



**EI @ Haas WP 239R**

## **Climate Amenities, Climate Change, and American Quality of Life**

David Albouy, Walter Graf, Ryan Kellogg, and Hendrik Wolff  
March 2013

**Revised version published in the  
*Journal of the Association of  
Environmental and Resource Economists*  
3, 205–246, March 2016.**

*Energy Institute at Haas working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to review by any editorial board.*

*© 2013 by David Albouy, Walter Graf, Ryan Kellogg, and Hendrik Wolff. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit is given to the source.*

<http://ei.haas.berkeley.edu>

# Climate Amenities, Climate Change, and American Quality of Life

David Albouy, Walter Graf, Ryan Kellogg, and Hendrik Wolff\*

## Abstract

We present a hedonic framework to estimate U.S. households' preferences over local climates, using detailed weather and 2000 Census data. We find that Americans favor an average daily temperature of 65 degrees Fahrenheit, will pay more on the margin to avoid excess heat than cold, and are not substantially more averse to extremes than to temperatures that are merely uncomfortable. These preferences vary by location due to sorting or adaptation. Changes in climate amenities under business-as-usual predictions imply annual welfare losses of 1 to 3 percent of income by 2100, holding technology and preferences constant. (JEL H49, I39, Q54, R10)

\* Albouy: Department of Economics, University of Michigan, and NBER (email: [albouy@umich.edu](mailto:albouy@umich.edu)); Graf: Department of Agricultural & Resource Economics, University of California, Berkeley (email: [walter.graf@berkeley.edu](mailto:walter.graf@berkeley.edu)); Kellogg: Department of Economics, University of Michigan, and NBER (email: [kellogg@umich.edu](mailto:kellogg@umich.edu)); Hendrik Wolff, Department of Economics, University of Washington (email: [hgwolff@uw.edu](mailto:hgwolff@uw.edu)). Acknowledgments: Albouy acknowledges financial support from NSF grant SES-0922340. We thank Kenneth Chay, Don Fullerton, Philip Haile, Kai-Uwe Kühn, Matthew Turner, and Wolfram Schlenker for particularly helpful criticisms and suggestions. We are also grateful for comments from seminar participants at Boston, Brown, Calgary, Columbia, CU-Boulder, Duke, EPA, Harvard, Houston, Illinois, Maryland, Michigan, Michigan State, Minnesota, NYU, Olin, Oregon, RFF, Rice, Stanford, Texas A&M, TREE, UC Berkeley, UCLA, UNR, Washington, Western Michigan, and Wisconsin, as well as conference attendees at the AEA, Cowles, NASM, NARSC, NBER EEE, RSQE, and WCERE.

The chemistry of the human body makes our health and comfort sensitive to climate. Every day, climate influences human activity, including diet, chores, recreation, and conversation. Households spend considerable amounts on housing, energy, clothing, and travel to protect themselves from extreme climates and to enjoy comfortable moderation. Geographically, climate impacts the desirability of different locations and the quality of life they offer; few seek to live in the freezing tundra or oppressively hot deserts. Given the undeniable influence climate has on economic decisions and welfare, we seek to estimate the dollar value American households place on climate amenities, including sunshine, precipitation, humidity, and especially temperature.

Valuing climate amenities not only helps us to understand how climate affects welfare and where people live, but also helps to inform policy responses to climate change. Global climate change threatens to alter local climates, most obviously by raising temperatures. *A priori*, the welfare impacts of higher temperatures are ambiguous: households may suffer from hotter summers but benefit from milder winters. Ultimately, these impacts depend on where households are located, the changes in climate amenities they experience, and how much they value these changes.

In this paper, we estimate the value of climate amenities in the United States by examining how households' willingness to pay (WTP) to live in different areas varies with climate in the cross-section. Following Rosen (1974, 1979) and Roback (1982), we measure WTP by developing a local quality-of-life (QOL) index based on how much households pay in costs-of-living relative to the incomes they receive. The U.S. is a particularly appropriate setting in which to use this method as it has a large population that is mobile over areas with diverse climates. Globally, the U.S. lies in a temperate zone, with some areas that are quite hot (Arizona) while others are quite cold (Minnesota), and some with extreme seasonality (Missouri) while others are mild year-round (coastal Califor-

nia). This variation allows us to identify preferences over a broad range of habitable climates.

We adopt this hedonic approach as there are no explicit markets for climate amenities, only an implicit market based on household location choices. Our estimates of amenity values primarily reflect impacts of exposure to climate on comfort, activity, and health, including time use (Graff Zivin and Neidell 2012) and mortality risk (Barreca *et al.* 2013, Deschênes and Greenstone 2011). They exclude costs from residential heating, cooling, and insulation. As such, the value of the climate amenities we estimate does not appear in national income accounts, and so neither would the impact of climate change on these amenities. Our study therefore complements work that assesses how climate directly affects national income through agricultural and urban productivity (for a survey of the climate and productivity literature, see Tol 2002 and 2009).

We adopt a cross-sectional estimation strategy in the tradition of Mendelsohn *et al.* (1994) rather than a time-series panel approach for several reasons. First, yearly changes in weather are unlikely to affect households' WTP to live in an area: WTP should depend on long-run climate rather than the outcome of the weather in the most recent year. Second, low frequency changes are not very informative as long-run secular climate changes so far have been slight, particularly relative to changes in technology—especially air conditioning (Barreca *et al.* 2013)—and local economic conditions. Third, Kuminoff and Pope (2012) have shown that temporal changes in the capitalization of amenities do not typically translate to measures of WTP. Finally, households can mitigate potential damages from climate through adaptation—say, by insulating homes, changing wardrobes, or adopting new activities—which cross-sectional methods account for, thereby making our estimates more relevant for assessing the impact of climate change.

An unavoidable drawback of our estimation strategy is that it requires climate amenities to be uncorrelated with the influence unobserved local amenities have on QOL. This untestable assumption is hard to circumvent, as there do not appear to be any viable instrumental variables for climate. Instead, we examine the potential for omitted variable bias by testing the robustness of our estimates to an array of specifications and powerful controls, following the intuition laid out by Altonji *et al.* (2005) and the literature on agricultural yields and farmland values (Schlenker *et al.* 2006, Deschênes and Greenstone 2007, Schlenker and Roberts 2009, Deschênes and Greenstone 2012, and Fisher *et al.* 2012).<sup>1</sup>

Prior hedonic studies investigating U.S. climate preferences have been few and conflicting. Estimates of WTP for incremental warming range from positive (Hoch and Drake 1974, Moore 1998), to approximately zero (Nordhaus 1996), to negative (Cragg and Kahn 1997 and 1999, Kahn 2009, Sinha and Cropper 2013).<sup>2</sup> Our paper makes three contributions to this literature. First, our QOL estimates account for local expenditures beyond housing and net out federal taxes from local wage differences. We show that previous work, by ignoring these factors, over-emphasized differences in wages relative to local costs in measuring QOL, which can lead to noisy and misleading estimates of climate valuations. Second, we characterize climates using the full distribution of realized daily temperatures rather than seasonal or monthly averages, allowing us to explore how households value progressively extreme temperatures. Prior research into climate ameni-

<sup>1</sup> Deschênes and Greenstone (2007, 2012) argue that cross-sectional agricultural damage estimates are not robust to minor changes in the econometric specification, while Schlenker *et al.* (2006), Schlenker and Roberts (2009), and Fisher *et al.* (2012) argue the opposite.

<sup>2</sup> Hedonic studies focusing on countries other than the U.S. include Maddison and Bigano (2003) in Italy, Rehdanz (2006) in Great Britain, Rehdanz and Maddison (2009) in Germany, and Timmins (2007) in Brazil. In addition, alternative non-hedonic approaches have been used to estimate the impact of climate change on amenities. Shapiro and Smith (1981) and Maddison (2003) use a household production function approach, Fritjers and Van Praag (1998) use hypothetical equivalence scales, and Rehdanz and Maddison (2005) link a panel of self-reported happiness across 67 countries to temperature and precipitation.

ties has ignored the importance of flexibly modeling temperature profiles, even though agricultural research (Schlenker *et al.* 2006, Schlenker and Roberts 2009) reports that extreme temperatures, not averages, are especially harmful to crop yields. Third, following methods introduced by Bajari and Benkard (2005), we model unobserved heterogeneity in households' preferences, thereby allowing for sorting based on differences in (dis)taste for cold or heat and for adaptation to local climates. We show that all three innovations are consequential in obtaining preferences for climate.

Our estimates of climate preferences yield four main results. First, we find that Americans most prefer average temperatures (average of daily high and low) near 65 degrees Fahrenheit (°F), agreeing with standard degree-day models that predict little need for heating or cooling at this temperature.<sup>3</sup> Second, on the margin, households pay more to avoid a degree of excess heat than a degree of excess cold. Third, we find that the marginal WTP (MWTP) to avoid extreme temperatures is not substantially greater than the MWTP to avoid moderate heat and cold. Put another way, households will pay more to turn a moderately cold day into a perfect day than to turn a bitterly cold day into a moderately cold day. This finding is consistent with evidence that households reduce their time outdoors as temperatures become uncomfortable, reducing their sensitivity to further temperatures changes (Graff Zivin and Neidell 2012). Fourth, we find that households in the South are particularly averse to cold. This result is consistent with models of both residential sorting and adaptation.

We apply our estimated climate preferences by simulating how climate change may impact welfare by improving or reducing the value of climate amenities. For our climate change forecast, we rely primarily on the business-as-usual A2 scenario used in the Inter-

<sup>3</sup> We define and discuss heating and cooling degree days in section II below.

governmental Panel on Climate Change (IPCC) fourth assessment report (2007), which predicts an 8.3°F increase in U.S. temperatures by 2100. Our simulated welfare effects are predicated on technology and preferences remaining constant and are therefore best interpreted as a benchmark case.<sup>4</sup> We view endogenizing technology and preferences as beyond the scope of this paper’s climate change application, and we leave this issue for future work.

We project that on average, Americans would pay 1 to 3 percent of their annual income to avoid forecasted climate change. While damages are rather severe in the South, we find that most areas in the North also suffer because they lose many pleasant summer days in exchange for only moderately warmer winter days and because Northerners are willing to pay less to reduce cold than are Southerners. Welfare impacts are reduced slightly when we model migration responses.

Below, we first discuss the model we use to derive local QOL estimates from wages and housing costs. Section II describes the current and projected climate data, and section III presents estimates of climate preferences assuming preferences are homogenous. These results provide a baseline for section IV, which allows for heterogeneous preferences. Section V then discusses the implications of our preference estimates for climate change impacts. Section VI considers alternative models of QOL, and section VII concludes.

## **I. Hedonic Estimates of Quality of Life**

The intuition underlying our approach is that households pay higher prices and accept lower wages to live in areas with desirable climate amenities. Below we discuss the he-

<sup>4</sup> The constant technology and preferences assumption is common to most estimates of climate change damages, including the agricultural and health literatures referenced above.

donic framework underlying this intuition, how we calculate wage and cost of living differences across areas, and how we combine these to create a single-index QOL measure for each location. Our approach is rooted in the conceptual framework of Rosen (1974, 1979) and Roback (1982) and adopts important recent contributions to this framework from the hedonic literature. In particular, we follow Albouy (2011) to properly weight wages and housing prices when creating the QOL measure, and we adopt Bajari and Benkard (2005) to allow for preference heterogeneity. These innovations ultimately prove to be consequential in obtaining preferences for climate, as we discuss further below.

#### *A. A model of QOL using local cost of living and wage differentials*

The U.S. economy consists of locations, indexed by  $j$ , which trade with each other and share a population of perfectly mobile, price-taking households, indexed by  $i$ .<sup>5</sup> These households have preferences over two consumption goods: a traded numeraire good  $x$  and a non-traded "home" good  $y$ , with local price  $p_j$  that determines local cost of living. Households earn wage income  $w_j^i$  that is location-dependent and own portfolios of land and capital that pay a combined income of  $R^i$ . Gross household income is  $m_j^i = R^i + w_j^i$ , out of which households pay federal taxes of  $\tau(m_j^i)$ . Federal revenues are rebated lump sum.<sup>6</sup>

Each location is characterized by a  $K$ -dimensional vector of observable amenities  $\mathbf{Z}_j$ , including climate, and a scalar characteristic  $\xi_j$  that is observable to households but not econometricians.<sup>7</sup> We assume a continuum of locations so that the set of available charac-

<sup>5</sup> In section VI below, we relax the perfect mobility assumption by allowing for mobility costs per estimates from Notowidigdo (2012). Doing so has only a minor impact on estimated preferences and welfare impacts.

<sup>6</sup> We also apply adjustments for state taxes and tax-benefits to owner occupied housing, discussed in Albouy (2011), which prove to be minor in practice.

<sup>7</sup> We do not explicitly include an idiosyncratic preference shock  $\varepsilon_{ij}$  in households' utility function. Such a shock would imply that some households are infra-marginal and that our QOL measure should depend on



teristics  $(\mathbf{Z}, \xi)$  is a complete, compact subset of  $\mathbb{R}^{K+1}$ .<sup>8</sup> Each household  $i$  seeks out the location  $j$  that maximizes its utility, given by  $u_j^i = V^i(p_j, w_j^i, \mathbf{Z}_j, \xi_j)$ . This indirect utility function is assumed to be continuous and differentiable in all its arguments, strictly increasing in  $w_j^i$  and  $\xi_j$ , and strictly decreasing in  $p_j$ . Households are permitted to have heterogeneous preferences over  $(\mathbf{Z}, \xi)$ .

On the supply side, we assume that firms face perfectly competitive input and output prices and earn zero profits, offering higher wage levels in more productive locations. We model each household  $i$ 's wage in location  $j$  as  $\varphi^i w_j$ , where  $\varphi^i$  is household-specific skill and  $w_j$  is the local wage level.<sup>9</sup>

Let  $\mathbf{p}(\mathbf{Z}_j, \xi_j)$  and  $\mathbf{w}(\mathbf{Z}_j, \xi_j)$  denote the functions relating wages and prices to local characteristics. These functions are determined in equilibrium by households' demands for local amenities, firms' location decisions, and the availability of land.<sup>10</sup> On the demand side, household utility maximization implies the following first-order condition for each characteristic  $k$ :

$$(1) \quad \frac{1}{m_j^i \lambda^i} \frac{\partial V^i(p_j, w_j^i, \mathbf{Z}_j, \xi_j)}{\partial Z_k} = \frac{p_j y_j^i}{m_j^i} \frac{\partial \ln \mathbf{p}(\mathbf{Z}_j, \xi_j)}{\partial Z_k} - \frac{w_j^i (1 - \tau^i)}{m_j^i} \frac{\partial \ln \mathbf{w}(\mathbf{Z}_j, \xi_j)}{\partial Z_k}$$

---

population in addition to prices and wages. We address this issue by controlling for population density in most of our specifications (we use density rather than population because our data are PUMA-level) and by explicitly adjusting QOL for population changes in section VI. See Albouy (2011) for a more general discussion of addressing idiosyncratic preferences in QOL estimates.

<sup>8</sup> In our empirical implementation, this completeness assumption is an approximation, since we examine a finite, though large, number of U.S. PUMAs. We discuss how this assumption can be relaxed in section IV and appendix 4 via a set identification procedure.

<sup>9</sup> By using a single index of skill, we abstract away from the possibility that some households have a comparative advantage in certain locations, which would be equivalent to allowing for an idiosyncratic preference shock  $\varepsilon_{ij}$  in households' utility function (see footnote 7 above and Albouy (2011)).

<sup>10</sup> The existence and differentiability of  $\mathbf{p}(\mathbf{Z}_j, \xi_j)$  and  $\mathbf{w}(\mathbf{Z}_j, \xi_j)$  hold under mild conditions. Rosen demonstrates the existence and differentiability of the quality of life function  $\hat{Q}(\mathbf{Z}_j, \xi_j)$  under the assumptions on demand given here and perfectly competitive land supply, while Bajari and Benkard (2005) show that a Lipschitz condition on  $V^i(p_j, w_j^i, \mathbf{Z}_j, \xi_j)$  is sufficient even under imperfect competition. Given this result, the separate existence and differentiability of  $\mathbf{p}(\mathbf{Z}_j, \xi_j)$  and  $\mathbf{w}(\mathbf{Z}_j, \xi_j)$  are given by the separate mobility conditions on households and firms and an assumption that local productivity differentials are continuous and differentiable in  $\mathbf{Z}_j$  and  $\xi_j$ .

where  $\lambda^i$  is the marginal utility of income and  $\tau'$  is the average marginal tax rate on labor income. This equation relates household  $i$ 's marginal valuation of characteristic  $k$ , as a fraction of income, to differential changes in the logarithms of the cost of living and wage differentials at  $j$ .

Operationally, we develop a QOL index to indicate the willingness-to-pay of households, averaged by income, from the right-hand side of (1). This measure at  $j$ , denoted  $\hat{Q}_j$ , is a weighted combination of  $\hat{p}_j$  and  $\hat{w}_j$ , the differentials in log housing-costs and wages relative to the U.S. income-weighted average, according to the formula

$$(2) \quad \hat{Q}^j = s_y \hat{p}_j - (1 - \tau') s_w \hat{w}_j = 0.33 \hat{p}_j - 0.50 \hat{w}_j$$

where  $s_y$  denotes the average share of income spent on local goods and  $s_w$  the average share of income from wages. The second equality substitutes in values for these parameters of  $s_y = 0.33$ ,  $s_w = 0.75$ , and  $\tau' = 0.33$ . For additional details, including the incorporation of local non-housing expenditures into  $s_y$ , see Albouy (2011).

Let  $\hat{Q}(\mathbf{Z}_j, \xi_j)$  denote QOL as a function of local characteristics, per (2) and the functions  $\mathbf{p}(\mathbf{Z}_j, \xi_j)$  and  $\mathbf{w}(\mathbf{Z}_j, \xi_j)$ . Then, by condition (1), for any household  $i$  in  $j$ , the marginal willingness to pay (MWTP) for characteristic  $k$  is equal to the derivative of  $\hat{Q}(\mathbf{Z}_j, \xi_j)$  with respect to  $k$ :

$$(3) \quad \frac{1}{m_j^i \lambda^i} \frac{\partial V^i(p_j, w_j, \mathbf{Z}_j, \xi_j)}{\partial Z_k} = \frac{\partial \hat{Q}(\mathbf{Z}_j, \xi_j)}{\partial Z_k}$$

Condition (3) is illustrated in figure 1 in the case of a single characteristic, average summer temperature  $T_s$ . The bold line denotes a hypothetical function  $\hat{Q}(T_s)$  that is decreasing in  $T_s$ , indicating that milder weather is “paid for” through either higher housing prices or lower wages. The slope of  $\hat{Q}(T_s)$  at any given location is the hedonic price for a marginal increase in temperature at that location. This hedonic price is equal to house-

holds' MWTP for  $T_s$  at that location, as shown for locations A and B on the figure. As shown, households at A have a higher MWTP to avoid heat than do households at B, consistent with sorting or adaptation.

### *B. Estimates of wage and housing cost differentials, and QOL*

We follow Albouy (2011) to estimate wage and housing-cost differentials using the 5 percent sample of Census data from the 2000 Integrated Public Use Microdata Series (IPUMS). Geographically, these data are available by Public Use Microdata Areas (PUMAs), which contain populations of 100,000–300,000 and form the main unit of our analysis. We summarize this procedure below; for more details see Albouy (2011).

We calculate wage differentials by PUMA,  $\hat{w}_j$ , using the logarithm of hourly wages for full-time workers aged 25 to 55. These differentials control for observable skill and occupation differences across workers to provide an analogue to the  $\varphi^i$  factors in the model. Specifically, we regress the log wage of worker  $i$  on PUMA-indicators  $\mu_j^w$  and extensive controls  $X_{ij}^w$  (each interacted with gender) for education, experience, race, occupation, and industry, as well as veteran, marital, and immigrant status, in an equation of the form  $\ln w_{ij} = X_{ij}^w \beta^w + \mu_j^w + \varepsilon_{ij}^w$ . The estimates  $\mu_j^w$  are used as the PUMA wage differentials  $\hat{w}_j$  and are interpreted as the causal effect of a PUMA's characteristics on a worker's wage. This interpretation requires that workers do not sort across places of work according to unobserved skills.<sup>11</sup> Finally, we refine the wage differentials so that they reflect wages by place-of-work rather than place-of-residence, netting out differences in commuting costs. The details of this procedure are discussed in Albouy and Lue (2011).

<sup>11</sup> Baum-Snow and Pavan (2012) find that unobserved skill-based sorting contributes little to the urban-rural wage gap, supporting our identification assumption.

To calculate housing costs, we use housing values and gross rents, including utilities. Following previous studies, we convert housing values to imputed rents at a discount rate of 7.85 percent (Peiser and Smith 1985) and add in utility costs to make them comparable to gross rents.<sup>12</sup> We then calculate housing-cost differentials with a regression of rents on flexible controls  $Y_{ij}$  (each interacted with renter status) for size, rooms, acreage, commercial use, kitchen and plumbing facilities, type and age of building, and the number of residents per room. This regression takes the form  $\ln p_{ij} = Y_{ij}\beta^p + \mu_j^p + \varepsilon_{ij}^p$ . The coefficients  $\mu_j^p$  are used as PUMA-level housing cost differentials  $\hat{p}_j$ . Proper identification of housing-cost differences requires that average unobserved housing quality does not vary systematically across locations.

We incorporate energy and insulation costs in our housing-cost measure because doing so allows us to interpret our QOL differentials as solely reflecting the value of nonmarket climate amenities rather than the effect of climate on utility costs. Hence, our QOL differentials will reflect the disamenity of outdoor exposure to climate and the disamenity of adverse indoor temperatures to the extent they are not completely mitigated by insulation and energy use. In addition, the QOL estimates will incorporate any disamenity from spending more time indoors to avoid uncomfortable outdoor temperatures.

Descriptive statistics for QOL are given at the bottom of table 1,<sup>13</sup> and QOL differentials across PUMAs for the year 2000 are mapped in figure 2. These estimates show that households find the amenities in urban areas, coastal locations, and certain mountain areas to be quite desirable. Areas in the middle of the country, where seasons are more extreme, tend to be less desirable, although the variation is considerable. As discussed in

<sup>12</sup> This approach follows the standard practice in the QOL literature from Blomquist *et al.* (1988) to Chen and Rosenthal (2008) and is required by the data because utility costs are included in gross rents.

<sup>13</sup> The mean QOL differential is not exactly zero in table 1 because the table shows unweighted data, while QOL differentials are defined so that the income-weighted mean is zero.

Albouy (2011), our QOL estimates correlate well with non-economic measures of QOL, such as the “livability” rankings in the *Places Rated Almanac* (Savageau 1999). Moreover, the QOL model correctly predicts the relationship between housing costs and wages, controlling for observable amenities.

## II. Data

We estimate our main specifications at the PUMA-level using 2,057 PUMAs covering the contiguous 48 states as of the 2000 census.<sup>14</sup> In this section, we summarize our dataset, covering recent historical climate, climate-change projections, and other variables. Additional details are provided in appendix 2.

### *A. Recent historical climate data*

Our daily temperature data over 1970–1999 originate from Schlenker and Roberts (2009), which provides historic daily temperatures at a 4 kilometer-square resolution. From these data, we create temperature bins with a width of 0.9°F (0.5°C) and calculate the average number of days at each 4 kilometer grid point, over 1970–1999, in which the average temperature—calculated as the mean of the daily high and low—fell within each bin. Within each PUMA, we average the bin distribution at each grid point to yield a PUMA-level dataset.

We obtain monthly precipitation and humidity data from the Parameter elevation Regressions on Independent Slopes Model at the same grid points for 1970–1999, averaging them by month-of-year at the PUMA-level. We obtain sunshine data, measured as the

<sup>14</sup> We have also aggregated our data to the msa-level and run some of our empirical specifications at an msa-level resolution. The point estimates for preferences and climate change welfare impacts are similar to those discussed below, with modestly larger standard errors (see appendix table A1.1, column R13). We believe that the msa-level results are less precise because msas are frequently too large to capture important micro-climates, particularly in densely-populated coastal areas such as San Francisco.

percent of daylight hours for which the sun is not obscured by clouds, by month-of-year from 156 weather stations from the National Climactic Data Center. We interpolate PU-MA-level data on sunshine from the four closest weather stations.

### *B. Projected climate data*

Predicted climate change data (temperatures, precipitation, humidity, and sunshine) are from the third release of the Community Climate System Model. These data were also used in the Intergovernmental Panel on Climate Change Assessment Report 4 (IPCC AR4) released in 2007. We use two business-as-usual scenarios in which no actions to reduce greenhouse gas emissions are taken: the A2 scenario and the A1FI scenario. In both models, data are provided at a resolution of 1.4 degrees longitude (120 km) by 1.4 degrees latitude (155 km), so we interpolate these data to the PUMA level. The A2 scenario predicts average (population-weighted) U.S. warming of 8.3°F from the baseline (1970-1999) to the end of the century (2090-2099), while the A1FI scenario predicts warming of 9.6°F. We focus on the A2 scenario but also provide results for A1FI.

Data for present and projected (A2 model) climate variables are given in the top half of table 1, which summarizes temperature distributions using annual heating degree days (HDD) and cooling degree days (CDD) statistics. HDD is a measure of how cold a location is: it equals the sum, over all days of the year in which it is colder than 65°F, of the difference between 65°F and each day's temperature. CDD, a measure of heat, is defined similarly for temperatures greater than 65°F. HDD and CDD are often used by engineers as predictors of heating and cooling loads, and since Graves (1979) they have often been used as measures of climate amenities.

According to the CCSM model predictions, the effects of climate change will be manifest primarily through changes in temperature, as shown in table 1. The average U.S.

PUMA will see its number of HDDs fall by 36% and its number of CDDs rise by 113% under the A2 scenario. In contrast, changes to precipitation, relative humidity, and sunshine are forecast to be minor on average.<sup>15</sup> The predicted temperature changes vary considerably by geography, as shown in figure 3. While substantial increases in both January and July temperatures are predicted nationwide, the interior South is forecasted to experience a particularly large increase in days for which the average temperature exceeds 90°F.

### *C. Other variables*

Table 1 also presents data on the control variables in our econometric specifications. The geographic controls, used in all estimates, include the minimum distances from each PUMA’s centroid to an ocean and Great Lakes coastline, as well as the average slope of the land, to measure hilliness. Demographic data include measures of population density,<sup>16</sup> educational attainment, age, and racial-ethnic composition. Table 1 also provides statistics on the distribution of population across PUMAs and on the QOL measure.

## **III. Estimation of WTP for Climate under Homogenous Preferences**

### *A. Specification*

We begin by estimating a simple hedonic model in which we assume that climate preferences are homogenous across the U.S. population and that factors (including climate) affecting QOL enter linearly. While this model is highly restrictive, it provides an intuiti-

<sup>15</sup> Some areas are forecast to have appreciable changes, though these cancel out on average. Also note that our study is meant to capture the impact of precipitation on QOL, not water supply.

<sup>16</sup> We use population density rather than population because PUMAs are drawn to have similar populations. Our population density measure is “weighted” in the sense that, within each PUMA, we calculate the population density of each of its census tracts and then take a population-weighted average of these densities. We feel that this weighted density gives a better sense of the population density experienced by households in the PUMA than does a conventional unweighted measure.

tive introduction to our approach that resembles the previous literature and provides a benchmark to compare with subsequent models.

We estimate the impact of marginal changes to climate on QOL using an OLS regression of each PUMA  $j$ 's QOL differential  $\hat{Q}_j$  on vectors of climate variables  $\mathbf{X}_j$  and other local characteristics  $\mathbf{D}_j$ :

$$(4) \quad \hat{Q}_j = \beta \mathbf{X}_j + \gamma \mathbf{D}_j + \xi_j$$

The parameters  $\beta$  and  $\gamma$  represent the WTP of households for an additional unit of each element of  $\mathbf{X}_j$  and  $\mathbf{D}_j$  respectively, measured as a fraction of income. The disturbance term  $\xi_j$  is a vertical location characteristic that is observed by households but not by the econometrician. We face two substantial challenges in estimating (4). First, we must select a functional form for how the climate variables – and in particular the temperature distribution – enter into  $\mathbf{X}_j$ . Our goal is to use a form that is both flexible and capable of providing precise estimates.<sup>17</sup> Second, consistent estimation of  $\beta$  and  $\gamma$  requires that unobserved factors  $\xi_j$  be uncorrelated with  $\mathbf{X}_j$  and  $\mathbf{D}_j$ . In the absence of instrumental variables, orthogonality between  $\xi_j$  and  $\mathbf{X}_j$  is a necessary assumption that cannot be tested. We therefore assess the reliability of our estimates of  $\beta$  by studying their sensitivity to alternative specifications for the control variables  $\mathbf{D}_j$ .

We model the WTP for exposure to an additional day at temperature  $t$ , relative to a day at 65°F, as the function  $f(t)$ .<sup>18</sup> 65°F is a natural normalization point because conventional

<sup>17</sup> We could in principle model rainfall, humidity, and sunshine as flexible distributions, as we do with temperature. We instead simply include these variables as linear regressors, for two reasons. First, these three variables are highly collinear even when entered linearly; allowing for substantial nonlinearities only exacerbates this problem. Second, we wish to focus primarily on temperatures, given that in our climate change application only temperatures change substantially.

<sup>18</sup> This specification, under homogenous preferences, disallows a preference for “seasonality”, which would be manifest as a lower WTP to avoid heat (cold) in a location with severe winters (summers). Preferences for seasonality are, however, captured by our heterogeneous preference specification, which allows the MWTP for any temperature bin to vary with the distribution of realized temperatures. Overall, we find



HDD and CDD calculations treat 65°F as a “bliss point” at which neither heating nor cooling is required. If discomfort from heat and cold, in terms of WTP, follow HDD and CDD, then  $f(t)$  is a two-piece linear function with the kink point at 65°F. One of our objectives is to test whether  $f(t)$  follows this functional form.

Given our data, the most flexible possible model for  $f(t)$  would include a dummy variable for each 0.9°F temperature bin, in which each coefficient would signify a WTP relative to the bin containing 65°F. This model is impractical, however, as we have too little data to provide estimates that are precise enough to be meaningful.<sup>19</sup> Instead, we model  $f(t)$  using cubic splines per (5) below, in which  $S_1(t)$  through  $S_S(t)$  are standard basis functions of a cubic B-spline of degree  $S$ .<sup>20</sup>

$$(5) \quad f(t) = \sum_{s=1}^S \beta_s S_s(t)$$

Defining  $N_{jt}$  as the average number of days per year at location  $j$  for which the average temperature is between  $t$  and  $t + 0.9$ , we flexibly estimate climate preferences by substituting into (4) the product of the spline function in (5) with  $N_{jt}$ , summed over all temperature bins:

---

little evidence of a preference for seasonality, as shown in appendix figure A1.1, which plots the MWTP for 40°F and 80°F as functions of CDD and HDD. These plots would be upward-sloping in the presence of a preference for seasonality.

<sup>19</sup> We have experimented with wider bins and found that bins approximately 10°F wide are necessary to achieve a reasonable degree of precision. While such a binned specification (with a constant WTP assumption over extreme temperatures beyond the present-day distribution) yields similar numerical results to those discussed below (see appendix table A1.1, column R12), we find the continuous splines more appealing as WTP functions because we do not believe that human comfort changes discontinuously every ten degrees.

<sup>20</sup> The support of the U.S. temperature distribution ranges from -39.1°F to 111.2°F, covering 167 0.9°F bins. The spline basis nodes are evenly spaced on the cdf of the population-weighted average temperature distribution (for a cubic spline of degree  $S$ , there are  $S - 2$  nodes, including the nodes at the endpoints). This spacing, rather than even spacing over the unweighted [-39.1°F, 111.2°F] interval, clusters the nodes in the center of the distribution where the data are richest, improving flexibility in this region.

$$\begin{aligned}
(6) \quad \hat{Q}_j &= \sum_t N_{jt} f(t) + \beta_r \text{Rain}_j + \beta_h \text{Humid}_j + \beta_s \text{Sun}_j + \gamma \cdot \mathbf{D}_j + \xi_j \\
&= \sum_{s=1}^S \beta_s \left( \sum_t N_{jt} S_s(t) \right) + \beta_r \text{Rain}_j + \beta_h \text{Humid}_j + \beta_s \text{Sun}_j + \gamma \cdot \mathbf{D}_j + \xi_j
\end{aligned}$$

As a reference case, we include in  $\mathbf{D}_j$  the “full set” of both geographic and demographic controls described above. In all of our regressions, we weight each PUMA (observation) by its population.<sup>21</sup> For inference, we use standard errors that are clustered to allow for arbitrary spatial correlation of residuals across PUMAs within each state (Arellano 1987, Wooldridge 2003).<sup>22</sup>

### *B. Homogenous preference WTP estimates*

Figure 4 presents estimated WTP curves for 5<sup>th</sup> through 10<sup>th</sup> degree splines given our reference case controls. These plots also depict, for reference, the present and future (2090–2099, A2 scenario) U.S. population-weighted average temperature distributions. While the 5<sup>th</sup> degree spline appears too restrictive, and the 10<sup>th</sup> degree too noisy, several regularities emerge from the set of plots. First, the WTP curves consistently have an interior maximum near 65°F. We view this result—which is driven by the data and not “forced” by our QOL variable or functional form—as an important validation of our model and empirical strategy, since it accords with the intuition underlying HDD and CDD that WTP is maximized at 65°F. Second, there are too few days with average temperatures over 90°F to permit precise inference over this range, as evinced by the extremely wide standard error bands. Third, it appears that WTP departs non-linearly

<sup>21</sup> As shown in appendix table A1.1, column R11, this weighting does not materially affect the estimates.

<sup>22</sup> We have also experimented with clustering at the metropolitan statistical area (msa) level (in which PUMAs that are not part of an msa are clustered within each state) and census division level. These approaches lead to estimated standard errors that are only slightly smaller and larger, respectively, than those presented here. When clustering on census divisions (of which there are nine) we use the cluster wild bootstrap to improve small sample performance (Cameron *et al.* 2008).

away from 65°F, undermining the restriction from the HDD/CDD model that  $f(t)$  is linear. In particular, the slopes of the WTP curves tend to be steep on either side of 65°F but then level off at more extreme temperatures. Fourth, the decline in WTP away from 65°F is steeper on the right than on the left, meaning there is a greater WTP to avoid heat than to avoid cold.

The result that WTP declines less steeply over extreme temperatures accords with the intuition that households protect themselves from extremes by taking shelter in climate-controlled environments. Once the temperature is sufficiently uncomfortable that households spend little time outside, further increases in extreme temperature are less important. This result is consistent with research by Graff Zivin and Neidell (2012) that uses time-diary data to show that households spend less time outside in cold or hot weather. It contrasts with the Schlenker and Roberts (2009) result that crop yields are severely reduced by extreme heat, following the logic that humans can take shelter inside while crops cannot.

The result that, on the margin, increases in heat are worse than increases in cold also follows intuition. Individuals can adapt to cold by wearing more clothing. However, options for thermoregulation are more limited in hot conditions.

The imprecision of our estimates at the extremes of the temperature distribution inhibit the ability of spline-estimated WTP functions to inform the welfare effects of climate change. This is true especially at the high end: days exceeding 90°F are rare at present but common in climate change projections. To address this issue, we examine two functional forms that maintain flexibility in the interior of the temperature distribution while projecting WTP at the far extremes.<sup>23</sup> The first is a 7<sup>th</sup> degree cubic spline, with WTP assumed to

<sup>23</sup> We are not the first to confront the issue of conducting inference at temperatures near and beyond the limit of what is realized in present-day data. Prior work in the crop yield and health literatures has assumed that the damage function is constant beyond the point at which inference is no longer feasible (Deschênes

be constant over the extreme 1% of realized temperatures.<sup>24</sup> The second is a four-piece linear spline, with kink points at 45°F, 65°F, and 80°F, allowing for a projection of decreasing WTP over extreme temperatures and permitting more straightforward hypothesis testing of how WTP changes over the temperature distribution.<sup>25</sup>

Panels A and B of figure 5, and columns I and II of table 2, present results for both specifications, controlling for both geography and demographics. In both specifications, WTP declines more steeply as temperatures increase away from 65°F than as they decrease from 65°F. With the linear spline, we reject a null hypothesis that the magnitudes of the slopes on either side of 65°F are equal with a p-value of 0.007. Both splines exhibit flatter slopes over the extremes than over the center of the distribution, particularly on the cold side of 65°F. For the linear spline, the change in slope at the 45°F kink point is statistically significant ( $p = 0.056$ ), though the change in slope at 80°F is not ( $p = 0.534$ ).

Thus, while we have some confidence in the result that extreme cold is not substantially more disamenable than moderate cold, we cannot say the same about extreme versus moderate heat. In fact, the linear spline projection is sufficiently imprecise to the right of 80°F that we also cannot reject the hypothesis that WTP is constant over this range ( $p = 0.179$ ). Overall, it is difficult to make a precise claim about the behavior of WTP over extreme heat. The restricted cubic and linear spline specifications can therefore be viewed

---

and Greenstone 2007 and 2008, Schlenker and Roberts 2009). The first of our two restricted specifications—a 7<sup>th</sup> degree cubic spline with a restriction that MWTP is constant at the extremes—accords with this prior practice.

<sup>24</sup> Specifically, we impose constant WTP over the bottom 0.5% and top 0.5% of the population-weighted temperature distribution. We apply this imposition after estimating the full 7<sup>th</sup> degree spline rather than beforehand, as reversing the order leads to unstable estimates at the cutoff points. We focus on a 7<sup>th</sup> degree spline because it visually appears to strike a balance between flexibility and precision. Using 6<sup>th</sup> or 8<sup>th</sup> degree splines instead does not substantially affect the results.

<sup>25</sup> The choices of 45°F and 80°F tend to yield the best fit to the data (per  $R^2$ ) in most specifications. Alternative choices such as 50°F or 75°F do not substantially impact the preference or welfare estimates.

as two plausible alternatives. The former is conservative in assessing the WTP to avoid extreme heat, whereas the second is more aggressive.

Regarding preferences for other climate variables, the results in columns I and II of table 2 indicate that households have a strong preference for sunshine, a mild preference for precipitation, and no statistically significant taste for humidity. Due to multicollinearity, it is difficult to disentangle preferences for these three climate variables: locations that are very sunny also tend to be dry and non-humid.<sup>26</sup>

The demographic controls may themselves be endogenous, as demographic groups may have different tastes and sort accordingly.<sup>27</sup> That said, we believe that the benefits of controlling for demographics outweigh the costs. First, their inclusion substantially improves the precision of the estimated WTP for climate. Panels C and D in figure 5, and columns III and IV of table 2, present estimates of specifications that do not include the demographic variables. These specifications have notably wider WTP confidence intervals and less than one-half the  $R^2$  of those that include demographics. Second, including demographic variables helps guard against omitted variables bias and, in this case, provides evidence that such bias may not be a substantial concern. Although the demographic variables are powerful correlates of QOL, as evidenced by their large effect on  $R^2$ , they have only a modest effect on the estimated WTPs for the climate variables. Nor, as shown in

<sup>26</sup> Amongst the geographic control variables, we estimate strong tastes for hilliness (average land slope), a taste for proximity to the ocean, but no preference for proximity to a Great Lake. The estimated coefficients on the demographic control variables indicate that QOL increases with population density, educational attainment, average age, and percent Hispanic, while it decreases in percent black.

<sup>27</sup> The coefficient on population density is particularly difficult to interpret, since it captures the amenities that come with population density (e.g., culture and restaurants), unobserved amenities that attract people, and idiosyncratic household-specific valuations of a location. The first two drive the coefficient on population density upward, while the latter drive it downward (because the cost of living must fall (or wages must rise) to attract additional households). Empirically, the amenity aspects of population density appear to dominate, since we estimate a positive coefficient on this variable. Our estimated WTP for climate is robust to excluding population density from the specification. In section VI, we discuss the related issue of adjusting our QOL measure to account for changes in population between 1990 and 2000.

section V below, do they substantially impact our estimates of the welfare impact of climate change. Following the logic of Altonji *et al.* (2005), these results suggest that if unobserved demographics substantially bias our climate preference and welfare estimates, they would have to be more strongly correlated with climate than the “headline” observable demographics included in our regressions.<sup>28</sup> In appendix table A1.1, panel (a), we further probe the robustness of our estimates by enriching our specification in various ways and find that doing so does not substantially affect our results.<sup>29</sup>

Specifications that include state fixed effects (FE) along with demographics are examined in Figure 5 panels E and F, and columns V and VI of table 2.<sup>30</sup> These specifications rely solely on within-state variation in climate for identification. Remarkably, we still find in panel E that WTP is maximized near 65°F despite this substantial change in identifying variation. The inclusion of state FE does result in a loss of precision and an increase in the estimated WTP to avoid cold, though this remains smaller than the estimated WTP to avoid heat (this difference is no longer statistically significant). Overall, the confidence bands are sufficiently wide that the state FE estimates are not statistically significantly different than those without FE.<sup>31</sup>

<sup>28</sup> In other settings for hedonic studies—for example, measurements of the WTP for clean air—it has been shown that estimates are sensitive to the inclusion or exclusion of demographic controls (see, for example, Chay and Greenstone 2005). In such cases, the source of bias is often intuitive. For example, emitters of pollutants are often concentrated in industrial areas that tend to be populated by lower-income households. With regard to climate, however, a source of omitted variable bias from demographics is less obvious *ex-ante* because a wide range of demographic groups and urban vs. rural locations can be found in essentially every climate zone. Our result that adding demographic variables to the specification does not substantially affect the estimated WTP for climate accords with this intuition.

<sup>29</sup> It is true, however, that excluding the geographic variables altogether substantially increases the estimated WTP for mild temperatures, since coastal and mountainous locations tend to be associated with mild weather and provide amenities directly. Based on the results in appendix table A1.1, it appears that a linear function in slope and a quadratic in inverse distance to the coast are sufficient to capture the effects of these variables.

<sup>30</sup> As shown in appendix table A1.1, specification R10, using census division fixed effects rather than state fixed effects does not result in a substantial difference relative to the non-FE estimate.

<sup>31</sup> We test the null hypothesis that the state FE are uncorrelated with the climate variables via block wild bootstrapping, clustering on state (200 repetitions). For each draw, we estimate preferences both with and

The statistical power of the FE estimates comes from large, populous states with a large number of PUMAs and a broad range of climates. Thus, states like California and Texas are more influential than states in the Northeast and Midwest. If residents of these large, warm states have sorted because they are more averse to cold—a possibility consistent with our findings in section IV below—this may explain the difference between the FE and non-FE estimators.

## IV. Estimation of WTP for Climate with Heterogeneous Preferences

### A. Empirical strategy

We now relax the assumption that households share homogenous preferences for climate, allowing them to sort into locations that suit their preferences.<sup>32</sup> Estimation under these more relaxed conditions is based on the framework developed by Bajari and Benkard (2005) and applied by Bajari and Kahn (2005). The intuition of this approach lies in the first order condition of equation (3) and is illustrated in figure 1, which depicts a hedonic equilibrium in which the only characteristic is average summer temperature,  $T_s$ . Given a nonlinear hedonic price function  $\hat{Q}(\mathbf{Z}_j, \xi_j)$ , the MWTP of households located at  $j$  for a characteristic  $k$  (e.g., a day in a given temperature bin) is simply given by  $\partial \hat{Q}(\mathbf{Z}_j, \xi_j) / \partial Z_k$ . Thus, flexible estimation of  $\hat{Q}(\mathbf{Z}_j, \xi_j)$  allows us to recover the distribution of MWTP for each characteristic  $k$  across the population of households. In figure 1,

---

without state FE. We then use all draws to obtain a standard deviation of the difference between the estimators. This procedure, unlike a traditional Hausman test, does not require that the non-FE estimator be efficient (see Cameron *et al.* 2005, p. 718). We find that we cannot reject equality of the WTP for 40°F across the estimators (the p-values are 0.708 for the cubic spline and 0.501 for the linear spline). Moreover, we cannot reject equality of the welfare impact of climate change across the two estimators (the p-values are 0.208 for the cubic spline and 0.203 for the linear spline).

<sup>32</sup> While this section emphasizes the sorting explanation for preference heterogeneity, we reiterate here that heterogeneity could also reflect households' adaptation to their local climate.

for example, the depicted equilibrium is consistent with positive sorting in which households with a high MWTP to avoid heat settle in areas with low summer temperatures.

To flexibly estimate  $\hat{Q}(\mathbf{Z}_j, \xi_j)$ , we use local linear regression per Fan and Gijbels (1996). We suppose that, local to location  $j^*$ ,  $\hat{Q}(\mathbf{Z}_j, \xi_j)$  satisfies (7) below:

$$(7) \quad \hat{Q}_j = \sum_k \beta_k^{j^*} Z_{jk} + \xi_j$$

In (7), the implicit prices  $\beta$  are superscripted by  $j^*$ , as we estimate a distinct set of prices at each location. We obtain the  $\beta^{j^*}$  at each location via weighted least squares per (8) and (9):

$$(8) \quad \beta^{j^*} = \arg \min_{\beta} (\mathbf{Q} - \mathbf{Z}\beta)' \mathbf{W} (\mathbf{Q} - \mathbf{Z}\beta)$$

$$(9) \quad \mathbf{Q} = [\mathbf{Q}_j]; \quad \mathbf{Z} = [\mathbf{Z}_j]; \quad \mathbf{W} = \text{diag}[K_h(\mathbf{Z}_j - \mathbf{Z}_{j^*})]$$

where  $\mathbf{W}$  is a matrix of kernel weights defined so that locations that are similar to  $j^*$  in characteristics receive the most weight in the regression.<sup>33</sup> This approach allows households in relatively hot or cold locations to have different MWTPs to avoid departures from mild climates. To calculate  $\mathbf{W}$ , we use a normal kernel function with a bandwidth  $h$  of 2, per (10) and (11) below.<sup>34</sup>  $\hat{\sigma}_k$  denotes the standard deviation of characteristic  $k$  across the sample:

<sup>33</sup> In principle,  $\mathbf{W}$  can include every characteristic in the model, so that the MWTP for a given temperature can vary in the cross-section not only with the climate variables but also with the controls. In practice, however, we include only the temperature spline basis functions in  $\mathbf{W}$ , since including the controls requires a very large bandwidth in order to avoid collinearity in the weighted covariate matrix. Moreover, when we implemented this procedure for the 7<sup>th</sup> degree cubic spline, we initially obtained very noisy estimates of temperature preferences at extreme locations (such as northwest Minnesota) that have few neighbors in climate space. Basing  $\mathbf{W}$  on only a 5<sup>th</sup> degree cubic spline solved this problem.

<sup>34</sup> The use of a normal kernel follows Bajari and Benkard (2005) and Bajari and Kahn (2005). We also follow these papers in choosing a bandwidth that yields visually appealing preference estimates, in the sense that estimated MWTP varies smoothly over characteristics space (as in figure 6). In contrast, using leave-one-out cross-validation we find an “optimal” bandwidth of 0.22 in the cubic spline preference specification. This small bandwidth leads to very noisy and imprecise preference estimates as well as an



$$(10) \quad K(\mathbf{Z}) = \prod_k N(Z_k / \hat{\sigma}_k)$$

$$(11) \quad K_h(\mathbf{Z}) = K(\mathbf{Z} / h) / h$$

This method yields point identification of MWTP for climate at each location so long as the choice set is continuous in characteristics. In our setting, however, there is a discrete number of 2,057 PUMAs, meaning that households' preferences are only set identified. That is, there is a range of MWTPs such that each household would prefer to stay in its PUMA rather than move. We examine set identification in appendix 4 using a Bayesian method per Bajari and Benkard (2005) and find that, so long as we allow preference heterogeneity over only the temperature variables, the identified sets are small enough to closely approximate point identification. Thus, our primary results presented below follow Bajari and Kahn (2005) in treating MWTP as point identified but restrict preferences to be heterogeneous over only the temperature variables, not the controls.<sup>35</sup>

Finally, we emphasize that we only identify the MWTP of each household for each temperature at its chosen climate and cannot identify its WTP for a large change in climate, over which MWTP may vary. In figure 1, for instance, the MWTPs at locations A and B are identified, but the shapes of the indifference curves are not. Estimating the welfare impacts from non-marginal changes in climate therefore requires assuming a functional form for households' utility. We focus on welfare estimates that assume the indifference curves are linear, with a slope equal to the estimated MWTP. This approach is transparent, permits an apples-to-apples comparison to the homogenous preference esti-

---

extremely large estimated amenity loss from climate change (19.9% of income under the A2 scenario). In general, the estimated amenity loss increases as the bandwidth becomes smaller, so that our bandwidth choice of 2 is conservative.

<sup>35</sup> Prior to implementing the local linear regression, we enforce preference homogeneity over the non-temperature variables by stripping their effects from  $\hat{Q}$  using the OLS estimates from section III (this approach follows an example given in Bajari and Benkard 2005). Thus, the only characteristics involved in the local linear regression are the temperature spline basis functions.

mates, and is conservative: allowing for concavity would unambiguously increase the estimated amenity losses, as the cost of each additional hot day would rise. In appendix 3, we investigate non-linear indifference curves based on the small amount of concavity necessary to satisfy households' second order condition for utility maximization (this calculation is possible because the estimated  $\hat{Q}(Z_j, \xi_j)$  function is slightly concave in some dimensions). We find that allowing for concavity increases the estimated losses from climate change by 0.6 percentage points relative to the results given in section V below.

### *B. Heterogeneous preference MWTP estimates*

Estimated MWTPs for temperature by PUMA are graphed in figure 6: panels A and C from the cubic spline, panels B and D from the linear spline, all controlling for geography and demography, but not state FE. Panels A and B present the distribution of MWTP, scaled up by 365, for an additional day at 40°F relative to a day at 65°F, as a function of HDD. Each of the plotted points represents the preferences of a specific PUMA in the data. The overall upward slope is consistent with households sorting based on their aversion to cold: those with the most negative MWTP for cold weather tend to be located in areas with the fewest HDDs. Panels C and D examine the MWTP for an additional day at 80°F as a function of CDD. These plots provide less conclusive evidence on sorting to avoid heat: while the linear spline estimates exhibit a slightly positive slope, the cubic spline estimates exhibit a weak downward slope. This latter result is consistent with residents of hot areas (up to 3,000 CDD) being more heat-averse on the margin, suggesting mild concavity in households' utility functions rather than sorting. This result further suggests that households have only a limited ability to adapt to heat, relative to their ability to adapt to cold.

An alternative depiction of heterogeneous climate preferences is given in appendix figure A1.2. This figure plots MWTP curves under the cubic spline model at a few select cities. Beyond showing that the magnitude of households' MWTP to avoid cold is negatively correlated with the amount of cold they face, this figure also shows that the estimated maximum of the MWTP function occurs at slightly higher temperatures in warmer locations, again consistent with sorting based on aversion to cold.

The average MWTP to avoid 40°F and 80°F for the cubic and linear spline models plotted in figure 6 are given in columns I and II of table 3. The average MWTP to avoid heat exceeds that to avoid cold, consistent with the homogenous preference model (columns I and II of table 2). Columns III through VI of table 3 proceed through the same control variable specifications that were used in table 2, with similar results.

## **V. Estimated Impacts of Climate Change on Climate Amenities**

In this section, we use climate preference estimates to calculate how forecasted climate change may impact local amenity values. These results maintain technology and preferences at their present levels and are best interpreted as welfare impacts that would occur if changes forecasted for the end of the century were to occur now, allowing for spatial adaptation (e.g., households in the North adopting southern behaviors and levels of air conditioning penetration) to occur instantly.

We begin in section V.A with welfare calculations that assume homogenous preferences and then compare them in section V.B with those assuming preference heterogeneity. In both cases, we focus on projections from the A2 scenario for 2090–2100 and estimate welfare impacts that do not allow for migration. We then examine migration in section V.C.

### *A. Impacts under homogenous preferences*

Table 4 displays estimated changes in welfare using the homogenous preference estimates. Our baseline model I, using the restricted cubic spline under A2, yields an average welfare loss equivalent to 2.24% of income per year by 2090-2100 (estimates under A1FI are given in appendix table A1.2). This net loss is driven by the result that the WTP to avoid heat exceeds the WTP to avoid cold, so that the welfare loss from hotter summers exceeds the gain from warmer winters.<sup>36</sup> These temperature-driven welfare losses are slightly offset by increases in amenities from the small forecast changes to precipitation, humidity, and sunshine.

The geographic distribution of welfare changes in this specification is shown in panel A of figure 7. The vast majority of the U.S. population will experience an amenity loss: the benefits of warmer winters outweigh the cost of hotter summers for only the 3.9% of the population that lives in the Pacific Northwest, northern Michigan, and the extreme Northeast. The amenity losses are largest along the southern Atlantic, Gulf Coast, and southern Pacific coastlines; in the upper Midwest; and in the Great Basin region. The interior South area of Oklahoma, Arkansas, north Texas, and north Louisiana is not projected to experience particularly severe amenity losses despite the fact that this region is forecasted to gain a large number of extremely hot summer days (recall figure 3). This result occurs because summer temperatures in these areas are already sufficiently hot that residents are on the flat portion of the WTP curve. Thus, contrary to conventional wisdom, it may be that Americans in relatively temperate climates—for whom summers are usually warm

<sup>36</sup> In a restricted model in which the WTP function is forced to be symmetric around 65°F, the welfare impact is estimated to only be -0.04% of income. We estimate this symmetric specification using the five-piece linear spline, since it is not possible to impose symmetry around 65°F using the cubic spline. We force the slopes of the segments below 45°F and above 80°F to equal in magnitude and we force the slopes of the segments between 45°F and 65°F and between 65°F and 80°F to be equal in magnitude.

but not severe—will suffer more from changes in climate amenities than southern households.

While our estimates based on the restricted cubic spline model suggest the welfare distribution depicted in figure 7, panel A, we cannot rule out an alternate distribution in which amenity losses are greatest in the hottest regions. Such a distribution is generated by our linear spline specification (corresponding to column II of tables 2 and 4), for which the estimated WTP to avoid an additional day beyond 80°F increases linearly with temperature. Panel B of figure 7 displays the estimated distribution of amenity losses under this specification; these losses primarily occur in the interior South and the desert Southwest. Table 4, column II shows that the average welfare loss is larger in magnitude than for the restricted cubic spline of column I but is estimated less precisely. This lack of precision reflects the fact that the present climate offers too few extremely hot days to allow a precise claim about the behavior of WTP over extreme heat.

Columns III and IV of table 4 present amenity losses that use the corresponding estimates from columns III and IV of table 2. Removing the demographic controls reduces the precision of the welfare impacts but does not substantially affect the point estimates. Appendix table A1.1 presents estimates of amenity losses from several other alternative specifications with varying control structures, finding generally similar results. Finally, columns V and VI use estimates that include state FE, which increase the disamenity from cold. These specifications yield losses of roughly half the magnitude of those from columns I and II, although they are not statistically significantly different (the Hausman test p-values are 0.208 and 0.203, respectively).

### *B. Impacts under heterogeneous preferences*

Table 5 presents preference and climate change welfare impact estimates under the heterogeneous preference models, with columns I through VI corresponding to the same specifications as in tables 2 through 4. The baseline restricted cubic spline model presented in column I yields an estimated overall amenity loss of 2.83% of income, larger in magnitude than the 2.24% from the homogenous preferences model. This difference is primarily driven by the estimated sorting pattern in which households in the North are the least averse to cold. This sorting implies that the forecasted reduction in cold days is spatially negatively correlated with the MWTP to avoid those cold days, reducing the benefit of warmer winters.

The distribution of amenity impacts under specification I is given in panel A of figure 8. Relative to the estimates that assumed preference heterogeneity (figure 7, panel A), the estimated amenity losses increase primarily in the North. Overall, preference heterogeneity reinforces the possibility that the amenity impact of climate change may be felt more in the North than the South.

The remaining columns of table 5 present estimates from alternative empirical specifications. As with the homogenous case, using the linear spline in column II yields an amenity loss that is larger and less precise than that of column I. The distributional impact obtained under this specification is shown in panel B of figure 8; as with panel B in figure 7, welfare losses tend to be greatest in the South. Omitting the demographic controls in columns III and IV does not dramatically change the point estimates but does decrease precision. Finally, including state FE in columns V and VI decreases the magnitude of the welfare impacts, as was the case with homogenous preferences.

### *C. Migration*

The welfare impacts we have estimated thus far do not account for household migration in response to climate change. If climate change is only marginal, then these welfare effects provide accurate first-order approximations, robust to migration via the envelope theorem. However, if climate changes are sufficiently large, the second-order effect of migration may materially mitigate welfare damages.

We assess the importance of migration using a parsimonious model based on homogeneous preferences.<sup>37</sup> Intuitively, households will leave areas with worsening climates and migrate towards areas with improving climates. We model the case in which mobility responses are proportional to changes in QOL,  $\Delta\hat{Q}_j$ , and households remain within the United States. Thus, changes in the logarithm of population density of location  $j$ ,  $\Delta\hat{N}_j$ , depend on changes in QOL at  $j$  relative to the average QOL change in the United States,  $\Delta\hat{Q}_{avg}$ , per the formula

$$(12) \quad \Delta\hat{N}_j = \varepsilon_{N,Q} (\Delta\hat{Q}_j - \Delta\hat{Q}_{avg})$$

where  $\varepsilon_{N,Q}$  is the long-run elasticity of population with respect to changes in quality of life. Although direct empirical evidence on this elasticity is scant, a calibration from Albouy and Stuart (2012) suggests a value of 8 over the very long run.<sup>38</sup> In other words, a one percent increase in QOL will lead to an eight percent increase in population density, once housing and migration fully adjust.<sup>39</sup>

<sup>37</sup> Modeling migration with preference heterogeneity is beyond the scope of this paper.

<sup>38</sup> This elasticity incorporates five different responses to a change in QOL: i) housing producer substitution away from land; i.e. housing supply elasticity ii) traded-good producers' substitution away from land, creating diminishing returns from fixed factors, iii) consumers' substitution away from housing; i.e. crowding into the existing housing stock, iv) willingness of households to consume less overall (net of housing) in high QOL areas; and v) endogenous agglomeration economies and crowding effects from changes in population density.

<sup>39</sup> Saiz (2010) has estimated local housing supply elasticities for 95 metropolitan areas in the U.S. The median estimated elasticity is 1.61 and the maximum is 5.45. In addition, in a medium run climate change

Allowing for adaptive migration per equation (12) we estimate that for our baseline specification I, the average absolute value of the population change across all PUMAs is 8.3%. This migration only modestly changes our estimated aggregate welfare impact from -2.24% to -2.05%. Adaptive migration only slightly mitigates welfare losses because only a few areas actually improve, and migration to these areas becomes constrained by crowding.<sup>40</sup>

## VI. Alternative measures of QOL

In section I above, we emphasized the importance of correctly weighting wage and housing-cost and wage differentials when calculating our QOL measure, while also making a correction for commuting costs. In table 6, we examine climate-preference and welfare estimates based on alternative measures. First, in column II, we use wage differentials by place of residence rather than place of work, without correcting for commuting. This change slightly attenuates the WTP estimates, leading to smaller inferred damages. Intuitively, the wage and commuting correction to QOL does not substantially impact the WTP estimates because these factors operate on a small geographic scale that is likely orthogonal to climate variation. We next consider the separate roles that wages and housing costs play in determining QOL. In column III, we measure QOL using only the additive inverse of wages (so that low wages denote higher QOL). Without state FE, this change causes extreme temperatures to be valued positively; with fixed effects, the specification finds a positive (though statistically insignificant) WTP for heat, recalling the results of

---

scenario, worsening areas may only depopulate slowly as the supply of housing diminishes slowly over time due to depreciation, consistent with a small downside supply elasticity (Glaeser and Gyourko 2005). It is therefore likely that our use of a supply elasticity of 8 results in an overestimate of migration.

<sup>40</sup> The ability of migration to mitigate welfare losses would be further constrained by adjustment costs, including the cost of physically moving.



Hoch and Drake (1974) and Moore (1998).<sup>41</sup> When only housing costs are used to measure QOL in column IV, the estimates exaggerate households' aversion to extreme temperatures and damages to climate change. Column V uses a QOL measure similar to Roback (1982), Nordhaus (1996), and Sinha and Cropper (2013) that, relative to our measure, puts 2.7 times more weight on wages relative to housing costs (by not accounting for taxation and non-housing local costs). Not surprisingly, these results resemble those based only on wages, albeit less extremely. The dramatically different results with and without fixed effects also evokes the lack of robustness Nordhaus (1996) finds using his QOL measure.

Columns VI and VII of table 6 return to our baseline QOL measures but alternatively use wage levels for workers with a college degree or with only a high school degree, to test whether these two groups value amenities differently. Our findings suggest that more educated workers may have a lower aversion to heat and cold, perhaps because their higher incomes allow them to better protect themselves or because they are less likely to work outdoors.

The last column adjusts our estimates for mobility costs. As discussed in Albouy (2011), mobility costs or idiosyncratic preference heterogeneity can cause new migrants to a city to have a higher WTP to live there than current (infra-marginal) residents. We adjust our QOL measure to account for mobility costs by adding a correction based on migration rates from 1990 to 2000, so that locations that are growing relatively quickly have their QOL increased. We multiply the percentage change in population in each PU-

<sup>41</sup> In the column III state FE specification, the estimated welfare impact is similar to that obtained with the reference case QOL measure despite the very different MWTPs at 40°F and 80°F. This result occurs only because the estimated MWTP is very large and positive over extreme cold <30°F. The estimates in column V with state FE behave similarly.

MA by a mobility cost estimate from Notowidigdo (2012).<sup>42</sup> This adjustment slightly reduces our estimated value of avoiding both extreme heat and extreme cold, reducing the estimated cost of climate change by only 0.3 percent of income. This reduction results from the fact that areas with relatively extreme temperatures (particularly extreme heat) are growing more quickly than areas with relatively moderate temperatures.

## VII. Conclusions

This paper illustrates the remarkable ability of the Rosen-Roback hedonic framework, when properly calibrated, to recover willingness to pay for non-market goods such as climate. Using a quality of life measure that is carefully constructed from local wage and housing price differentials, we find estimates of households' valuation of climate amenities that both accord with intuition and are generally robust to a wide range of empirical specifications. We find that households: (1) place the most value on temperatures near 65°F, (2) tend to dislike marginal increases in heat more than marginal increases in cold, and (3) care less about marginal changes in outdoor temperature once the temperature is sufficiently uncomfortable that they are unlikely to go outside. Finally, we find evidence of heterogeneity in these preferences, with households that are most averse to cold living in the South, consistent with models of both sorting and adaptation.

Even though our estimates are generally robust, they may nonetheless still be biased by omitted variables. Despite this limitation, we believe that households value climate amenities considerably and that our work is an important step forward in this area as it

<sup>42</sup> The mobility cost elasticity, which dictates how the mobility cost of the marginal migrant increases with the number of migrants, comes from a weighted average of Notowidigdo's (2012) estimated elasticities for high-skill and low-skill workers (0.066 and 0.065 respectively, see table 7 in his paper). These estimates are identified from plausibly exogenous Bartik (1991) labor demand shocks and correspond to migration over a ten-year period, thereby aligning with our 1990 to 2000 migration rates. We also use Notowidigdo's (2012) curvature parameters (betas) in specifying the mobility cost function.

properly calculates the quality of life measure, flexibly models local temperature distributions across the contiguous U.S., and accounts for preference heterogeneity.

When we apply our estimates to value the amenity impacts of climate change, we estimate an average U.S. welfare loss between one and three percent of income per year by 2090–2099 under “business as usual”. Most areas in the United States are forecast to experience losses, as many pleasantly warm days will be ruined while uncomfortably cold days will only be moderately reduced. These estimates are similar in magnitude to the percentage losses in GDP (market goods) forecasted by Nordhaus (2007) and the Stern report (2007) for 2100 under business as usual, suggesting that nonmarket amenity benefits alone may justify the mitigation costs discussed in their studies.<sup>43</sup>

The greatest limitation to our empirical method is that we have little statistical power to make inferences regarding households’ willingness to pay (WTP) to avoid temperatures beyond the limits of the current temperature distribution. The data provide only limited guidance as to how WTP should be extrapolated over extremely hot temperatures that have yet to be realized. Should the WTP to avoid heat actually increase non-linearly over the extremes, our damage estimates will be too conservative, particularly in the southern U.S.

<sup>43</sup> The Nordhaus and Stern market damage estimates of 3% and 2.9% apply to worldwide GDP, not U.S. GDP.

## References

- Albouy, David Y. (2011) "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas." NBER Working Paper #14472 (original 2008, revised December 5, 2011).
- Albouy, David Y. and Bert Lue (2011) "Driving to Opportunity: Local Wages, Commuting, and Sub-Metropolitan Quality of Life," mimeo, University of Michigan.
- Albouy, David Y. and Bryan Stuart (2012) "Urban Population and Amenities," mimeo, University of Michigan.
- Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber (2005) "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy* 113, pp. 151-184.
- Arellano, Manuel, "Computing Robust Standard Errors for Within-Groups Estimators," *Oxford Bulletin of Economics and Statistics*, 49 (1987), pp. 431-434.
- Bajari, Patrick and C. Lanier Benkard (2005) "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," *Journal of Political Economy* 113, pp. 1239-1276.
- Bajari, Patrick and Matthew E. Kahn (2005) "Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities," *Journal of Business and Economic Statistics* 23, pp. 20-33.
- Barreca, Alan, Karen Clay, Olivier Deschênes, Michael Greenstone, and Joseph S. Shapiro (2013) "Adapting to Climate Change: The Remarkable Decline in the U.S. Temperature-Mortality Relationship over the 20<sup>th</sup> Century," NBER working paper #18700.
- Bartik, Timothy J. (1991) "Who Benefits from State and Local Economic Development Policies," W.E. Upjohn Institute for Employment Research: Kalamazoo, MI.
- Baum-Snow, Nathaniel and Ronni Pavan (2012) "Understanding the City Size Wage Gap," *Review of Economic Studies* 79, pp. 88-127.
- Blomquist, Glenn C, Mark C. Berger and John P. Hoehn (1988) "New Estimates of Quality of Life in Urban Areas." *American Economic Review* 78, pp. 89-107.

- Cameron, A. Colin and Pravin K. Trivedi (2005) *Microeconometrics: Methods and Applications*, New York: Cambridge University Press.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2008) "Bootstrap-Based Improvements for Inferences with Clustered Errors," *Review of Economics and Statistics* 90, pp. 414-427.
- Chay, Kenneth Y. and Michael Greenstone (2005), "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy* 113, pp. 376-424.
- Chen, Yong and Stuart S. Rosenthal (2008) "Local Amenities and Life-Cycle Migration: Do People Move for Jobs or Fun?" *Journal of Urban Economics* 64, pp. 519-537.
- Cragg, Michael I. and Matthew E. Kahn (1997) "New estimates of climate demand: evidence from location choice. *Journal of Urban Economics* 42, pp. 261-284.
- Cragg, Michael I. and Matthew E. Kahn (1999) "Climate consumption and climate pricing from 1940 to 1990," *Regional Science and Urban Economics* 29, pp. 519- 539.
- Deschênes, Olivier and Michael Greenstone (2007) "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather," *American Economic Review*, 97, pp. 354-385.
- Deschênes, Olivier and Michael Greenstone (2011): Climate Change, Mortality and Adaptation: Evidence from Annual Fluctuations in Weather in the U.S.," *American Economic Journal: Applied Economics* 3, pp. 152-185.
- Deschênes, Olivier and Michael Greenstone (2012) "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather: Reply," *American Economic Review*, 102, pp. 3761-3773.
- Fan, Jianqing and Irène Gijbels (1996), *Local Polynomial Modelling and Its Applications*. London: Chapman and Hall.
- Fisher, Anthony C., W. Michael Hanemann, Michael J. Roberts, and Wolfram Schlenker (2012) "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather: Comment," *American Economic Review* 102, pp. 3749-3760.
- Frijters, P. and Van Praag, B.M.S. (1998) "The effects of climate on welfare and wellbe-

- ing in Russia.” *Climatic Change* 39, pp. 61–81.
- Glaeser, Edward L. and Joseph Gyourko (2005) “Urban Decline and Durable Housing.” *Journal of Political Economy* 113, pp. 345-375.
- Graff Zivin, Joshua and Matthew Neidell (2012) “Temperature and the Allocation of Time: Implications for Climate Change,” *Journal of Labor Economics*, forthcoming.
- Graves, Philip E. (1979) “A Life-Cycle Empirical Analysis of Migration and Climate, by Race.” *Journal of Urban Economics* 6, pp. 135-47.
- Hoch, Irving and Judith Drake (1974) “Wages, Climate, and the Quality of Life.” *Journal of Environmental Economics and Management* 1: pp. 268-296.
- International Panel on Climate Change Working Group I (2007) *Climate Change 2007: The Physical Science Basis*. Published for the International Panel on Climate Change.
- Kahn, Matthew E. (2009) “Urban Growth and Climate Change.” *Annual Review of Resource Economics* 1, pp. 333-350.
- Kuminoff, Nicolai V. and Jaren C. Pope (2012) “Do ‘Capitalization Effects’ for Public Goods Reveal the Public’s Willingness to Pay?” working paper.
- Maddison, David (2003) “The Amenity Value of the Climate: the Household Production Function Approach,” *Resource and Energy Economics* 25, pp. 155-175.
- Maddison, David and Bigano, A. (2003) “The Amenity Value of the Italian Climate.” *Journal of Environmental Economics and Management* 45, pp. 319-332.
- Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw (1994) “The Impact of Global Warming on Agriculture: A Ricardian Analysis,” *American Economic Review* 84, pp. 753-771.
- Moore, Thomas Gale (1998) “Health and Amenity Effects of Global Warming.” *Economic Inquiry* 36, pp. 471–488.
- Nordhaus, William D. (1996) “Climate amenities and global warming”, in *Climate Change: Integrating Science, Economics, and Policy*. N. Nakicenovic, W.D. Nordhaus, R. Richels, and F. Toth, eds. IIASA, pp. 3–45.
- Nordhaus, William D. (2007) *The Challenge of Global Warming: Economic Models and Environmental Policy*. Vol 4. Yale University.

- Notowidigdo, Matthew J. (2012), "The Incidence of Local Labor Demand Shocks," mimeo.
- Peiser, Richard B. and Lawrence B. Smith (1985) "Homeownership Returns, Tenure Choice and Inflation," *American Real Estate and Urban Economics Journal* 13, pp. 343-360.
- Rehdanz, Katrin (2006) "Hedonic Pricing of Climate Change Impacts to Households in Great Britain." *Climatic Change*, 74, pp. 413-434.
- Rehdanz, Katrin, and David J. Maddison (2005) "Climate and Happiness." *Ecological Economics* 52, pp. 111-25.
- Rehdanz, Katrin and David J. Maddison (2009) "The Amenity Value of Climate to German Households." *Oxford Economic Papers* 61, pp. 150-167.
- Roback, Jennifer (1982) "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90, pp. 1257-1278.
- Rosen, Sherwin (1974) "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82, pp. 34-55.
- Rosen, Sherwin (1979) "Wages-based Indexes of Urban Quality of Life," in *Current Issues in Urban Economics*. P. Mieszkowski and M. Straszheim, eds., Baltimore: John Hopkins Univ. Press.
- Saiz, Albert (2010) "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics* 125, pp. 1253-1296.
- Savaeau, David (1999) *Places Rated Almanac*. Foster City, CA: IDG Books Worldwide.
- Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher (2006) "The Impact of Global Warming on U.S. Agriculture: An Econometric Analysis of Optimal Growing Conditions." *Review of Economics and Statistics* 88, pp. 113-125.
- Schlenker, Wolfram, and Michael J. Roberts (2009) "Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields Under Climate Change." *Proceedings of the National Academy of Sciences*, 106, pp. 15594-15598.
- Shapiro, P. and Smith, T., (1981) "Preferences on Non Market Goods Revealed Through Market Demands." *Advances in Applied Microeconomics: A Research Annual* 1, pp.

105-122.

Sinha, Paramita and Maureen L. Cropper (2013) “The Value of Climate Amenities: Evidence from U.S. Migration Decisions.” NBER Working Paper 18756.

Stern, Nicholas (2007) *The Economics of Climate Change: the Stern Review*. Cambridge University Press.

Timmins, Christopher (2007) “If You Can’t Take the Heat, Get Out of the Cerrado...Recovering the Equilibrium Amenity Cost of Nonmarginal Climate Change in Brazil.” *Journal of Regional Science* 47, pp. 1-25.

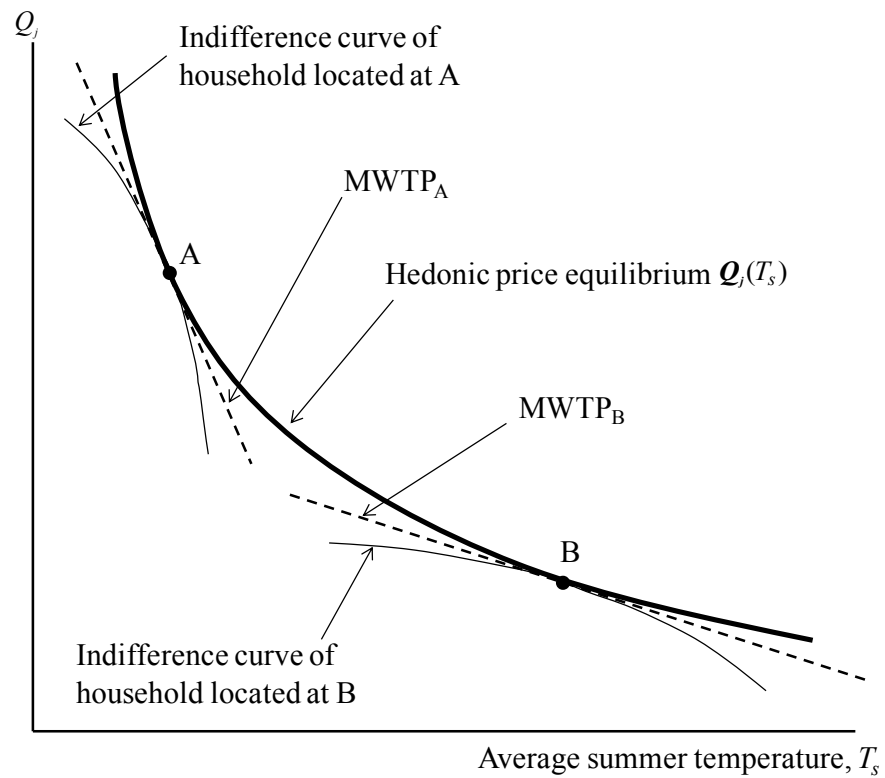
Tol, Richard S. J. (2002) “Estimates of the Damage Costs of Climate Change, Part II. Dynamic Estimates.” *Environmental and Resource Economics* 21, pp. 135–160.

Tol, Richard S. J. (2009) “The Economic Effects of Climate Change.” *Journal of Economic Perspectives* 23, pp. 29–51.

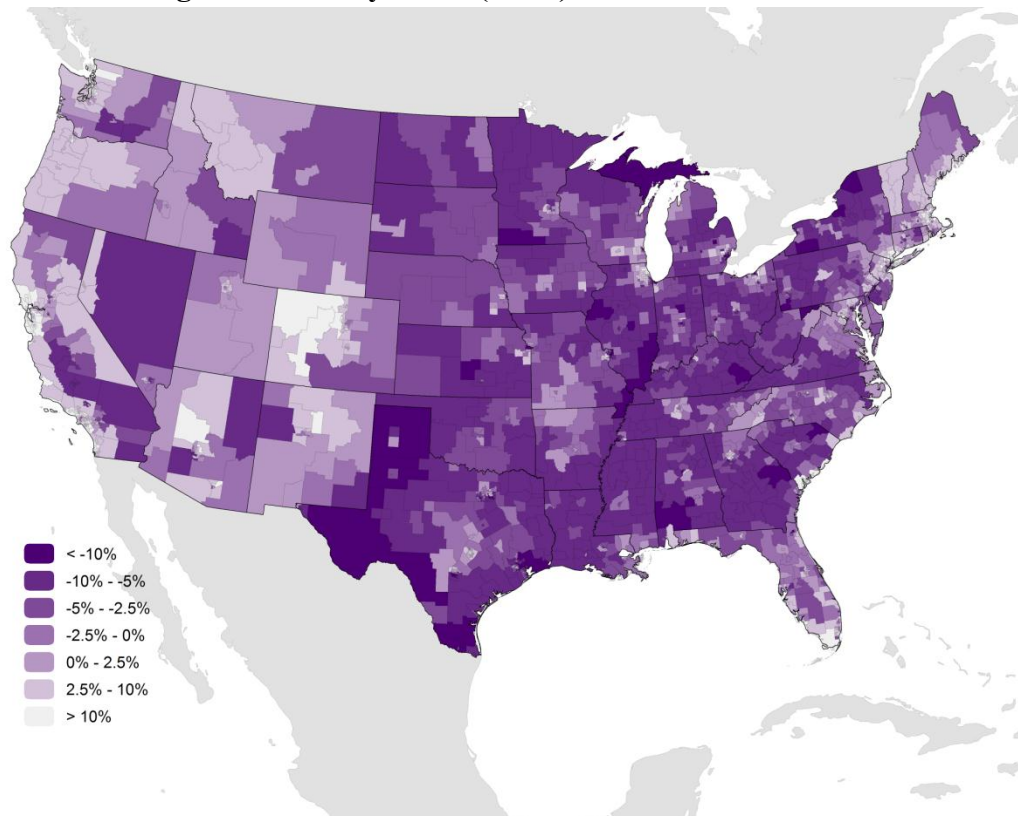
Wooldridge, Jeffrey M., “Cluster-Sample Methods in Applied Econometrics,” *American Economic Review Papers and Proceedings*, 93 (2003), 133-138.



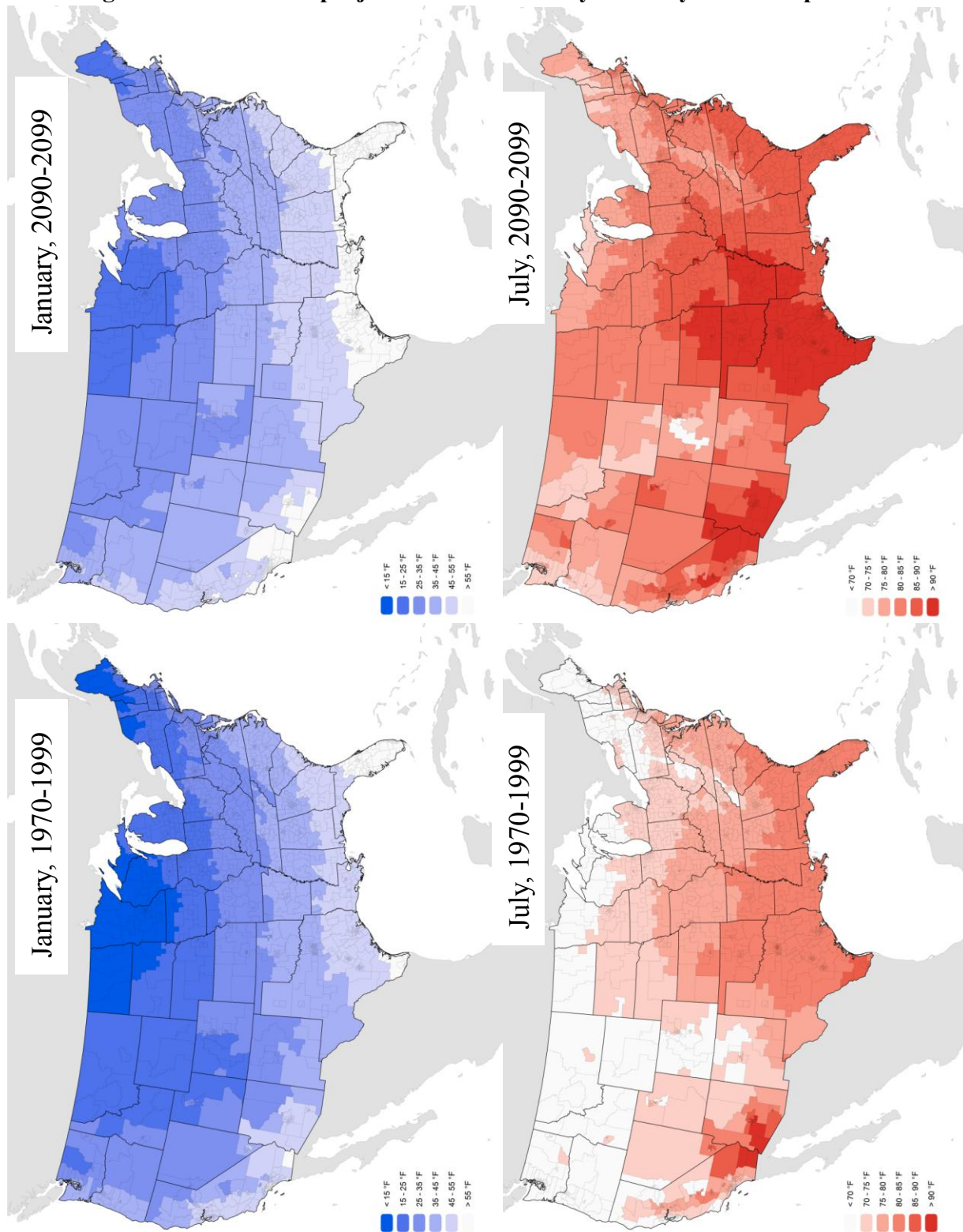
**Figure 1: Illustrative hedonic price function with demand-side equilibrium FOC's satisfied**



**Figure 2: Quality of life (QOL) differentials in 2000**



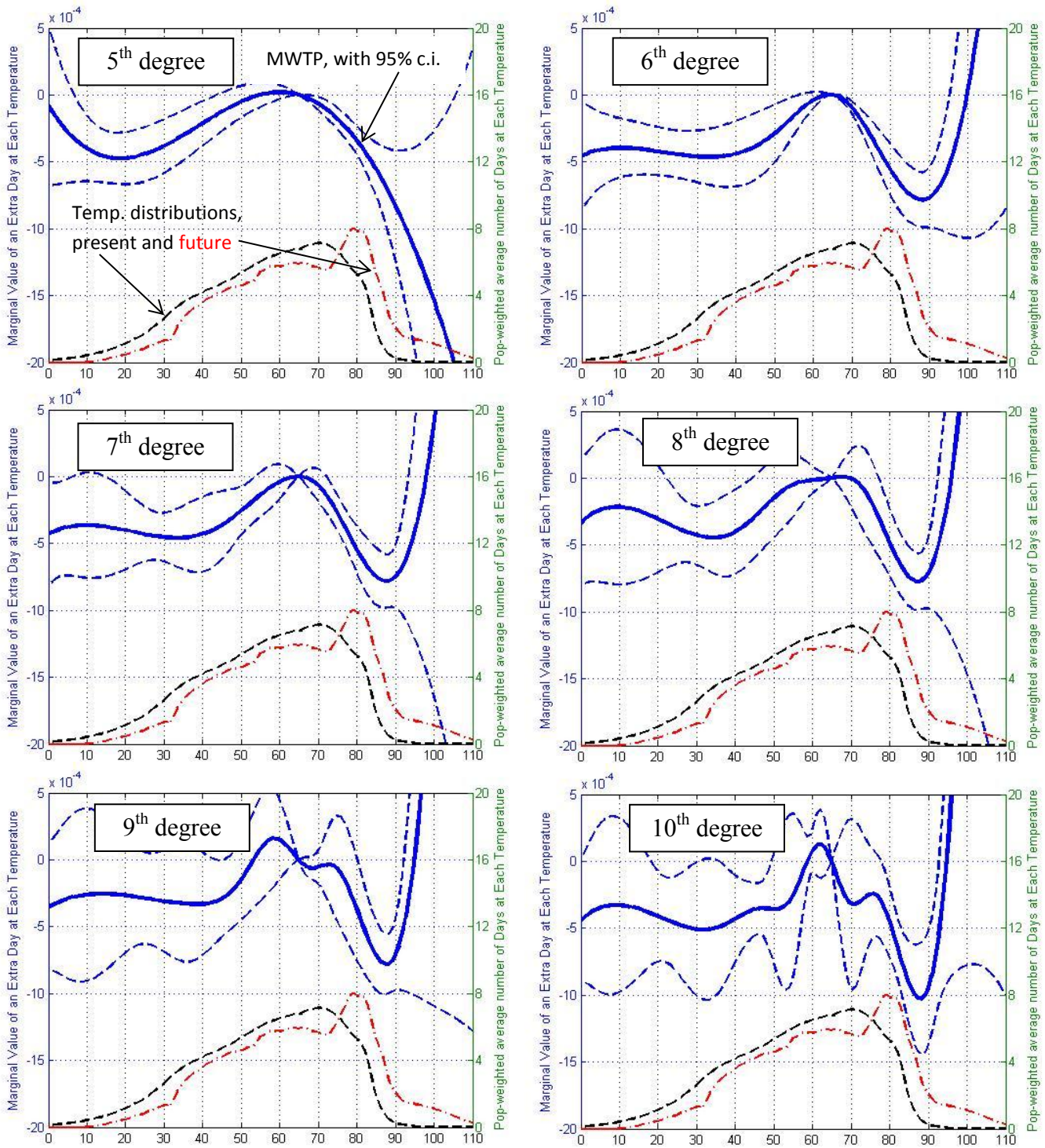
**Figure 3: Present and projected future January and July mean temperatures**



*Notes:* Present-day temperatures are 1970-1999 average. Projections use the A2 scenario for 2090-2099.

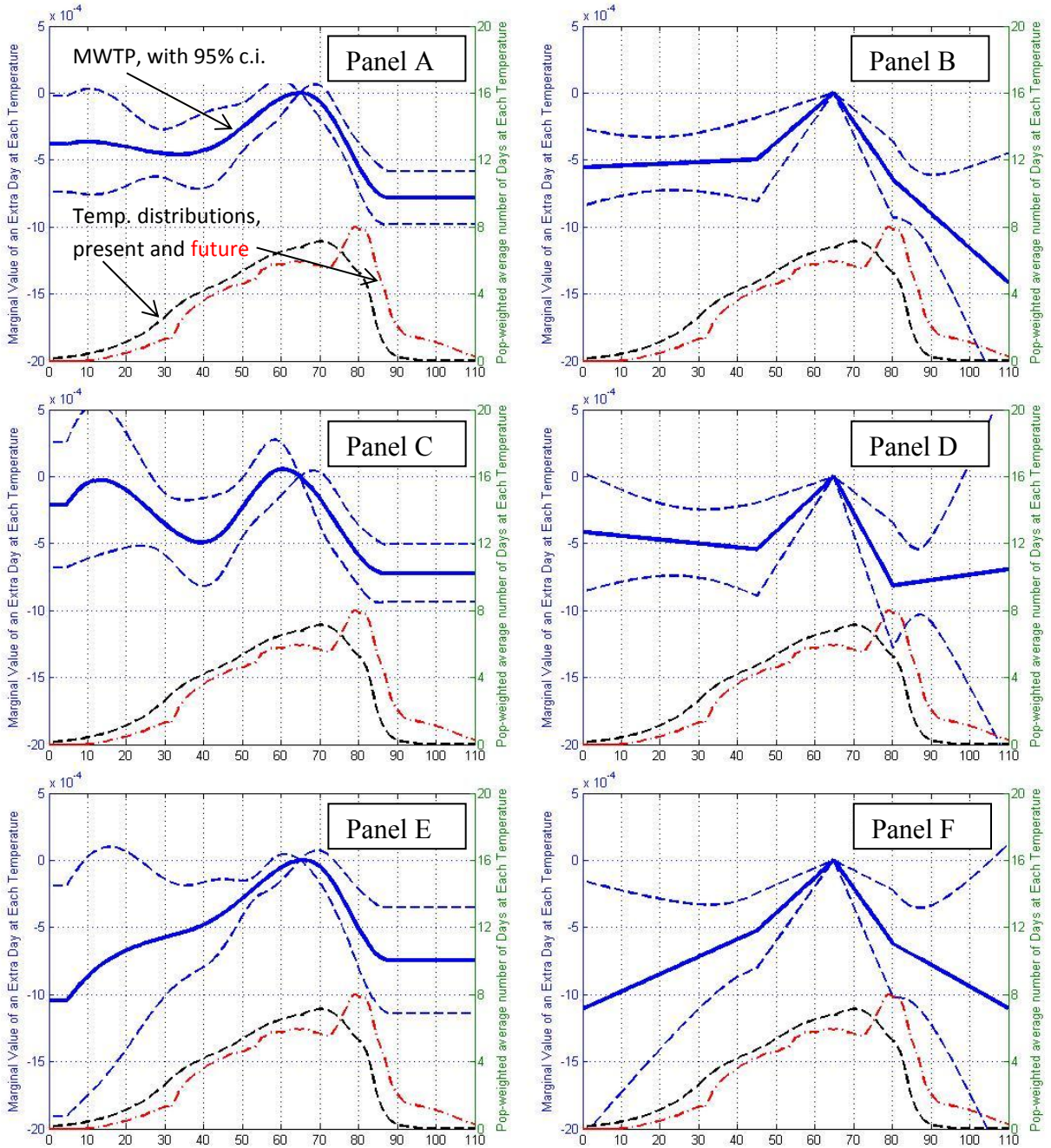


**Figure 4: Estimated WTP for temperature, homogenous preference cubic spline models**



*Notes:* All specifications use the full set of geographic, “other weather”, and demographic controls. MWTPs are expressed as a fraction of income and normalized to zero at 65°F. Each observation (PUMA) is weighted by its population in the regressions. Standard errors are clustered on state. “Present” temperature distribution denotes 1970-1999, and “future” denotes 2090-2099. Both distributions are population-weighted U.S. averages.

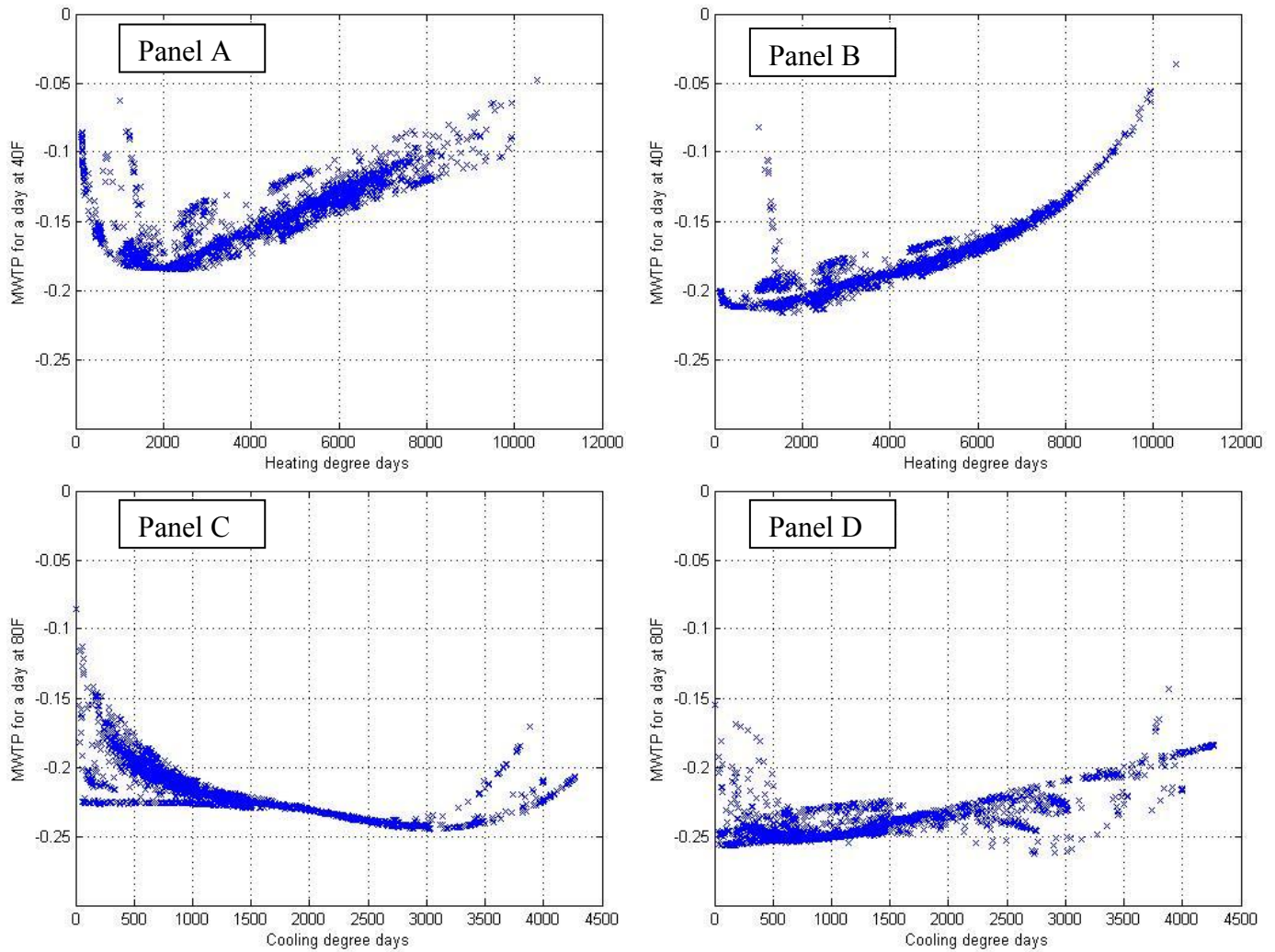
**Figure 5: Estimated WTP for temperature, homogenous preference models**



*Notes:* Panels A, C, and E, use a 7<sup>th</sup> degree cubic spline WTP model, restricted to be constant over the extreme 1% of the temperature distribution. Panels B, D, and E use a four-piece linear spline. Panels A and B include all geographic, "other weather", and demographic controls. Panels C and D omit demographic controls. Panels E and F include all controls and state fixed effects. MWTPs are expressed as a fraction of income and normalized to zero at 65°F. Each observation (PUMA) is weighted by its population in the regressions. Standard errors are clustered on state. "Present" temperature distribution denotes 1970-1999, and "future" denotes 2090-2099. Both distributions are population-weighted U.S. averages.

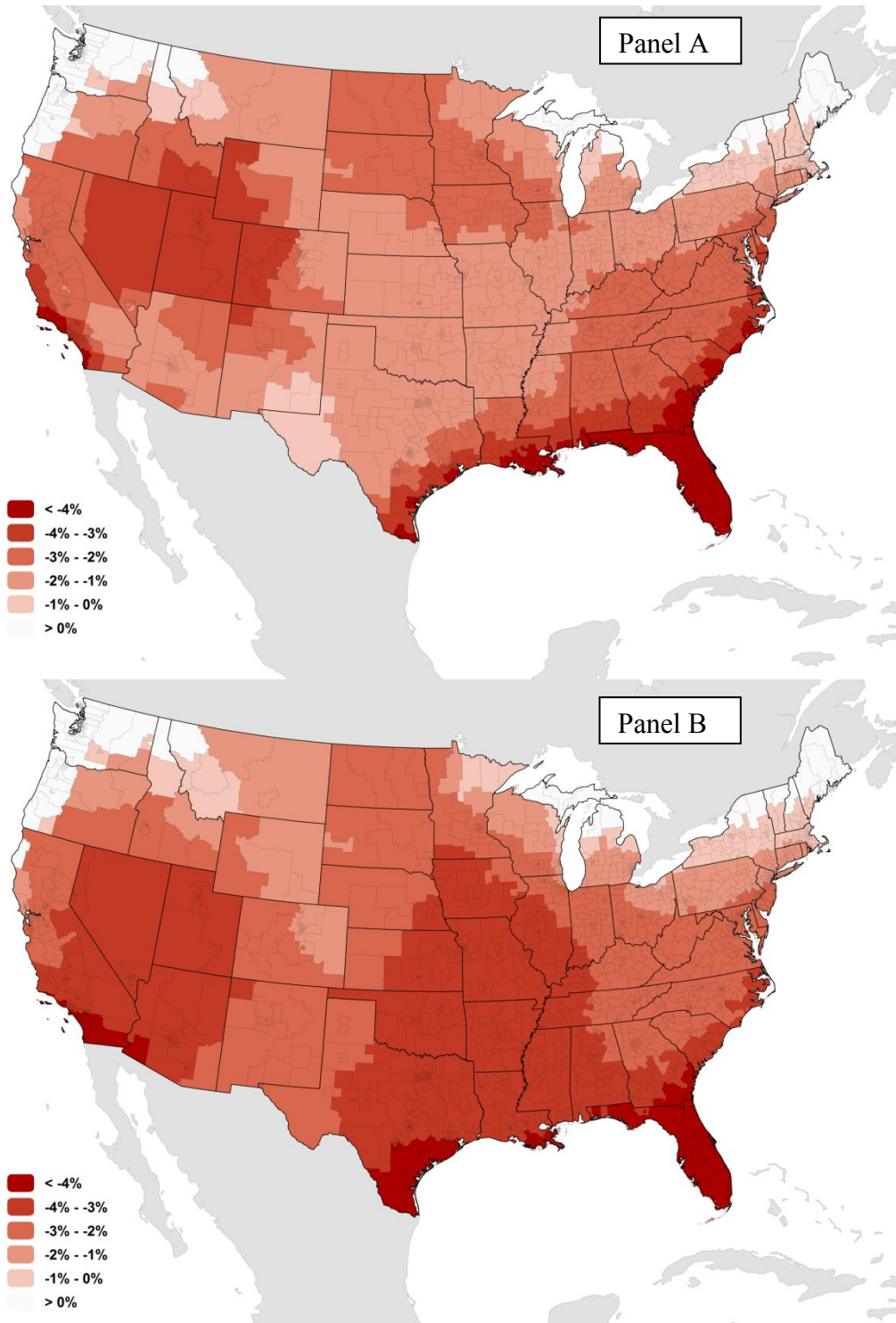


**Figure 6: Estimated distribution of MWTPs for heat and cold from the heterogeneous preferences models**



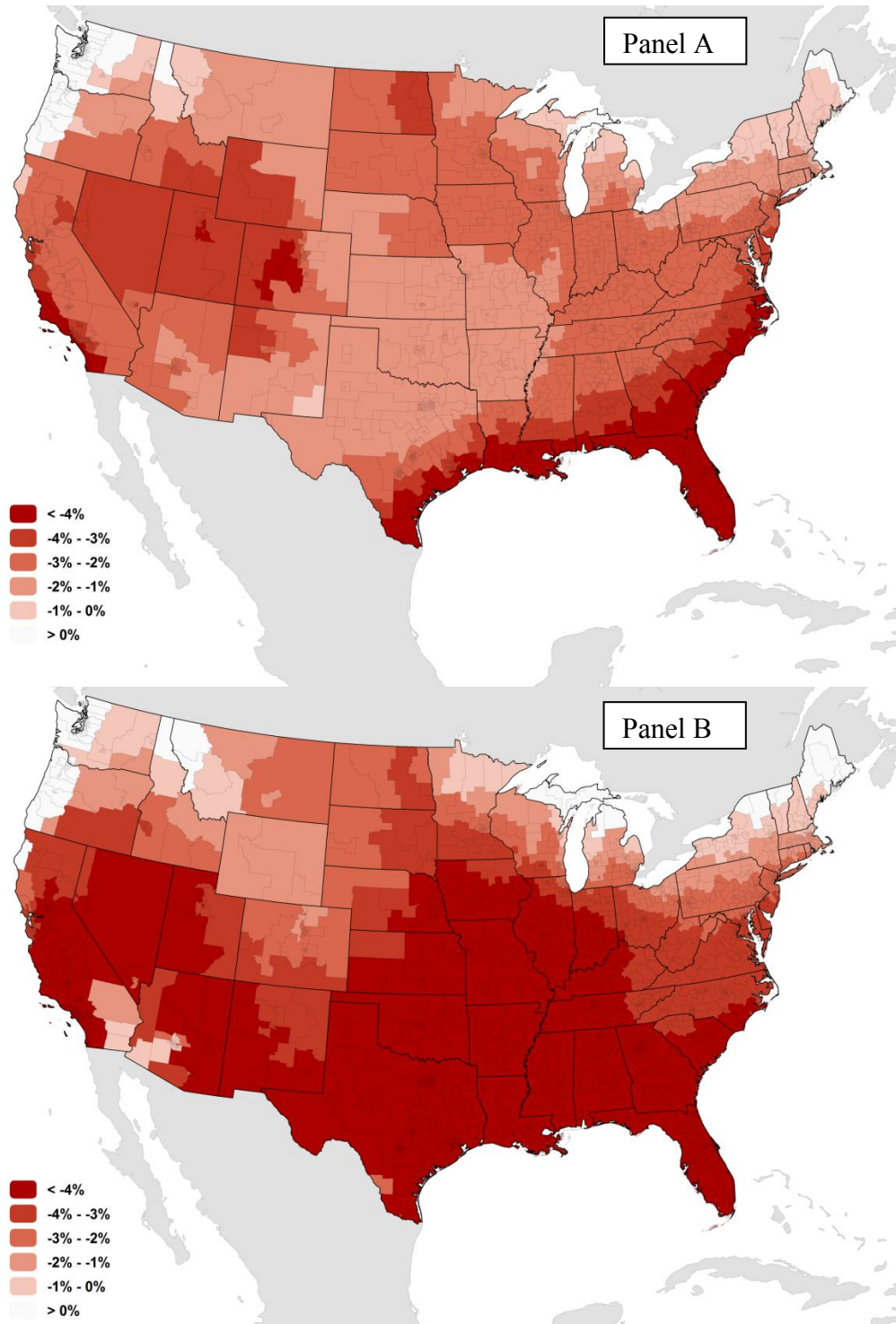
*Notes:* MWTPs are expressed as a fraction of income, scaled up by 365, and normalized to zero at 65°F. Each plotted point denotes a PUMA-specific estimate of MWTP. Panels A and C use the restricted 7<sup>th</sup> degree cubic spline model, and panels B and D use the four-piece linear spline. Models used in all panels include the geographic, “other weather”, and demographic controls. Each observation (PUMA) is weighted by its population in the regressions.

**Figure 7: Predicted changes in QOL as percent of income under the A2 scenario for 2090-2099, with the homogenous preference model WTP estimates**



*Notes:* Panel A uses the restricted 7<sup>th</sup> degree cubic spline WTP model, per panel A of figure 5 and specification I of table 2. Panel B uses the four-piece linear spline WTP model, per panel B of figure 5 and specification II of table 2.

**Figure 8: Predicted changes in QOL as percent of income under the A2 scenario for 2090-2099, with the heterogeneous preference model MWTP estimates**



*Notes:* Panel A uses the restricted 7<sup>th</sup> degree cubic spline WTP model, per panels A and C of figure 6 and specification I of table 4. Panel B uses the four-piece linear spline WTP model, per panels B and D of figure 6 and specification II of table 4.

**Table 1: Descriptive statistics for primary dataset.**  
**Sample consists of 2057 Public Use Microdata Areas (PUMAs)**

	Mean	Std. Dev	10th percentile	90th percentile
<i>Temperature data, 1970-1999 average</i>				
Average annual heating degree days (1000s)	4.384	2.204	1.326	7.009
Average annual cooling degree days (1000s)	1.290	0.929	0.374	2.762
<i>Temperature data, 2090-2099 projected (CCSM A2)</i>				
Projected avg. annual heating degree days (1000s)	2.821	1.698	0.425	4.845
Projected avg. annual cooling degree days (1000s)	2.747	1.234	1.447	4.719
<i>Other climate data, 1970-1999 average</i>				
Average annual precipitation (inches)	39.26	14.09	16.25	53.85
Average annual relative humidity (%)	63.58	8.10	53.31	70.52
Average annual sunshine (% of available daylight)	60.17	8.64	49.74	73.08
<i>Other climate data, 2090-2099 projected (CCSM A2)</i>				
Projected avg. annual precipitation (inches)	41.60	15.56	14.68	56.91
Projected avg. annual relative humidity (%)	62.41	8.94	50.64	69.44
Projected avg. annual sunshine (% of available daylight)	61.55	8.63	51.29	72.81
<i>Geographic data</i>				
Distance from centroid of puma to ocean (miles)	250.1	272.1	4.3	729.2
Distance from centroid of puma to Great Lake (miles)	763.2	715.4	54.0	2128.4
Average land slope (degrees)	1.677	2.131	0.191	4.270
<i>Demographic data (2000 census)</i>				
Weighted population density (people per sq. mile)	5,466	11,997	360	9,981
Percent high school graduates	80.1%	10.0%	67.2%	90.9%
Percent of population with bachelors degree	24.1%	12.4%	11.3%	41.0%
Percent of population with graduate degree	8.7%	5.6%	3.7%	20.3%
Average age	48.7	2.6	45.3	51.8
Percent hispanic	8.9%	13.6%	0.6%	81.6%
Percent black	12.3%	17.6%	0.5%	96.7%
Population (1000s)	135.9	32.9	103.8	182.2
Quality of life differential (in logs)	0.034	0.071	-0.043	0.128



**Table 2: Estimation results for the linear, homogenous preferences model.**  
**Dependent variable is quality of life (QOL) differential as fraction of income**

	I	II	III	IV	V	VI
	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline
Temperature function						
MWTP for a day at 40F (x365)	-0.158 (0.052)	-0.184 (0.051)	-0.179 (0.061)	-0.193 (0.056)	-0.177 (0.059)	-0.216 (0.053)
MWTP for a day at 80F (x365)	-0.203 (0.033)	-0.236 (0.053)	-0.217 (0.043)	-0.297 (0.087)	-0.188 (0.055)	-0.227 (0.075)
Average annual precipitation (inches, x10)	0.008 (0.002)	0.008 (0.002)	0.003 (0.003)	0.003 (0.003)	0.008 (0.003)	0.008 (0.003)
Average annual relative humidity (percent)	0.013 (0.051)	-0.001 (0.045)	0.023 (0.077)	0.025 (0.059)	-0.019 (0.051)	-0.031 (0.052)
Average annual sunshine (fraction of all daytime hours)	0.198 (0.054)	0.200 (0.052)	0.176 (0.078)	0.179 (0.068)	0.104 (0.079)	0.118 (0.079)
Average land slope (degrees)	0.005 (0.001)	0.004 (0.001)	0.007 (0.002)	0.007 (0.002)	0.004 (0.001)	0.004 (0.001)
Inverse distance to ocean (miles)	0.233 (0.032)	0.231 (0.032)	0.341 (0.044)	0.340 (0.045)	0.225 (0.036)	0.218 (0.036)
Inverse distance to Great Lakes (miles)	0.144 (0.110)	0.154 (0.110)	0.004 (0.191)	0.034 (0.183)	0.080 (0.112)	0.089 (0.111)
Squared inverse distance to ocean (miles)	-0.219 (0.029)	-0.217 (0.030)	-0.296 (0.061)	-0.294 (0.059)	-0.224 (0.031)	-0.219 (0.032)
Squared inverse distance to Great Lakes (miles)	-0.145 (0.122)	-0.156 (0.121)	0.089 (0.209)	0.054 (0.197)	-0.092 (0.122)	-0.101 (0.122)
Log of weighted population density	0.004 (0.002)	0.004 (0.002)	-	-	0.002 (0.002)	0.002 (0.002)
Fraction high school graduates	0.109 (0.029)	0.107 (0.030)	-	-	0.098 (0.027)	0.095 (0.027)
Fraction bachelors degrees	0.278 (0.047)	0.272 (0.049)	-	-	0.284 (0.048)	0.286 (0.049)
Fraction graduate degrees	0.039 (0.078)	0.048 (0.086)	-	-	0.062 (0.074)	0.060 (0.076)
Average age (x100)	0.152 (0.050)	0.148 (0.049)	-	-	0.154 (0.041)	0.147 (0.041)
Fraction hispanic	0.048 (0.028)	0.043 (0.032)	-	-	0.057 (0.022)	0.055 (0.024)
Fraction black	-0.072 (0.007)	-0.074 (0.007)	-	-	-0.064 (0.008)	-0.066 (0.007)
State fixed effects	N	N	N	N	Y	Y
Number of observations	2057	2057	2057	2057	2057	2057
R <sup>2</sup>	0.782	0.780	0.366	0.364	0.822	0.822

*Notes :* The unit of observation is a PUMA. All regressions are weighted by population in each PUMA. The linear splines are four-piece with breakpoints at 45F, 65F, and 80F, and the restricted cubic spline is 7th degree with a constant MWTP restriction for the lower and upper 1% of the population-weighted U.S. average temperature distribution. MWTPs shown for a day at 40F or 80F are relative to 65F and expressed as a fraction of income. Parenthetical values indicate standard errors clustered on state.

**Table 3: Estimation results for the heterogeneous preferences model.**  
**Dependent variable is quality of life (QOL) differential as fraction of income**

	I	II	III	IV	V	VI
Temperature function	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline
Average MWTP for a day at 40F (x365)	-0.145 (0.047)	-0.179 (0.045)	-0.169 (0.057)	-0.201 (0.051)	-0.175 (0.051)	-0.215 (0.044)
Average MWTP for a day at 80F (x365)	-0.215 (0.023)	-0.241 (0.045)	-0.283 (0.045)	-0.357 (0.063)	-0.205 (0.046)	-0.236 (0.060)
Geog & other weather controls	Y	Y	Y	Y	Y	Y
Demographic controls	Y	Y	N	N	Y	Y
State fixed effects	N	N	N	N	Y	Y

*Notes :* The unit of observation is a PUMA. All regressions are weighted by population in each PUMA. The linear splines are four-piece with breakpoints at 45F, 65F, and 80F, and the restricted cubic spline is 7th degree with a constant MWTP restriction for the lower and upper 1% of the population-weighted U.S. average temperature distribution. MWTPs shown for a day at 40F or 80F are relative to 65F, expressed as a fraction of income, and are income-weighted averages across all PUMAs. Parenthetical values indicate standard errors wild cluster bootstrapped on state using 200 draws.

**Table 4: Projected U.S. average amenity value impacts from climate change for 2090-2100 using the A2 scenario and preference estimates from the homogenous preference models.**

	I	II	III	IV	V	VI
Temperature function	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline
Change in welfare as percent of income	-2.24% (0.37%)	-2.52% (0.69%)	-2.72% (0.50%)	-2.16% (0.84%)	-1.23% (0.81%)	-0.97% (1.04%)
Component driven by warmer winters	1.72% (0.33%)	2.23% (0.43%)	1.13% (0.42%)	2.13% (0.51%)	2.58% (0.87%)	3.19% (0.86%)
Component driven by hotter summers	-4.38% (0.58%)	-5.20% (0.83%)	-4.13% (0.64%)	-4.57% (0.89%)	-4.15% (1.08%)	-4.54% (1.31%)
Component driven by other weather changes	0.42% (0.09%)	0.44% (0.08%)	0.28% (0.13%)	0.28% (0.12%)	0.34% (0.14%)	0.38% (0.13%)
Percent of population losing	96.1% (2.4%)	96.1% (3.9%)	97.3% (1.9%)	97.3% (11.5%)	80.2% (16.3%)	76.1% (20.3%)
Geog & other weather controls	Y	Y	Y	Y	Y	Y
Demographic controls	Y	Y	N	N	Y	Y
State fixed effects	N	N	N	N	Y	Y

*Notes :* Parenthetical values indicate standard errors wild cluster bootstrapped on state using 200 draws.

**Table 5: Projected U.S. average amenity value impacts from climate change for 2090-2100 using the preference estimates from the heterogeneous preference models.**

	I	II	III	IV	V	VI
Temperature function	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline
A2 scenario change in welfare as percent of income	-2.83% (0.30%)	-3.84% (0.76%)	-3.34% (0.39%)	-2.49% (0.97%)	-1.24% (0.82%)	-0.87% (1.31%)
A1FI scenario change in welfare as percent of income	-3.37% (0.36%)	-4.74% (1.01%)	-3.88% (0.48%)	-2.85% (1.33%)	-1.69% (0.98%)	-1.25% (1.67%)
Geog & other weather controls	Y	Y	Y	Y	Y	Y
Demographic controls	Y	Y	N	N	Y	Y
State fixed effects	N	N	N	N	Y	Y

Notes : Parenthetical values indicate standard errors wild cluster bootstrapped on state using 200 draws.

**Table 6: Preference estimates and projected U.S. average welfare impacts from climate change for 2090-2100. Heterogeneous preferences model with alternative quality of life (QOL) measures**

	I	II	III	IV	V	VI	VII	VIII
QOL measure	Reference case, 7th degree cubic spline	Wages by place of residence	Wages only (negative)	Prices only	QOL = 0.25p - 1.00w	High school educated only	College educated only	Mobility cost adjustment
<u>Demographics, no state FE</u>								
Average MWTP for a day at 40F (x365)	-0.145 (0.047)	-0.105 (0.042)	0.248 (0.115)	-0.730 (0.177)	0.068 (0.087)	-0.257 (0.045)	-0.124 (0.054)	-0.126 (0.048)
Average MWTP for a day at 80F (x365)	-0.215 (0.023)	-0.165 (0.041)	0.584 (0.149)	-1.424 (0.145)	0.240 (0.120)	-0.268 (0.024)	-0.226 (0.028)	-0.203 (0.024)
CCSM A2 change in welfare as percent of income	-2.83% (0.30%)	-2.47% (0.37%)	5.53% (1.53%)	-15.97% (1.99%)	1.54% (1.13%)	-2.87% (0.35%)	-3.04% (0.36%)	-2.53% (0.34%)
<u>Demographics, state FE</u>								
Average MWTP for a day at 40F (x365)	-0.175 (0.051)	-0.170 (0.043)	-0.118 (0.066)	-0.365 (0.189)	-0.214 (0.044)	-0.240 (0.054)	-0.153 (0.062)	-0.161 (0.050)
Average MWTP for a day at 80F (x365)	-0.205 (0.046)	-0.206 (0.055)	0.048 (0.104)	-0.730 (0.256)	-0.131 (0.072)	-0.245 (0.048)	-0.167 (0.052)	-0.190 (0.047)
CCSM A2 change in welfare as percent of income	-1.24% (0.82%)	-1.92% (0.72%)	-2.18% (1.58%)	-2.84% (3.63%)	-2.89% (1.00%)	-1.33% (0.98%)	-0.53% (1.00%)	-0.86% (0.95%)

Notes : The unit of observation is a PUMA. All regressions are weighted by population in each PUMA. All specifications use a 7th degree restricted cubic spline with a constant MWTP restriction for the lower and upper 1% of the population-weighted U.S. average temperature distribution. All regressions include geographic, other weather, and demographic controls. MWTPs shown for a day at 40F or 80F are relative to 65F, expressed as a fraction of income, and are income-weighted averages across all PUMAs. Parenthetical values indicate standard errors wild cluster bootstrapped on state using 200 draws. Column V uses wages by residence.

**NOT FOR PUBLICATION**

**ONLINE APPENDICES FOR “CLIMATE AMENITIES, CLIMATE CHANGE,  
AND AMERICAN QUALITY OF LIFE”**

**by**

**David Albouy, Walter Graff, Ryan Kellogg, and Hendrik Wolff**

Appendix 1: Supplementary figures and tables

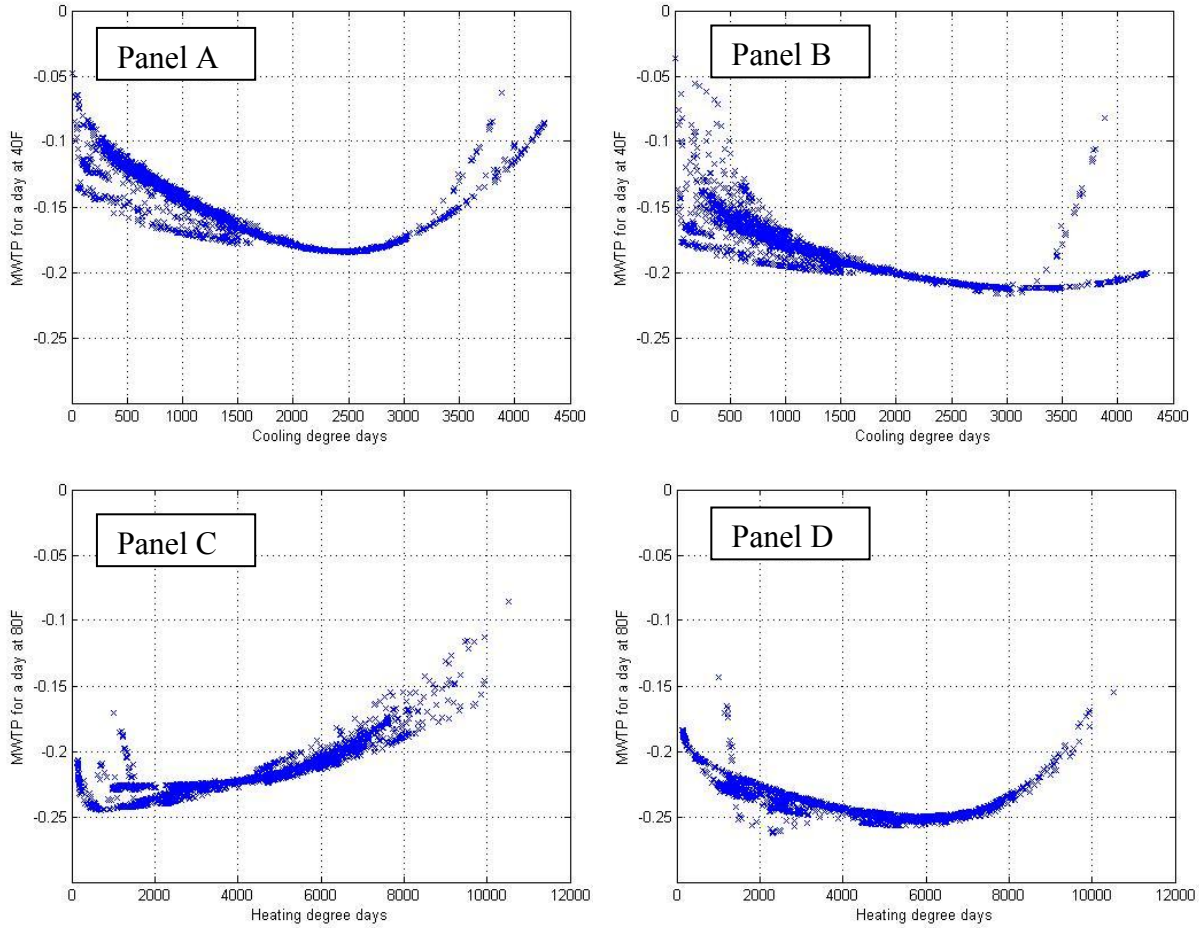
Appendix 2: Data construction

Appendix 3: Utility function concavity

Appendix 4: Set identification

## Appendix 1: Supplementary Figures and Tables

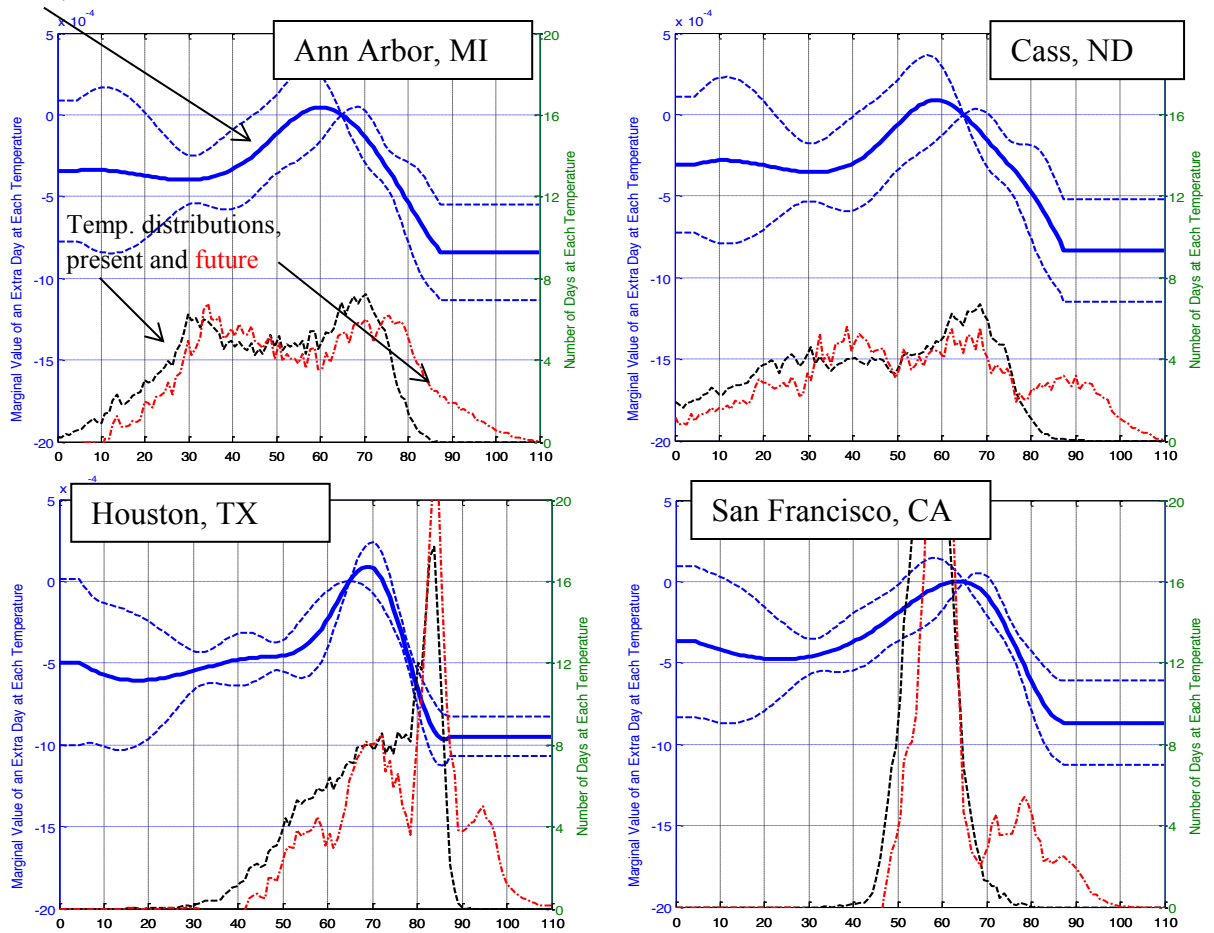
**Figure A1.1: Estimated MWTPs for heat and cold from the heterogeneous preferences models; MWTP to avoid cold (40°F) plotted against CDD and MWTP to avoid heat (80°F) plotted against HDD**



*Notes:* MWTPs are expressed as a fraction of income, scaled up by 365, and normalized to zero at 65°F. Each plotted point denotes a PUMA-specific estimate of MWTP. Panels A and C use the restricted 7<sup>th</sup> degree cubic spline model, and panels B and D use the four-piece linear spline. Models used in all panels include the geographic, “other weather”, and demographic controls. Each observation (PUMA) is weighted by its population in the regressions.

**Figure A1.2: Estimated WTP for temperature and four specific locations, heterogeneous preference model**

MWTP, with 95% c.i.



*Notes:* All panels use a 7<sup>th</sup> degree cubic spline WTP model, restricted to be constant over the extreme 1% of the temperature distribution, and include all geographic, “other weather”, and demographic controls. MWTPs are expressed as a fraction of income and normalized to zero at 65°F. Each observation (PUMA) is weighted by its population in the regressions. Standard errors are clustered on state. “Present” temperature distribution denotes 1970-1999, and “future” denotes 2090-2099.

Table A1.1, part 1: Alternative specifications for "geographic" and "other weather" controls

	Reference Case I	R1	R2	R3	R4	R5	R6	R7	R8
	Restricted cubic spline as in Specification I of tables 2-4	As in I, plus a cubic added in inverse distance to water bodies	As in I, plus a quadratic added in the average PUMA mountainousness	As in I, except "Relative Humidity" replaced by "Dew Point"	As in I, except "Relative Humidity" replaced by "Relative Humidity in summer months only"	As in I, except "Relative Humidity" replaced by "Dew Point in summer months only"	As in I, except all "other weather" variables split up into summer and winter variables*	As in R6, except "Relative Humidity" replaced by "Dew Point"	As in I, except "Sunshine" calculated as interpolated using all available stations (and not just the four closest stations)
<i>Panel a: Preference estimates and projected U.S. average amenity value impacts from climate change for 2090-2100 using the homogenous preference models</i>									
MWTP for a day at 40F (x365)	-0.158 (0.052)	-0.143 (0.053)	-0.164 (0.050)	-0.144 (0.070)	-0.197 (0.056)	-0.222 (0.052)	-0.152 (0.059)	0.000 (0.080)	-0.150 (0.053)
MWTP for a day at 80F (x365)	-0.203 (0.033)	-0.184 (0.032)	-0.208 (0.033)	-0.209 (0.032)	-0.218 (0.036)	-0.186 (0.028)	-0.190 (0.032)	-0.189 (0.021)	-0.199 (0.033)
A2 scenario change in welfare as percent of income	-2.24% (0.37%)	-2.01% (0.35%)	-2.30% (0.38%)	-2.17% (0.44%)	-2.53% (0.40%)	-3.42% (0.40%)	-2.35% (0.27%)	-2.57% (0.40%)	-2.24% (0.33%)
A1FI scenario change in welfare as percent of income	-2.71% (0.44%)	-2.41% (0.41%)	-2.78% (0.45%)	-2.62% (0.53%)	-3.09% (0.49%)	-4.18% (0.48%)	-2.82% (0.34%)	-3.05% (0.52%)	-2.68% (0.41%)

*Panel b: Preference estimates and projected U.S. average amenity value impacts from climate change for 2090-2100 using the heterogeneous preference models*

Average MWTP for a day at 40F (x365)	-0.145 (0.047)	-0.130 (0.048)	-0.150 (0.045)	-0.130 (0.067)	-0.184 (0.056)	-0.214 (0.051)	-0.142 (0.055)	0.002 (0.082)	-0.135 (0.048)
Average MWTP for a day at 80F (x365)	-0.215 (0.023)	-0.192 (0.024)	-0.222 (0.023)	-0.222 (0.024)	-0.231 (0.027)	-0.188 (0.022)	-0.199 (0.032)	-0.176 (0.024)	-0.211 (0.024)
A2 scenario change in welfare as percent of income	-2.83% (0.30%)	-2.56% (0.28%)	-2.91% (0.32%)	-2.76% (0.37%)	-3.10% (0.37%)	-3.71% (0.44%)	-2.79% (0.31%)	-2.65% (0.43%)	-2.83% (0.28%)
A1FI scenario change in welfare as percent of income	-3.37% (0.36%)	-3.03% (0.33%)	-3.47% (0.38%)	-3.29% (0.46%)	-3.73% (0.45%)	-4.50% (0.53%)	-3.33% (0.40%)	-3.13% (0.55%)	-3.35% (0.35%)

*Notes:* The unit of observation is a PUMA, except for R13, where the unit of observation is a MSA. MWTPs shown for a day at 40F or 80F are relative to 65F, expressed as a fraction of income, and are income-weighted averages across all observations. Except for R11, all regressions are weighted by population. Parenthetical values indicate standard errors cluster bootstrapped on state using 200 draws (except for first two rows, which are not bootstrapped). Dew point is highly collinear with temperature, so that in specifications R3, R5, and R7 the MWTPs for temperature decrease as this MWTP is loaded onto the dew point coefficient instead. The welfare predictions do not change substantially because the forecasted increase in dew point is similar to the forecasted increase in temperature.

\* This split effectively adds three variables to the specification

Table A1.1, part 2: Alternative specifications and standard errors

	Reference Case I	R9	R10	R11	R12	R13	R14
	Restricted cubic spline as in Specification I of tables 2-4	As in I, except "Age" and all "Minorities" excluded from the specification	As in I plus "Census Division Fixed Effects" added	As in I, except regressions are unweighted	As in I, except temperature modeled by each 10 degree bin**	As in I, except observations based on MSA level	As in I, except standard errors clustered by Census Division
<i>Panel a: Preference estimates and projected U.S. average amenity value impacts from climate change for 2090-2100 using the homogenous preference models</i>							
MWTP for a day at 40F (x365)	-0.158 (0.052)	-0.179 (0.052)	-0.197 (0.056)	-0.155 (0.054)	-0.186 (0.077)	-0.158 (0.055)	-0.158 (0.039)
MWTP for a day at 80F (x365)	-0.203 (0.033)	-0.205 (0.031)	-0.218 (0.036)	-0.203 (0.036)	-0.220 (0.039)	-0.179 (0.039)	-0.203 (0.030)
A2 scenario change in welfare as percent of income	-2.24% (0.37%)	-2.26% (0.36%)	-2.53% (0.40%)	-2.27% (0.36%)	-2.41% (0.58%)	-2.01% (0.46%)	-2.24% (0.23%)
A1FI scenario change in welfare as percent of income	-2.71% (0.44%)	-2.65% (0.44%)	-3.09% (0.52%)	-2.75% (0.45%)	-2.87% (0.73%)	-2.31% (0.57%)	-2.71% (0.31%)

*Panel b: Preference estimates and projected U.S. average amenity value impacts from climate change for 2090-2100 using the heterogeneous preference models*

Average MWTP for a day at 40F (x365)	-0.145 (0.047)	-0.167 (0.046)	-0.184 (0.055)	-0.141 (0.047)	-0.174 (0.100)	-0.147 (0.042)	-0.145 (0.039)
Average MWTP for a day at 80F (x365)	-0.215 (0.023)	-0.250 (0.020)	-0.231 (0.034)	-0.218 (0.025)	-0.222 (0.068)	-0.165 (0.036)	-0.215 (0.020)
A2 scenario change in welfare as percent of income	-2.83% (0.30%)	-2.81% (0.29%)	-3.10% (0.48%)	-2.91% (0.30%)	-3.70% (0.46%)	-2.62% (0.47%)	-2.83% (0.15%)
A1FI scenario change in welfare as percent of income	-3.37% (0.36%)	-3.27% (0.37%)	-3.73% (0.63%)	-3.48% (0.38%)	-4.40% (0.58%)	-2.99% (0.61%)	-3.37% (0.18%)

*Notes:* The unit of observation is a PUMA, except for R13, where the unit of observation is a MSA. MWTPs shown for a day at 40F or 80F are relative to 65F, expressed as a fraction of income, and are income-weighted averages across all observations. Except for R11, all regressions are weighted by population. Parenthetical values indicate standard errors cluster bootstrapped on state using 200 draws (except for first two rows, which are not bootstrapped).

\*\* The temperature bins are defined in the ranges (-inf, 35], (35, 45], (45, 55], ..., (75, 85], and (85, +inf).



**Table A1.2: Projected U.S. average amenity value impacts from A1FI climate change for 2090-2100 using preference estimates from the homogenous preference models.**

	I	II	III	IV	V	VI
Temperature function	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline	Restricted cubic spline	Linear spline
<u>Hadley A1FI Scenario</u>						
Change in welfare as percent of income	-2.71% (0.44%)	-3.08% (0.87%)	-3.13% (0.58%)	-2.47% (1.06%)	-1.68% (0.97%)	-1.35% (1.32%)
Component driven by warmer winters	1.98% (0.38%)	2.56% (0.50%)	1.34% (0.47%)	2.46% (0.59%)	2.91% (0.97%)	3.61% (0.96%)
Component driven by hotter summers	-5.26% (0.70%)	-6.24% (1.03%)	-4.90% (0.76%)	-5.37% (1.10%)	-4.98% (1.30%)	-5.42% (1.61%)
Component driven by other weather changes	0.56% (0.12%)	0.60% (0.12%)	0.44% (0.18%)	0.44% (0.17%)	0.39% (0.18%)	0.46% (0.19%)
Percent of population losing	96.8% (1.6%)	96.9% (3.8%)	97.3% (3.1%)	96.0% (15.1%)	86.9% (13.9%)	82.7% (20.1%)
Demographic controls	Y	Y	N	N	Y	Y
State fixed effects	N	N	N	N	Y	Y

*Notes :* Parenthetical values indicate standard errors wild cluster bootstrapped on state using 200 draws.

## Appendix 2: Data Construction

### 1 Overview

Here we document the construction of the dataset used to estimate the models in the paper. Specifically, we collect the following data for each PUMA:

- historical temperature
- other historical weather variables
- climate model data
- demographic controls
- natural controls
- quality of life measure
- MSA assignment

We detail the data assembly process in each of the categories listed above in the following sections.

### 2 Historical temperature

Our source of historical temperature data is the Schlenker-Roberts daily weather data for the contiguous United States,<sup>1</sup> available from 1950–2005. They construct these data by combining high resolution *monthly* average PRISM temperature data (discussed more in detail in section 3) with *daily* temperature data obtained from weather stations, which are at a much lower spatial resolution. They regress the monthly average PRISM temperature data on monthly average weather station data, and then interpolate the daily temperature records at each weather station to obtain daily records at each PRISM grid point. For more details regarding the construction of these data, see their article “Nonlinear Effects of Weather on Corn Yields”.<sup>2</sup>

The data are set on the same fine grid as the PRISM data: 2.5 arcmin  $\times$  2.5 arcmin (roughly 4 km  $\times$  4 km) grid.

#### 2.1 Binning temperature

Starting with the daily temperature data, we calculate, for each gridpoint-year, the number of days with an average temperature (mean of minimum and maximum temperature) falling in each of 222 temperature bins. The number of days in each bin is calculated as follows:

$$\sum_{day=1}^{365} 1 \{avgTemp_{day} \in range(bin)\}$$

We define 220 bins ranging from  $-50^{\circ}\text{C}$  to  $60^{\circ}\text{C}$ , with two more bins for days below  $-50^{\circ}\text{C}$  or above  $60^{\circ}\text{C}$ , for a total of 222 bins; see table A2.1 for more detail.

---

<sup>1</sup> Schlenker and Roberts (2009). “Nonlinear Temperature Effects indicate Severe Damages to U.S. Crop Yields under Climate Change,” Proceedings of the National Academy of Sciences, 106(37), September 15 2009, p.15594-15598.

<sup>2</sup> Schlenker and Roberts. “Nonlinear Effects of Weather on Corn Yields,” *Review of Agricultural Economics*, 28(3), Fall 2006, p. 391–398.

Bin	Temperature Range, °C
bb001	$(-\infty, -50.0)$
bb002	$[-50.0, -49.5)$
bb003	$[-49.5, -49.0)$
...	...
bb100	$[-1.0, -0.5)$
bb101	$[-0.5, 0.0)$
bb102	$[0.0, 0.5)$
bb103	$[0.5, 1.0)$
...	...
bb220	$[59.0, 59.5)$
bb221	$[59.5, 60.0)$
bb222	$[60.0, \infty)$

Table A2.1: Temperature ranges covered by each bin.

Because there are no days with temperatures lower than  $-50$  °C or higher than  $60$  °C, we can simply convert daily average temperature to its corresponding bin using:

$$bin = \text{floor}(2 \times \text{avgTemp}) + 102$$

The raw bins add up to slightly over 365 per year due to leap years, so we rescale the bins to add up to exactly 365 over a year. We perform this gridpoint-year binning for each year between 1970 and 1999, inclusive. As temperature and other weather indicators are generally averaged over 30-year “normals” (more specifically, the PRISM data we use follows the 1970–1999 normals), we generate normals of our binned data by averaging over these same 30 years.

## 2.2 Mapping binned temperature to PUMAs

We map the binned daily temperature data to PUMAs using one of the following two methods, depending on the size of the PUMA.

We use PUMA boundaries available in shapefile format from IPUMS USA.<sup>3</sup> For each PUMA that has at least one PRISM grid point within its boundary, we match the PUMA to the binned temperature data for their gridpoints and average across the PUMA, weighting all grid points equally.

For those few (33) very small PUMAs with that contain no grid points, we match the PUMA to the binned temperature data for the four (or fewer) gridpoints surrounding the PUMA centroid, and take the inverse-distance-squared weighted average, using great-circle distances. PUMA centroids are calculated using ArcMap from PUMA boundaries.

Most PUMAs are matched to four grid points. However, for some coastal PUMAs, the corresponding over-water grid points do not have data, so they are matched to three or fewer gridpoints. The following two tables summarize the mapping from PRISM grid points to PUMAs:

<sup>3</sup> <https://usa.ipums.org/usa/volii/2000pumas.shtml>

Count of PUMAs	
with grids inside	2,024
without grids inside	33

Count of grid points	
avg. per PUMA	228.9
falling inside a PUMA	470,773
not falling inside a PUMA	386
total	471,159

### 3 Other historical weather data

We assemble data on dew point, relative humidity, precipitation, and sunshine at the PUMA level.

Dew point, relative humidity, and precipitation come directly from, or are derived from, PRISM data. PRISM, or the Parameter-elevation Regressions on Independent Slopes Model climate mapping system is a “unique knowledge-based system that uses point measurements of precipitation, temperature, and other climatic factors to produce continuous, digital grid estimates of monthly, yearly, and event-based climatic parameters.”<sup>4</sup> We used PRISM monthly average temperature, precipitation, and dew point data from the years 1970–1999. The data are available on a high-resolution grid of 2.5 arcmin  $\times$  2.5 arcmin (roughly 4 km  $\times$  4 km).

#### 3.1 Dew point

The PRISM data give average dew point in each month for each PRISM grid point. We average over all 30 years to obtain the average dew point, for each grid point, in each calendar month, 1970–1999, and map these gridded data to PUMAs as described in section 2.2.

#### 3.2 Relative humidity

We calculate relative humidity from PRISM temperature and dew point data, following conversions given by the equations below. We first convert temperature  $t$  and dew point  $d$  to degrees Celsius. Actual vapor pressure  $e$  and saturated vapor pressure  $e_s$  are calculated as follows:

$$e = 6.112 \cdot \exp\left(\frac{17.67d}{d + 243.5}\right)$$

$$e_s = 6.112 \cdot \exp\left(\frac{17.67t}{t + 243.5}\right)$$

Then relative humidity is given by:

$$rh = 100 \frac{e}{e_s}$$

One problem is that the equations defining relative humidity as a function of temperature and dew point are nonlinear, so true average relative humidity is likely not exactly the same as relative humidity calculated from average temperature and average dew point. Because dew point and

<sup>4</sup> <http://www.prism.oregonstate.edu/https://usa.ipums.org/usa/volii/2000pumas.shtml>

temperature are not extremely variable within a given month at a given location, this is unlikely to significantly change the relative humidity calculation.

These gridded relative humidity data are then mapped to PUMAs as described in section 2.2.

### 3.3 Precipitation

We import gridded precipitation data from PRISM, which give average precipitation in each month over the periods 1970–1999, for each PRISM grid point. We average to obtain the mean precipitation for each grid point, in each calendar month, 1970–1999. These gridded data are then mapped to PUMAs as described in section 2.2.

### 3.4 Sunshine

Because high resolution gridded sunshine or cloud cover data are not available from PRISM or other sources, we use weather station data from the National Climatic Data Center.<sup>5</sup> This data set contains data on the amount of time the sun is not obscured by clouds, as a percent of possible daylight hours, by month-of-year, for 174 weather stations across the United States; this includes some stations in Alaska, Hawaii, Puerto Rico, and other US territories that we do not use. We use only the 156 stations in the contiguous United States, mapped in figure A2.1. The data set also lists the total number of years of sunshine observations for each station, which ranges from only a few years to over 100. These data do not give the longitude or latitude of each weather station, but rather give us a unique identifier for each station. We look up each station in the *North American Master Station Database*<sup>6</sup> to obtain the longitude and latitude.



Figure A2.1: Weather stations used for construction of sunshine dataset.

We then use inverse distance squared weighting to interpolate the weather station data to the PUMA level. The closest four weather stations, appropriately weighted, are used for the interpolation to each PUMA centroid. We use the great circle distances because the distances

<sup>5</sup> <http://www.ncdc.noaa.gov/oa/climate/online/ccd/pctpos.txt>

<sup>6</sup> <http://www.weathergraphics.com/identifiers/master-station.dat>

between PUMA centroid and nearest four weather stations can be potentially large. Interpolated values of sunshine are not sensitive to the number of surrounding weather stations used in the interpolation. These data are mapped in figure A2.2.

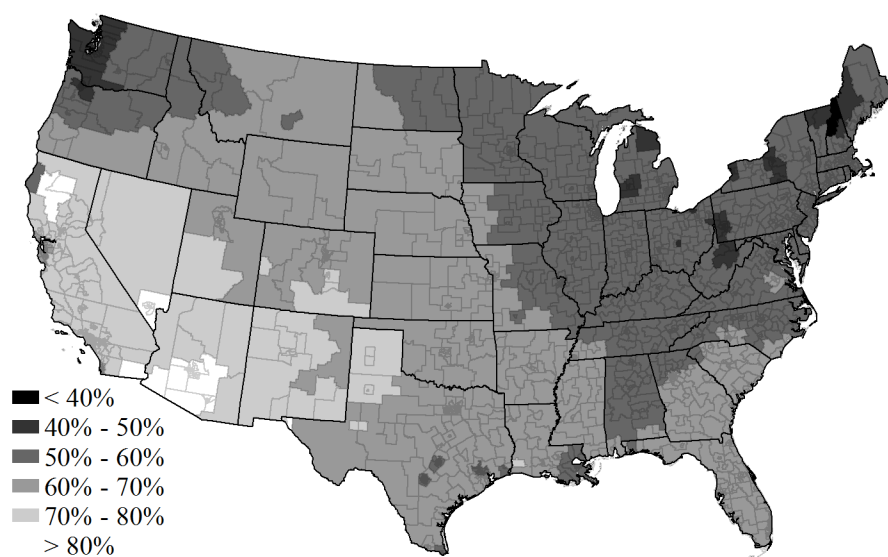


Figure A2.2: Annual sunshine, in percent of possible daylight hours.

## 4 Climate model data

All climate change scenario data are from the third release of the Community Climate System Model (CCSM 3.0), available from the Earth System Grid Gateway at the National Center for Atmospheric Research.<sup>7</sup> These data were also used for the 2007 Fourth Assessment Report of the United Nations Intergovernmental Panel on Climate Change.

We use average daily and monthly data from five runs (a–e) of the *Climate of the 20th Century experiment* from the years 1970–1999, five runs (a–e) of the *A2 scenario* from 2000–2099, and the only available run (a) of the *A1FI scenario* from 2000–2099. The data are available on a 1.4 degree  $\times$  1.4 degree resolution grid.

### 4.1 Daily data

The climate model provides daily data for all variables; we use only daily average surface temperature data (CCSM field name *TS*) to construct binned temperature, similarly to the construction of binned data from the Schlenker-Roberts daily weather data described in section 2. For each decade, we calculate the average number of days that fall in each bin as described in section 2.1.

### 4.2 Monthly data

We use monthly average surface temperature, relative humidity, surface pressure, convective precipitation rate, and large-scale (stable) precipitation rate (field names *TS*, *RELHUM*, *PS*, *PRECC*, and *PRECL*, respectively), calculating decadal averages of each variable in each calendar month. Convective and large-scale precipitation rates are added together to obtain total

<sup>7</sup> <http://earthsystemgrid.org>

precipitation.

### 4.3 Mapping to PUMAs

These monthly average and binned daily average data are then mapped to PUMAs. Because the CCSM grids are significantly bigger than the PRISM grids described in 2, there are many PUMAs that do not have even one CCSM grid point falling within its boundary. Thus, all PUMAs are assigned the inverse-distance-squared weighted value of the four grid points surrounding the centroid of the PUMA, as for the 33 PUMAs interpolated using the alternate method in section 2.2. We use the true great-circle distances, as the distances between PUMA centroid and surrounding grid point can be large, up to 100 km.

We calculate PUMA-level changes in the temperature distribution as the bin-by-bin difference between 1990–1999 *20CE* climate model data and the *A2* or *AIFI* climate model data. However, this yields a strange result for some PUMAs: if the climate model does not predict the actual temperature distribution correctly, it is possible that the projected decrease in the number of days in some bin will be larger than the number of days that we actually observe in that bin originally. This yields a predicted future temperature distribution with a few bins that have a negative number of days.

Fortunately, these errors are small. We calculate the number of bins with a negative value for 2090–2099 predicted temperatures under the *A2* scenario; across all PUMAs, this has a mean of 9.13 bins and a median of 5.05 bins. As a percentage of binned days, the number of negative predicted binned days is likewise small, with a mean of 2.5% and a median of 1.4%. Earlier decades have even fewer bins displaying such errors.

## 5 Demographic controls

We collect various demographic variables, including population, population change, schooling, age, percent married, percent veteran, percent in each minority group, and other variables. we calculate PUMA-level statistics for various variables over all individuals in each PUMA (not just workers). We use data from IPUMS.

Also, we calculate a population-weighted version of population density at the PUMA level, and impute population change from 1990–2000, as described in the following sections.

### 5.1 Population density data

We construct a population density measure that more accurately reflects the actual population density experienced by the population, relative to a simple *population/area* (that is, *area-weighted*) measure over an entire county PUMA. We acquire population and (land) area data by census tract from the 2000 Census.<sup>8</sup> Census tracts are subdivided much more finely than PUMAs; there are generally only between 2000–8000 people in each tract.

---

<sup>8</sup> <http://dataferrett.census.gov/>

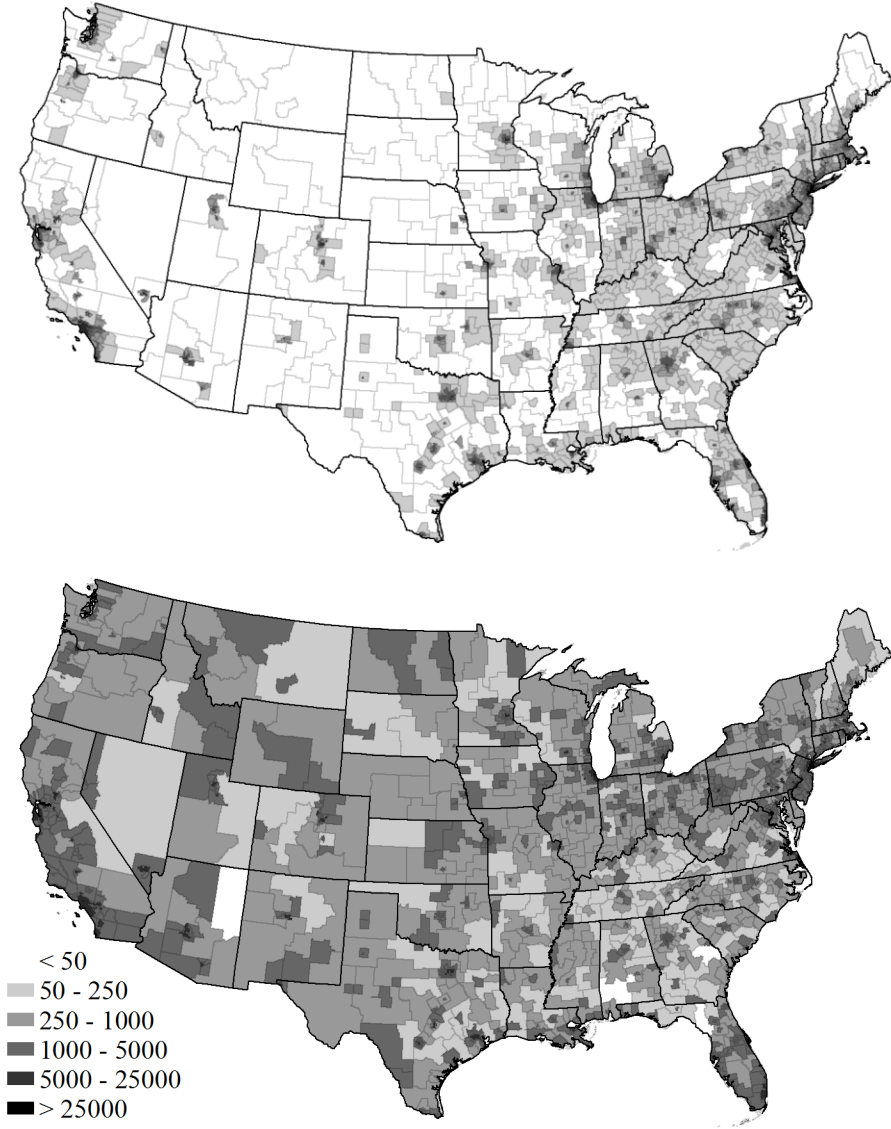


Figure A2.3: Comparison of unweighted and weighted population density by PUMA, in persons per square mile.

Our new population density data on PUMA level are computed by first calculating simple population density ( $population/area$ ) at the census tract level. Because census tracts do not cross county lines, computing this average would be simple for counties. However, census tracts do sometimes cross PUMA boundaries; thus, we use MABLE/Geocorr2K correspondence files to assign the correct areas and populations of the census tracts to each PUMA. Thus, for each PUMA  $p$ , the population weighted density is computed as follows:

$$weightedDensity_p = \frac{\sum_{t \in p} density_t \cdot population_t \cdot fract(t \in p)}{\sum_{t \in p} population_t \cdot fract(t \in p)},$$

where, for each tract  $t$  and each PUMA  $p$ , MABLE/Geocorr2K gives the allocation factor:



$$\text{fract}(t \in p) = \frac{\text{total population in } t \text{ and in } p}{\text{total population in } t}.$$

The weighted and unweighted PUMA densities are mapped in figure A2.3.

## 5.2 Population change data

We are interested in the PUMA-level population change between the 1990 census and 2000 census. However, because PUMA definitions change every census, we cannot directly use the 1990 PUMA population data, as the PUMAs do not line up. So, we impute 1990 population in our 2000 PUMAs using 1990 census tract level data, as follows.

In ArcMap, we intersect the 1990 tract shapefile with the 2000 PUMA shapefile; this splits each tract into one or more pieces, each one lying inside a different PUMA. Then, for each tract slice, we calculate the area, and compute the fraction of the whole tract area lying inside each PUMA. Then, we attribute that same fraction of the 1990 tract population to that PUMA, and add up over tracts to get the whole PUMA population.

The true 1990 population, not including AK and HI, is 247,051,601, while the total population we calculate using this method is 246,956,773. The discrepancy comes from a few hundred tracts that are not included in the PUMA shapefiles, and thus were not attributable to any PUMAs. These are tracts ending .99, which according to the US Census Bureau<sup>9</sup> is reserved for tracts consisting only of “crews-of-vessels populations.” We are unfortunately unable to assign this small sub-population properly to PUMAs, resulting in a missing population on the order of 0.038%.

For a consistency check, we repeat the above procedure to impute 2000 PUMA populations from 2000 census tracts. For the 2000 census, the crews-of-vessels populations are part of the related census tracts, so we do not have the same problem as above; the true 48-state population is 279,583,437, which is exactly the total population we calculated. However, the population we assigned to each PUMA is not exactly correct, and thus will not match the true PUMA populations, as it is imputed using this imperfect area-weighting method.

We approximate population change by taking the difference between imputed PUMA population in 2000 and imputed PUMA population in 1990.

## 6 Natural controls

We construct data on distance to coasts and mountainousness at the PUMA level from raw elevation data and lake and ocean shapefiles, as detailed in the following sections. Specifically, we construct

- average elevation (in meters)
- average slope (in percent)
- distance to ocean (in miles)
- indicator for whether PUMA is on the ocean
- distance to a great lake (in miles)
- indicator for whether PUMA is on a great lake

### 6.1 Elevation data

We use high-resolution elevation data from the *Global 30 Arc Second Elevation Data*

---

<sup>9</sup> [http://www.census.gov/geo/www/cob/tr\\_metadata.html](http://www.census.gov/geo/www/cob/tr_metadata.html)

(GTOPO30) digital elevation model (DEM) available from the the United States Geological Survey.<sup>10</sup> These data are set on a high resolution grid of 30 arcsec  $\times$  30 arcsec, or roughly 1 km  $\times$  1 km. We map the gridded elevation data to PUMAs by averaging the value of all grid points falling within the boundary of each PUMA. Due to the high resolution of the data, every PUMA has at least one grid falling within its boundary. Elevation data are mapped in figure A2.4.

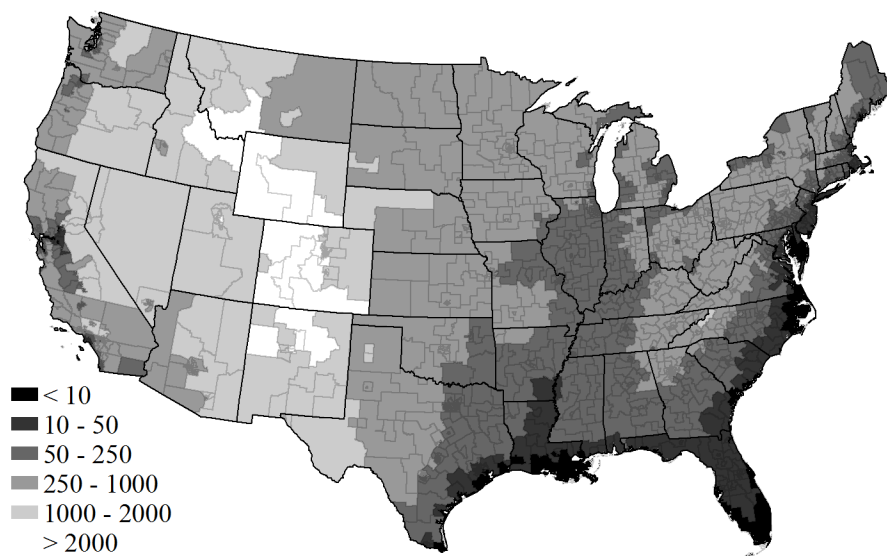


Figure A2.4: Average elevation by PUMA, in meters.

## 6.2 Land slope data

We construct a land slope dataset using the DEMs described in section 6.1. We calculate slope using the average maximum technique, where the slope at each grid point is the maximum rate of change of elevation from that grid point to its eight neighbors.<sup>11</sup> <sup>12</sup> Consider a 3 $\times$ 3 neighborhood of cells, where each cell is identified as follows:

a	b	c
d	e	f
g	h	i

Then the slope at gridpoint e, in radians, is given by:

$$slope = \arctan \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2},$$

where

<sup>10</sup> [http://eros.usgs.gov/#/Find\\_Data/Products\\_and\\_Data\\_Available/gtopo30\\_info](http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30_info)

<sup>11</sup> ArcToolbox help, section "How Slope Works," ESRI.

<http://webhelp.esri.com/arcgisDesktop/9.3/index.cfm?TopicName=How%20Slope%20works>

<sup>12</sup> Burrough, P. A. and McDonell, R.A., 1998. *Principles of Geographical Information Systems* (Oxford University Press, New York), p. 190.

$$\frac{\partial z}{\partial x} = \frac{(c + 2f + i) - (a + 2d + g)}{8 \cdot cellSize_x}$$

$$\frac{\partial z}{\partial y} = \frac{(g + 2h + i) - (a + 2b + c)}{8 \cdot cellSize_y}$$

As in section 6.1 above, we map the gridded slope data to PUMAs by averaging the slope of all grid points falling within the boundary of each PUMA. Slope data are mapped in figure A2.5.

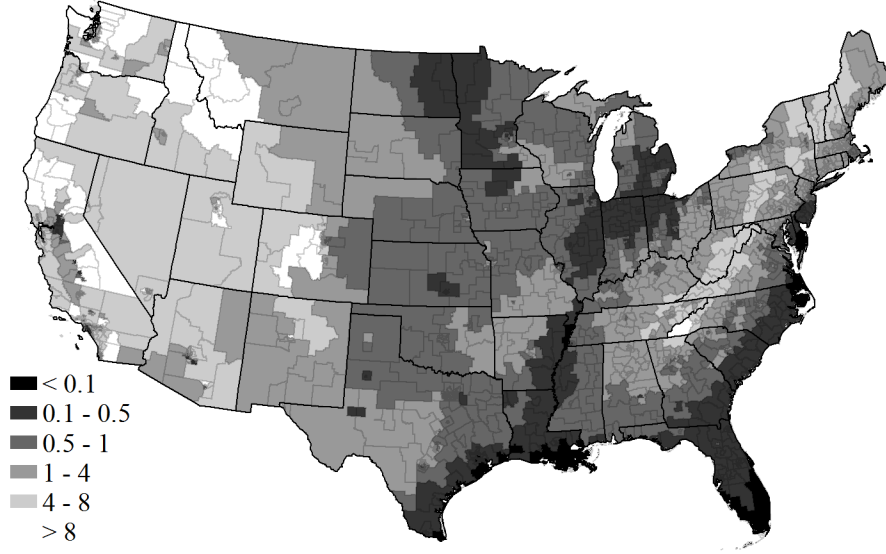


Figure A2.5: Average slope by PUMA, in percent

### 6.3 Coast data

For each geographic level, we construct the following variables:

1. Distance from centroid to oceanic coast
2. Indicator variable: 1 {*onOceanCoast*}
3. Distance from centroid to Great Lakes coast
4. Indicator variable: 1 {*onGreatLakesCoast*}

To obtain the distance variables (1) and (3), we first export the set of points defining the oceanic and Great Lakes from shapefiles available from Esri, creator of ArcMap and the ArcGIS suite. We then calculate the minimum distance from each PUMA centroid to any oceanic or Great Lakes coast point.

Indicator variables for a PUMA being on an oceanic or Great Lakes coast were calculated similarly, with the difference that each PUMA is first assigned the *minimum* distance between any point defining its border and any point on a coast. If this minimum distance is zero (or nearly so, allowing for some small errors in the shapefiles), the coastal indicators are assigned a value of one, and otherwise, a value of zero.

## 7 Quality of life data

Quality of life data at PUMA level come from Albouy and Lue (2011). We use both commuting-adjusted and non-adjusted QOL, data on average QOL and QOL by education level

(high school or college), and price and wage differentials. Construction of these data are detailed in the Albouy and Lue paper.<sup>13</sup>

## **8 MSA assignment**

We use MABLE/Geocorr2K<sup>14</sup> to map each PUMA to its corresponding consolidated metropolitan statistical area (CMSA). Each PUMA is matched to a single CMSA, but CMSAs contain more than one PUMA; the fraction of the CMSA population contained within each PUMA is also given by MABLE/Geocorr2K.

---

<sup>13</sup> Albouy and Lue (2011). "Driving to Opportunity: Local Wages, Commuting, and Sub-Metropolitan Quality of Life." Working Paper.

<sup>14</sup> <http://mcdc2.missouri.edu/websas/geocorr2k.html>

### Appendix 3: Concavity of the utility function

The specifications allowing for unobserved preference heterogeneity use households' first order condition to identify their MWTP for climate amenities at their location, following the approach of Bajari and Benkard (2005) under an assumption that the choice set is continuous. This approach cannot identify, however, the shape of households' utility functions as one moves away from the characteristics associated with their location. Because climate change is non-marginal, this shape affects the estimates of amenity loss associated with climate change. To be conservative, all results presented in the main text of the paper assume that households' utility functions are linear rather than concave in local characteristics.

The linearity assumption may be problematic, however, because it can lead to violations of households' utility maximization should the second order condition not hold. If the hedonic price function is not convex, then households satisfying the first order condition with linear utility functions will be locally minimizing rather than maximizing their utility at any location where the price function is locally concave. Further, even if a particular household is locally maximizing utility at a location where the price function is locally convex, concavities elsewhere in the price function may imply that the household would realize a higher utility at a different location.

In our application, we do find some concavities in the estimated price function, suggesting that violations of utility maximization may be a problem. To assess the empirical importance of this issue, we calculate the amount of utility function concavity necessary, at each location, to satisfy utility maximization. To do so, we use a utility function parameterized per equation (A3.1) below to denote the utility that a household located at  $i$  would obtain from residing at location  $j$ :

$$u_j^i = \beta^i Z_j - \alpha^i \sqrt{\sum_{k=1}^{K'} \left( \frac{Z_{ik} - Z_{jk}}{\sigma_k} \right)^2} - q_j + \xi_j \quad (\text{A3.1})$$

The quadratic term in (A3.1) allows the MWTP for each characteristic  $k$  to decrease as the difference between the amount of  $k$  provided at locations  $i$  and  $j$  increases.<sup>1</sup> All characteristics are scaled by  $\sigma_k$ , their standard deviation over all locations, so that their magnitudes are comparable.<sup>2</sup>

The household's first order condition yields estimates of the  $\beta$  per the local linear regressions discussed in the main text. These regressions also yield estimates of the  $\xi_j$ . Given these estimates, for each location  $i$  we loop through all other locations  $j$  and find the smallest positive value of  $\alpha^i$  such that  $u_i^i \geq u_j^i \forall j$ . The resulting estimates of  $\alpha^i$  denote the minimum concavity necessary so that all households are maximizing their utility by residing at their present location.

In our baseline specification (restricted 7<sup>th</sup> degree cubic spline with full controls, and no fixed effects), the mean value of  $\alpha^i$  obtained via this procedure is 0.0017. When calculating losses from climate change, allowing for this degree of concavity yields an additional welfare loss of 0.59% of income.

---

<sup>1</sup> The quadratic terms only involve characteristics  $1, \dots, K'$  for which preferences are permitted to be heterogeneous. For all models presented in the main text, these characteristics are those corresponding to the basis functions for the linear or cubic splines over the temperature bins.

<sup>2</sup> In a more general specification, the  $\alpha^i$  coefficients on the quadratic term could vary across characteristics. For example, in a model with both HDD and CDD, utility maximization could be achieved with: (i) high concavity on HDD but none on CDD; (ii) high concavity on CDD but none on HDD; or (iii) moderate concavity on both HDD and CDD. For simplicity, we assume that the each household's concavity coefficient  $\alpha^i$  is the same for all characteristics in the model.

## Appendix 4: Set identification of preferences

The estimated heterogeneous preferences presented in the text (and the climate change welfare impacts derived from them) assume point identification, following Bajari and Kahn (2005). That is, at each location  $i$ , the estimates assume that households' MWTP for a given characteristic  $k$  is given by the gradient of the price function at  $i$ ,  $\beta_k^i$ , per households' first order condition. Point identification, however, requires that the set of available locations is complete in characteristics space (itself a subset of  $\mathbb{R}^K$ ), a condition that cannot hold in a finite choice set such as our sample of PUMAs. Thus, in reality, preferences are only set identified. However, if the number of choices (in our data, the 2,057 PUMAs) is sufficiently large, the identified sets may be small enough that estimates assuming point identification are close approximations to the identified set. In particular, the distribution of point identified preferences across PUMAs may closely match the distribution of set identified preferences. This appendix examines the extent to which this approximation holds in this setting.

At each location  $i$ , finding the identified set of preferences requires finding the set of  $\beta^i$ 's such that  $u_i^i \geq u_j^i \forall j$ . We use a utility function that is concave in characteristics and given by equation (A4.1) below (copied from equation (A3.1) in appendix 3):

$$u_j^i = \beta^i Z_j - \alpha^i \sqrt{\sum_{k=1}^{K'} \left( \frac{Z_{ik} - Z_{jk}}{\sigma_k} \right)^2} - q_j + \xi_j \quad (\text{A4.1})$$

An allowance for concavity is required to ensure that the identified set is not null. Given values for the characteristics  $Z$  and quality of life “prices”  $q$ , and given estimates of the unobservables  $\xi_j$  and concavity parameters  $\alpha^i$  from the local linear regressions and the procedure discussed in appendix 3, respectively, equation (A4.1) can be used to find the identified set of  $\beta^i$  at each location.<sup>1</sup> To do so, we use apply the Bayesian estimation procedure

---

<sup>1</sup> As discussed in appendix 3, the  $\alpha^i$  are set to ensure that utility maximization holds exactly at each location, given point identified estimates of the  $\beta^i$ . As evidenced by the fact that the set identified preferences are not singletons, these  $\alpha^i$  do not imply that other values of  $\beta^i$  are inconsistent with utility maximization. Increases in  $\alpha^i$  do imply

discussed in Bajari and Benkard (2005) and described in detail in Benkard and Bajari (2004).<sup>2</sup> This procedure assumes that the distribution of preferences, within and across locations, is point identified and uses Gibbs sampling to obtain this distribution (an analytical characterization of these sets is very difficult because they are high-dimensional).

The Gibbs sampler requires a prior on the preference distribution. Following Benkard and Bajari (2004), we assume that preferences are distributed uniformly over a rectangular region  $\mathcal{B}$ . We must place bounds on this region so that, at locations with extreme characteristics, preferences are not unbounded (for example, at the hottest PUMA in the data, residents could in theory have an unbounded taste for heat). For each characteristic  $k$  over which preferences may be heterogeneous, we bound the set  $\mathcal{B}$  in the  $k$ 'th dimension by taking the support of  $\beta_k$  from the point identified estimates and doubling its width. Thus, for example, if the point identified MWTPs for characteristic  $k$  range from -0.1% to -0.2% of income, our prior for the Gibbs sampler is that preferences for  $k$  are uniformly distributed from -0.05% to -0.25% of income.

Given our prior, the Gibbs algorithm proceeds one location  $i$  at a time. A summary of the algorithm's steps in each iteration is given below. For details, see Benkard and Bajari (2004).

1. At a given iteration  $s$ , start with characteristic  $k = 1$ . Find the upper and lower bounds of  $\beta_k^i$  such that  $u_i^i \geq u_j^i \forall j$ , given the  $\beta_{-k}^i$  from iteration  $s-1$  (use the point-identified  $\beta_{-k}^i$  if  $s = 0$ ). Denote these bounds by  $\bar{\beta}_k^i$  and  $\underline{\beta}_k^i$ . If location  $i$  has the minimum or maximum value of  $X_k$  in the choice set, then the lower or upper bound, respectively, will be given by the relevant bound of  $\mathcal{B}$ .
2. The posterior distribution of  $\beta_k^i$  for iteration  $s$  is uniform on  $[\underline{\beta}_k^i, \bar{\beta}_k^i]$ . Randomly draw a  $\beta_k^i$  from this distribution.

---

increases in the size of the identified set of  $\beta^i$ , however, this sensitivity is not strong. Doubling the values of  $\alpha^i$  from those calculated in appendix 3 yields only a minor change in the kernel density plots discussed below.

<sup>2</sup> Benkard, C. Lanier and Patrick Bajari (2004) "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," NBER working paper 10278.



3. Repeat steps 1 and 2 for each characteristic  $k \in 2, \dots, K'$  ( $K'$  denotes the number of characteristics over which preferences are permitted to be heterogeneous). When finding the bounds  $\bar{\beta}_k^i$  and  $\underline{\beta}_k^i$ , use  $\beta_l^i$  from iteration  $s$  rather than  $s-1$  if  $l < k$ .

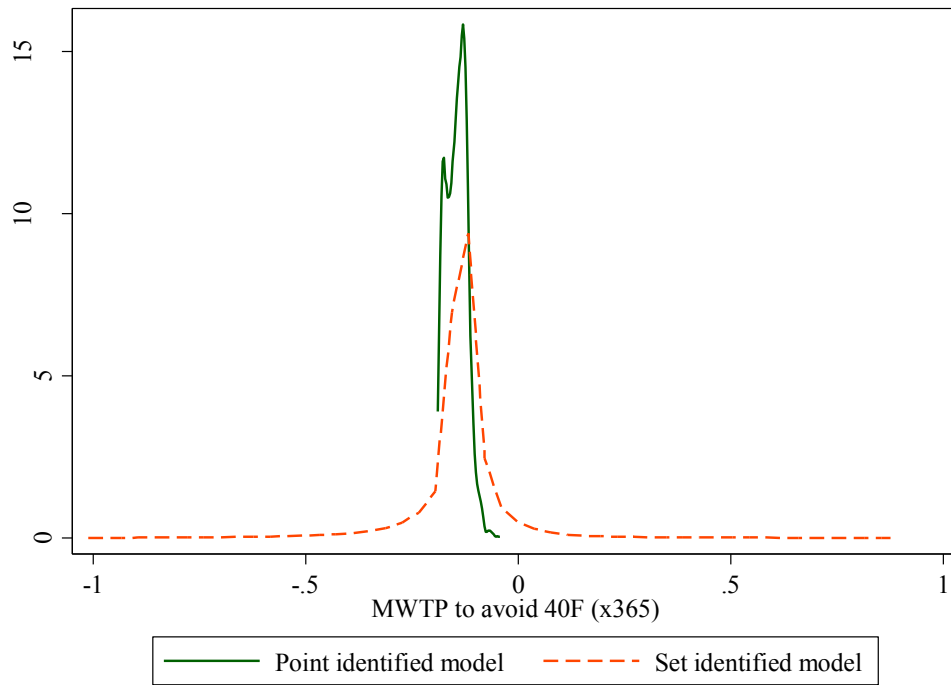
We use 1,500 draws in our algorithm and discard the first 500 as a “burn in” period. The remaining 1,000 draws at each location  $i$  then give a numerical estimate of the distribution of preferences at  $i$ . Aggregating the draws over all locations yields a numerical estimate of the distribution of preferences over the population. Figures A4.1 and A4.2 below show kernel density plots of the MWTPs for a 40 degree day and an 80 degree day (both relative to 65°F) for the 7<sup>th</sup> degree cubic spline specification that includes demographic controls. The primary difference between the point and set identified MWTP estimates is that set identification yields long, thin tails of extreme preferences that are driven by locations at or near the extremes of the characteristics space.<sup>3</sup> Otherwise, these two figures show strong agreement between the preference distributions implied by point identification and set identification, validating the use of the point identified estimates in the main text.

As noted in the text, we only allow for preference heterogeneity over the characteristics associated with temperatures and enforce preference homogeneity for the controls. We do so because the close agreement between the set and point identified preference distributions breaks down when we allow for heterogeneous preferences for the control characteristics. Figures A4.3 and A4.4 illustrate kernel density MWTP plots when preferences for the geographic and “other weather” controls are permitted to be heterogeneous. When preferences for this many characteristics are permitted to vary, the identified sets become too large to allow precise inference.

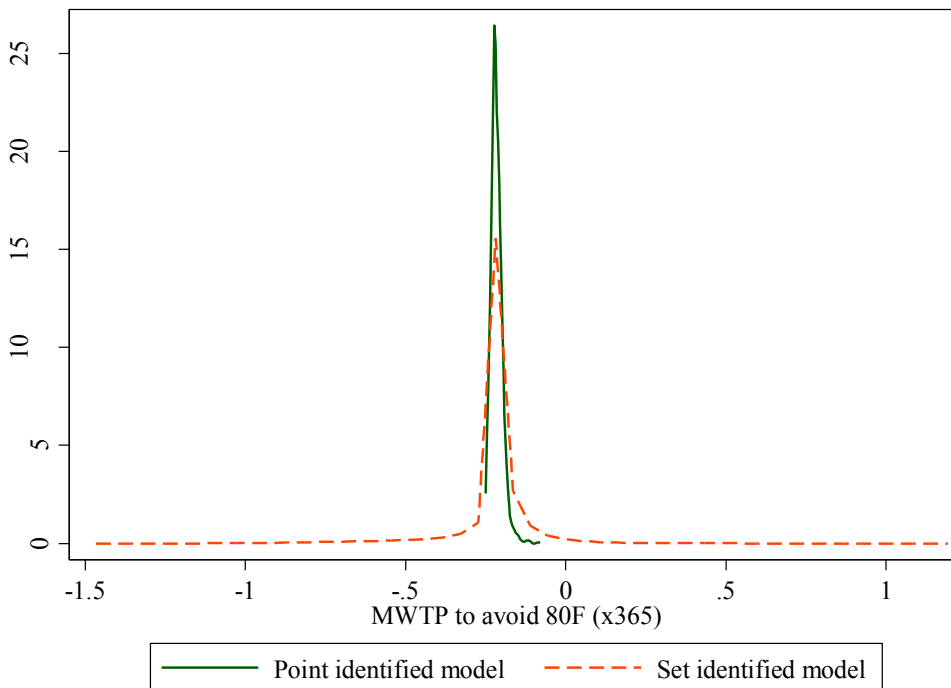
---

<sup>3</sup> Note that, even though the space  $\mathcal{B}$  has bounds given by a doubling of the support of the set of point identified preferences, the set identified MWTPs for a given temperature can have bounds given by more than a doubling of the support of point identified MWTPs. This outcome is due to the fact that the MWTP for a given temperature is given by a linear combination of multiple characteristics (seven of them in the cubic spline models).

**Figure A4.1: Kernel density plots of the MWTP for a 40°F day**

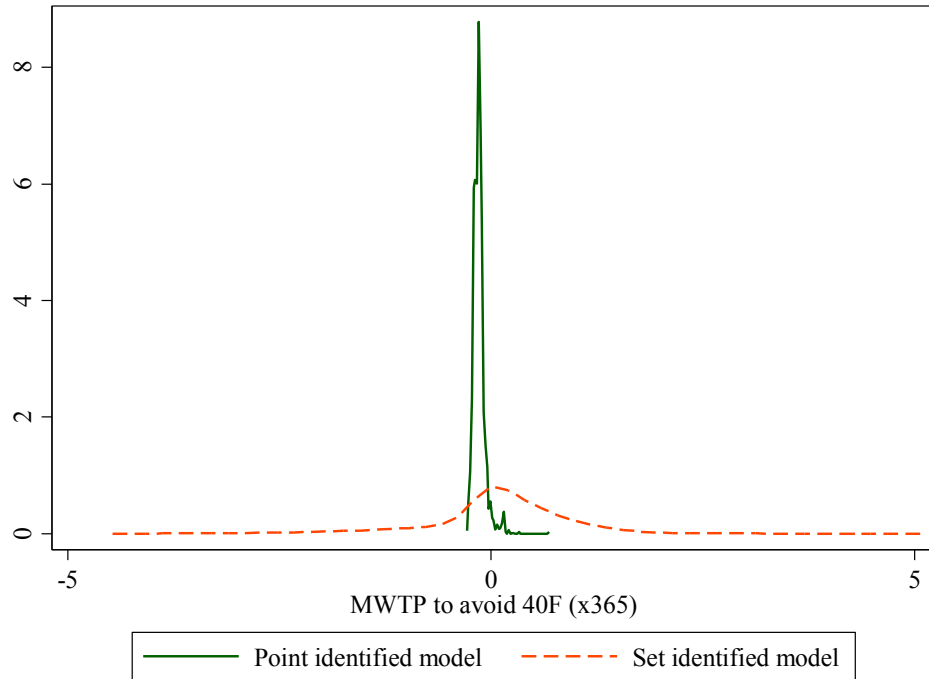


**Figure A4.2: Kernel density plots of the MWTP for an 80°F day**

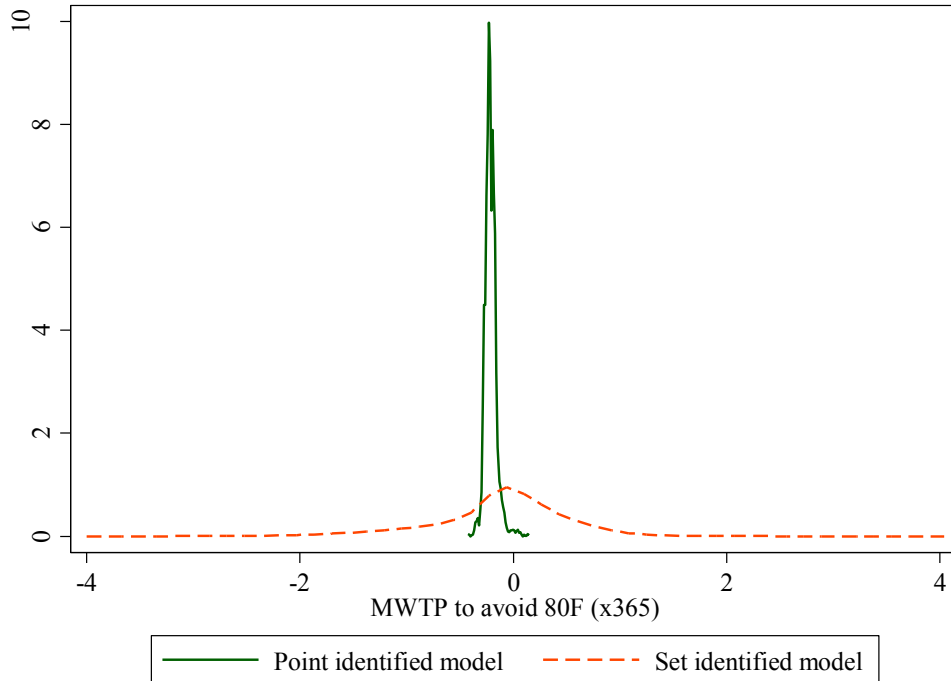


*Notes:* Both figures use the 7<sup>th</sup>-degree restricted cubic spline model with demographic controls. The only random coefficients are those on the temperature basis functions.

**Figure A4.3: Kernel density plots of the MWTP for a 40°F day**  
**Estimates allow preference heterogeneity for geographic and “other weather” controls**



**Figure A4.4: Kernel density plots of the MWTP for an 80°F day**  
**Estimates allow preference heterogeneity for geographic and “other weather” controls**



*Notes:* Both figures use the 7<sup>th</sup>-degree restricted cubic spline model with demographic controls. Random coefficients are used on the temperature basis functions, geographic controls, and other weather controls.