chicago, IL (CDD=992)

- - - Dallas, TX (CDD=2,669) ——— Miami, FL (CDD=4,215)
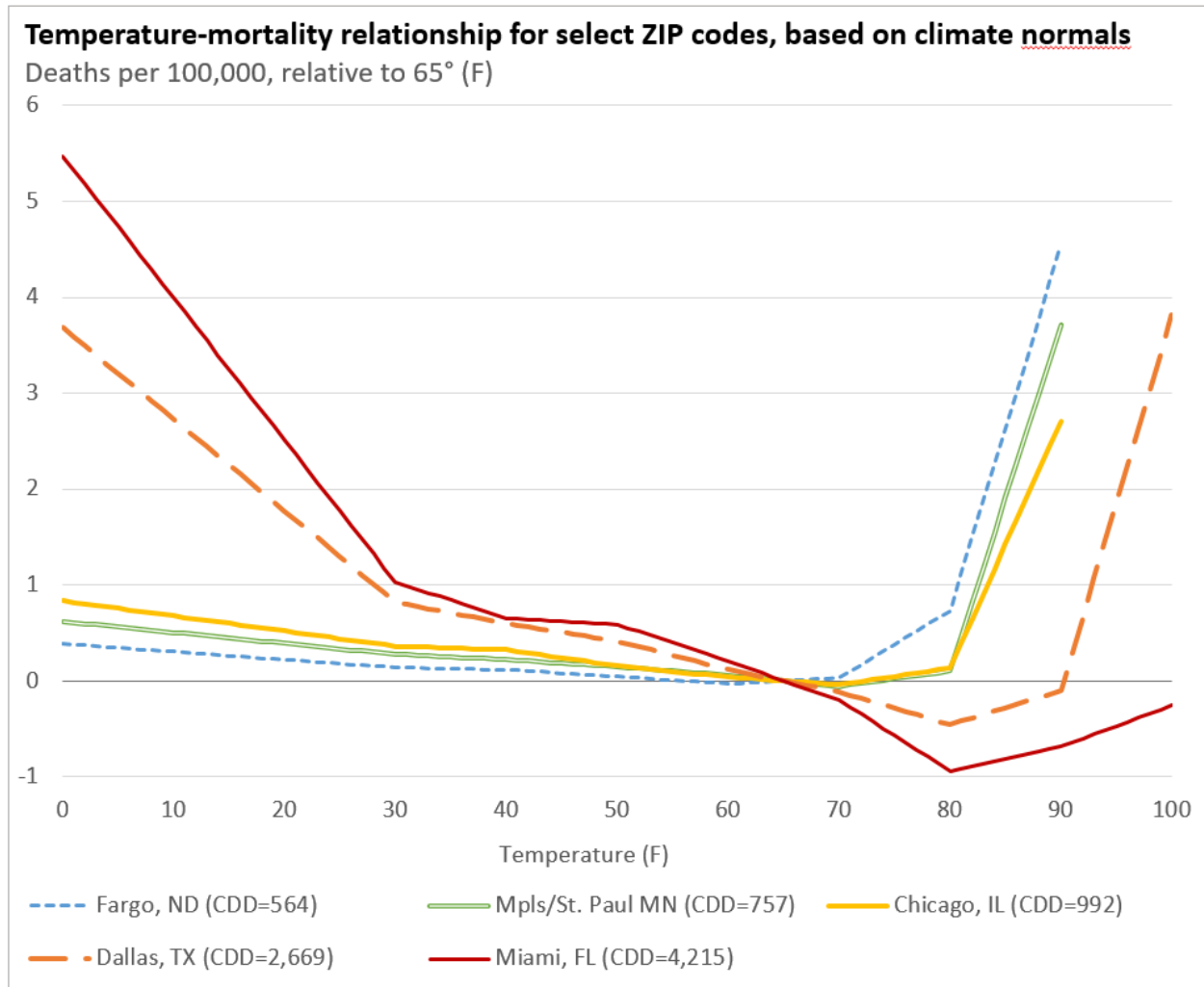
Notes: This figure summarizes the results of estimating Equation (3) by plotting the temperature-mortality relationship for two cold ZIP codes (Fargo, ND, and Minneapolis, MN); one middle ZIP code (Chicago, IL); and two hot ZIP codes (Dallas, TX, and Miami, FL) using the estimated temperature effects $f(t, CDD)$ evaluated at each ZIP code's current CDD normal.

Appendix Figure A12

Parametric versus non-parametric mortality effects of temperature, by climate tercile

Panel A: coolest third of US ZIP codes

Panel B: middle third of US ZIP codes

Panel C: warmest third of US ZIP codes

Notes: This figure summarizes how the estimated semi-parametric function of temperature and climate $f(t, CDD)$ performs relative to the non-parametric estimates of temperature effects by climate tercile reported in Figure 2 Panel B. The estimates plotted in yellow come from re-estimating Equation (1) (used to generate the non-parametric estimates plotted in blue) but with the fitted mortality values $\hat{f}(t, CDD)$ as the outcome and controlling only for temperature and climate tercile indicators.

Appendix Table A1

Year	Number of Monitors	Average, per-monitor number of days with missing temperature observations		Fraction of ZIP codes matched with at least one monitor	Average number of matched monitors per ZIP code
		Tmin	Tmax		
1992	6786	342	341	0.982	3.47
1993	6828	339	337	0.982	3.50
1994	6867	339	337	0.982	3.53
1995	6902	337	336	0.982	3.52
1996	7043	336	334	0.980	3.62
1997	7248	333	331	0.980	3.76
1998	7546	331	328	0.983	4.03
1999	7700	334	331	0.984	4.11
2000	7867	335	333	0.985	4.22
2001	8113	331	329	0.985	4.32
2002	8343	331	328	0.985	4.44
2003	8576	333	330	0.985	4.53
2004	8738	334	332	0.986	4.65
2005	8815	332	330	0.987	4.68
2006	8808	334	332	0.985	4.69
2007	8820	336	334	0.985	4.72
2008	8814	335	333	0.984	4.70
2009	8868	335	334	0.984	4.72
2010	8817	339	338	0.983	4.67
2011	9017	336	335	0.984	4.86

Notes: This table presents summary statistics related to the number of weather monitors available in the NCEI dataset. Number of monitors includes just monitors with at least one observation of Tmax (daily maximum temperature) or Tmin (daily minimum temperature). A monitor is matched to a ZIP code if it is located within 20 miles of the ZIP code centroid.

Appendix Table A2

Age Group	Fraction of Elderly Population	3-day Mortality Per 100,000 Beneficiaries	Average Life Expectancy (Based on Age, Sex)
65-69	26.6%	14.5	17.5
70-74	24.9%	21.6	14.1
75-79	20.3%	33.8	10.9
80-84	14.8%	54.1	8.1
85-100	13.5%	115.4	5.1
Total	100.0%	39.6	12.3

Notes: This table shows the fraction of our sample within each age group and the mortality rate across the age groups for the regressions shown in Figure 4. Statistics are calculated over daily observations at the ZIP code level. Statistics in all columns are weighted by the Medicare age group population in each ZIP code.