Climate Change Impacts on U.S. Migration and Household Location Choice

Qin Fan[1], Department of Agricultural Economics and Rural Sociology, The Pennsylvania State University, University Park, PA, 16801, USA
H. Allen Klaiber, Department of Agricultural, Environmental and Development Economics, The Ohio State University, Columbus, OH, 43210, USA
Karen Fisher-Vanden, Department of Agricultural Economics and Rural Sociology, The Pennsylvania State University, University Park, PA, 16801, USA

Keywords: Climate Change; Extremes; Tiebout Sorting; Locational Equilibrium; Migration

ABSTRACT

This paper employs a two-stage residential sorting model[2] to examine climate change impacts on residential location choices in the US. The estimated coefficients are used to simulate population changes and US migration patterns across regions under hypothetical changes in climate. The main dataset used for estimation is the Integrated Public Use Microdata Sample (IPUMS), which provides demographic characteristics of approximately 2.4 million households located in 283 Metropolitan Statistical Areas (MSAs) of the US in the year 2000. Projected climate data (i.e. extreme temperatures) used for simulation are obtained from the North American Regional Climate Change Assessment Program (NARCCAP). In the estimation component, a two-stage random utility sorting model (RUM) is employed. The first-stage discrete choice model employs a multinomial logit specification to recover heterogeneous parameters associated with MSA specific variables, migration costs, along with the mean indirect utility of each MSA. In particular, the interaction terms of temperature extremes and individual-specific characteristics, such as one's birth region, age and educational attainment, are used to recover valuations of temperature extremes for different classes of people with potentially different preferences. The second stage of this model decomposes the mean indirect utility obtained from the first stage into its MSA-specific attributes controlling for unobservables using region fixed effects. Migration costs are statistically significant. If migration costs are high, individuals are less likely to relocate for the sake of moderate changes in weather extremes. In the simulation component, the estimated coefficients are used to simulate population changes across regions in the US under hypothetical changes in extreme temperatures. We find that extreme temperature and extreme precipitation reduce utility, and people's preferences for temperature extremes are heterogeneous. The climate of one's place of birth and demographic characteristics such as age and educational attainment, are significant factors that lead to preference heterogeneity. In addition, we find that population share in the Southern region drops, while population share in Northeastern region increases under hypothetical changes in climate.

[1] Qin Fan is the corresponding author at: Department of Agricultural Economics and Rural Sociology, The Pennsylvania State University, 309 Armsby Building, University Park, PA, 16801, USA. Email address: quf101@psu.edu

[2] Sorting model is based on the logic of sorting households into local jurisdictions where they maximize utility and obtain a desired level of public goods.

1. Introduction

The Intergovernmental Panel on Climate Change (IPCC) projected that average surface air temperature has increased by $0.74^{\circ}C$ since 1900 and the sea level will rise by 0.6-1.6m by 2100 (IPCC, 2010). Projected climate change will directly reduce extreme cold days, but increase extreme heat days. Extreme events such as tornado, drought, and flood will occur with a higher probability each year as a result of climate change (IPCC, 2011). It has been recognized that there is a significant economic loss associated with temperature extremes. For example, extreme heat and the natural disasters that result from it (e.g. drought and tornado occurrences) lead to large economic costs in different sectors such as transportation, agriculture, energy, and public health. In contrast, the aggregate effect of extreme cold on public health is found to have higher costs and long-lasting impacts than effects of extreme heat. Deschenes and Moretti's (2007) find that the mortality rate attributable to extreme cold roughly amounts to 1.3% of average annual deaths in the U.S. over their sample period, while an increase in mortality rate attributable to extreme heat is much lower and the impact is short-lived.

Impacts of changes in weather extremes, such as extreme temperatures and extreme events have not been well examined in previous literature. For example, most previous studies examine climate change impacts on location choice in terms of mean temperature (e.g. Timmins, 2007), but few studies estimate people's valuation of climate change in terms of weather extremes. The empirical results on people's valuations of climate change in terms of weather extremes can provide evidence for analyzing the cost effectiveness of relevant climate change policies, particularly those aimed at reducing economic costs from the negative impacts of climate extremes. More efforts are needed, therefore, to study the impacts of temperature extremes that reflect climate variability and extreme events that have low-probability but can cause substantially large damages.

Heterogeneity in regional impacts is a key component in studying the effect of weather extremes on residential location choices, since climate change impacts are heterogeneous across both regions and individuals. Warm regions in the U.S. may be negatively affected by an increase in extreme heat days under climate change, while cold regions may benefit from reduced extreme cold days. Factors such as different climates of individuals' birth places, one's age and mobility choices may lead to preference heterogeneity. For example, people born in cold regions are potentially more sensitive to extreme heat, while those born in hot regions are potentially more sensitive to extreme cold. Older individuals after retirement may relocate for the sake of nice amenity and pleasant weather, and it is possible that they are more sensitive to temperature extremes than young people. Highly educated people (e.g. college graduates) are more mobile, and they have more options to move than those without college degrees. Changes in temperature extremes may have a greater impact on highly mobile people.

To better address these issues, this paper presents an analysis on how climate change affects where people choose to live in terms of weather extremes. In this paper, we allow for preference heterogeneity across individuals focusing on factors such as the climate of one's birth place, an individual's age and education level. This paper employs an empirical Tiebout sorting model that has been widely used to analyze the demand for public goods across space. The equilibrium sorting model used in this paper models the way households sort into local jurisdictions to maximize utility and obtain an optimal level of local public goods given prices and the location choices of other households. There are two main types of sorting models: (1) pure characteristics, which requires all households to have the same ordering of preference across locations (homogenous preference within communities with the same ordering of preference); and (2) random utility sorting (RUM), which allows preferences for attributes to vary distinctly across households. We employ the latter model in this paper as we believe preference heterogeneity is likely to be important in understanding the impacts of climate change. For a further discussion of sorting model, see Kuminoff (2009).

In order to understand the relationship between climate change impacts, migration, and household location choice, this paper incorporates migration costs while examining the tradeoff between the gains from local amenities and the loss in real income associated with migration. After incorporating migration costs, the true value of climate amenities is expected to be higher than what has been shown in the case where free mobility is assumed. Intuitively, if migration costs are high, people are not willing to migrate for the sake of a moderate change in amenable climate. An individual's valuation of climate (e.g. willingness to pay to reduce frequency of temperature extremes and number of tornado watches) must be higher when migration is costly in order to give individuals more incentive to move. In this sense, the results from conventional hedonic model with free mobility may be misleading when migration costs are significantly high. In addition, we simulate population changes across five regions in the US under changes in extreme temperatures projected in the year 2065, based on estimated coefficients and projected temperatures. We find that population share in the Northeastern region increases as extreme cold days decrease under climate change.

This paper tests the hypothesis that changes in climate extremes (i.e. extreme temperatures, extreme precipitation, and tornado frequencies) negatively affect an individual's location choice on where to live. We also estimate the magnitude of these impacts by allowing for preference heterogeneity and migration costs. Changes in population shares across regions in the US are predicted under changes in extreme temperatures.

Results suggest that climate change in terms of extremes have negative impacts on household location choice. In addition, we find that individuals' preferences are heterogeneous. People born in relatively cold regions (e.g. Northeast and West) are more sensitive to extreme heat than people born in warmer regions (e.g. South), while those born in California are more sensitive to extreme cold than people born in other regions. Besides the climate of one's birth place, demographic characteristics also contribute to preference heterogeneity. People over 65 years old after retirement generally favor pleasant amenity, and therefore are more sensitive to extreme temperatures than younger people. Weather extremes have larger impacts on the location decisions of individuals with higher education levels (i.e. college graduates). One reason might be that college graduates may have more options to move and are therefore more mobile than those without college degrees.

2. Literature Review

The traditional framework of non-market valuation has its roots in the early theoretical papers, which estimate marginal valuation without considering spatial relationships (e.g. Rosen, 1974). In the 1990s, Anselin (1988) incorporated spatial effects (e.g. spatial dependence, spatial autocorrelation, and spatial heterogeneity) into the hedonic model. Although the first-stage hedonic model that estimates marginal willingness to pay (MWTP) for public goods has been widely used and spatial effects are captured to some extent in the hedonic framework (Brown, 1980; Smith, 1985; Irwin, 2002), there are several limitations. Since the first stage of the hedonic model estimates an aggregate preference instead of each individual household, it is impossible to estimate the difference in valuations across households. Besides that, there is a strong assumption in the hedonic model that mobility is costless, which is not the case in reality. In addition, there are econometric challenges to identify demand functions in the second-stage hedonic model. Therefore, it is difficult to estimate non-marginal valuation through the hedonic model.

Residential sorting models, which were developed over recent years based on the logic of Tiebot sorting, have the potential to overcome several of the limitations discussed above (Epple, et. al. 2001; Walsh, 2006; Timmins, 2007; Bayer et. al. 2009). The Tiebout sorting model assumes that households sort into local jurisdictions where they maximize utility based on housing property, utility characteristics, and local attributes. Empirically, this model is often categorized into pure characteristics and random utility model (RUM). The former assumes that all households have the same ordering of communities, while the latter allows household preferences to vary distinctly over each household and space (Klaiber, 2010). Therefore, the horizontal sorting model may be preferred when preference heterogeneity and potentially different rankings of commodities are desired. Besides the advantages in capturing preference heterogeneity, the RUM sorting model can relax the assumption of free mobility and can incorporate migration variables that are left out of hedonic models. To allow for migration costs, Bayer et al. (2009) use a sorting model to estimate MWTP for air quality by using dummy variables that indicate whether an individual moves out of one's birth place. In this paper, we use a RUM model that incorporates heterogeneous preferences towards changes in climate by allowing for migration costs. In terms of estimating non-marginal value, sorting models can simulate the welfare effects of non-marginal changes in attributes, which is challenging in the hedonic framework (Timmins, 2007).

Another important motivation of our research is that most previous studies examine climate change impacts on location choices in terms of mean temperature and mean precipitation (Timmins, 2007). Although there are some studies that examine impacts of climate change in terms of weather extremes on agricultural output (Deschenes and Greenstone, 2007) and public health (Deschenes and Moretti, 2007), there are few studies that examine impacts of climate extremes on migration and household location choice. The study conducted by Poston et al. (2009) is one of the few examples. In this study, authors examine the effects of climate on three migration variables (in-migration, out-migration, and net-migration) by incorporating eleven climate variables including extreme heat days and extreme cold days. They use factor analysis to define a new variable TEMPERATURE as a climate factor, which accounts for the variance in these eleven correlated climate variables. They find that this climate factor is positively correlated with in-migration and net-migration rates and is negatively correlated with out-migration rate. This study, however, does not consider preference heterogeneity and migration costs in migration decisions. Ignoring this heterogeneity may lead to incomplete or at worst invalid inference. Huhtala (2000) use both the parametric and non-parametric methods to verify the importance of incorporating heterogeneity in valuation analyses of public goods. He finds that ignoring heterogeneity leads to a biased WTP estimates, and the bias tends to be significantly large in parametric estimation. In our paper, we not only examine climate change impacts on household location choice in terms of weather extremes, but also consider preference heterogeneity that is critical for assessing the potential responses of different groups of people to changes in extreme temperatures.

3. Theoretical Model

A two-stage random utility sorting model is used to estimate the valuation of weather extremes controlling for migration costs. A sorting model captures the process by which households sort into different jurisdictions as they seek to maximize utility and obtain an optimal level of public goods. The first-stage discrete choice model employs a multinomial logit specification to recover heterogeneous parameters associated with MSA specific variables, migration costs, along with the mean indirect utility of each MSA common across households. In particular, the interaction terms of temperature extremes and individual-specific characteristics, such as one's birth region, age and educational attainment are used to recover valuations of temperature extremes for different classes of people with potentially different preferences. The economic variable (i.e. service wage rate) is interacted with one's educational attainment (i.e. college degree) to examine the preference difference towards service wage rates between college graduates and those without college degrees. A dummy variable that indicates whether an individual migrates out of his/her birth region is used to recover long-term psychological costs of moving away from family roots. Immigrants are excluded in this study. MSA fixed effects are incorporated in this stage to recover the mean indirect utility—quality of life—for each MSA. In the second stage, we decompose the mean indirect utility recovered from the first stage into MSA specific attributes, such as economic activities, entertainment, natural amenities, and climate extremes including temperature extremes, precipitation extreme, number of tornado watches, and so on.

Following the methodology of Bayer et al. (2009), we use a simple version of this model to develop our theoretical framework. The head of the household i is assumed to be the decision maker who chooses a specific location j to live along with the consumption of utility characteristics and housing property. Each location j is characterized by local attributes such as economic activities, entertainment, natural amenities, and climate. Each decision maker chooses location j to maximize utility subject to a linear in income budget constraint. When migration costs are incorporated, there is an additional term entered into the utility function that includes psychological costs of moving away from one's place of birth. People move to a location where they achieve maximum utility and a desired level of public goods. A locational equilibrium is achieved if nobody has an incentive to move given prices and the location decision of all others. The general function is defined as:

$$\max_{\{C_i, H_i, X_j\}} U(C_i, H_i, Z_j, M_{ij}) \; s.t. \; C_i + \rho_j H_i = I_{ij} \qquad (1)$$

where C_i represents commodity demanded by individual i, H_i represents the quantity of housing services demanded by individual i, Z_j represents MSA specific attributes, M_{ij} represents whether a specific location j is out of individual i's place of birth, ρ_j represents housing price index for each location j. I_{ij} is an individual i' income in location j which we predict using an income regression described in the Section 5.2.

4. Data

The main dataset used for the empirical analysis is obtained from Integrated Public Use Microdata Sample (IPUMS), which comprise a 5% microdata sample from the 2000 US Population Census. There were 2,417,253 households who lived in the 283 Metropolitan Statistical Areas (MSAs) of the U.S. in this sample. Assuming the head of household is the decision maker, we focus on his/her demographic factors. The main dataset contains housing attributes (Appendix A) and demographic characteristics of head of household (Appendix B). The IPUMS dataset also provides information on the birth state of each head of household, which allows us to create a migration dummy variable that indicates whether location j is out of the head of household i's birth region. The dataset is used in the first-stage sorting model, which requires a two-dimension matrix for each variable: the row dimension has 2,417,253 observations that represent households, while the column dimension has 283 observations that represent MSAs.

MSA-specific amenity and disamenity data that are used in the second-stage sorting model are obtained from a variety of sources. We have 283 observations, one for each of the 283 MSAs. Wage rates by sector are obtained from the U.S. Bureau of Labor Statistics. Total establishments of arts, entertainment and recreation, and water area at the MSA level are obtained from the U.S. Census. (Descriptive statistics are presented in Appendix E). Climate data that includes snowfall, and number of tornado watches are acquired from National Climate Data Center (NCDC). In particular, downscaled temperature and precipitation data (1/8 degree spatial resolution) are used to calculate extreme heat days (annual number of days with daily maximum temperature above 90F), extreme cold days (annual number of days with daily minimum temperature below 32F), and extreme precipitation day (annual number of days with daily maximum precipitation over 1 inch) (Maurer et al., 2002)[3]. We use ArcGIS to intersect gridded data with each MSA, and calculate the arithmetic mean value of exceedance days for each MSA (the map is shown in Appendix D).

The projected temperature data is obtained from North American Regional Climate Change Assessment Program (NARCCAP). We obtain the projected data from runs of the Canadian Regional Climate Model (CRCM), and we count the mean extreme days in the projected 5-year period (2061-2065). Both extreme heat days (mean annual number of days with daily maximum temperature above 90F) and extreme cold days

[3] Gridded data on temperature and precipitation extremes were provided by Rob Nicolas from Department of GeoSicence at the Pennsylvania State University.

(mean annual number of days with daily minimum temperature below 32F) are counted from daily maximum and minimum temperature data. This temperature data (1/2 degree spatial resolution) is interacted with polygons that represent MSAs on the ArcGIS map. The arithmetic mean values of the projected extreme temperatures for each MSA are calculated. We divide the U.S. into five regions (i.e. California, South, Northeast, Midwest, and West). The division of these five regions matches economic regions from the U.S. Census with the U.S. Department of Agriculture (USDA) Plant Hardiness Zones, which are directly connected to different climates (Appendix F). Summary statistics that describe the projected extreme data by region are shown in Appendix F.

5. Empirical Model

5.1 Two-Stage Sorting Model

We follow the model framework by Bayer et. al. (2009), and add the interaction terms of each individual's characteristics and weather extremes (both extreme heat days and extreme cold days), along with the interaction term of college graduates and MSA-specific service wage rate in the utility function. The utility function for household i in location j is defined as:

$$U_{ij} = C_i^{\beta_c} H_i^{\beta_h} Z_j^{\beta_x} e^{\sum_{q=1}^{Q} \beta_q^T (HH_q^i \times T^j) + \beta_{edu}^w (EDU_i \times W_j) + \beta_m M_{ij} + \xi_j + \eta_{ij}}$$

$$\text{s.t. budget constraint: } C_i + \rho_j H_i = I_{ij} \qquad (2)$$

where C_i represents the numeraire good consumed by an individual i, H_i represents the quantity of housing services, Z_j denotes local attributes including economic activities, entertainment, natural amenities, extreme temperatures, extreme precipitation, and number of tornado watches; HH_q^i represents demographic factors of the head of household i, and q represents different types—birth region, age, and educational attainment. T_j includes both extreme cold days (annual number of days with minimum daily temperature below 32F) and extreme hot days (annual number of days with maximum daily temperature above 90F) in a specific MSA j; EDU_i represents whether the head of household i is college graduates; w_j represents service wage rate in MSA j; M_{ij} is a dummy variable which indicates whether a specific MSA is out of one's birth region. Five regions are defined as shown in Appendix F. ξ_j captures the MSA-specific unobservables; η_{ij} represents an individual-specific idiosyncratic component of utility that is assumed to be independent of mobility costs and MSA-specific characteristics. We assume that this idiosyncratic error term is independently and identically distributed type I extreme value, and the multinomial logit model is used in the first stage of our model.

In appendix G, we derive both the first-stage and second-stage equations along with the calculation of the coefficient of housing price index. Equation (3) is the linear in log random utility model (RUM) derived for the first-stage sorting model. (Also see equation G.7 in appendix G):

$$\ln U_{ij} = \ln V_{ij} + \eta_{ij} = \beta_I \ln \hat{I}_{ij} + \sum_{q=1}^{Q} \beta_q^T (HH_q^i \times T^j) + \beta_{edu}^w (EDU_i \times W_j) + \beta_m M_{ij} + \Theta_j + \eta_{ij} \tag{3}$$

where $\Theta_j = -\beta_h \ln \rho_j + \beta_x \ln Z_j + \beta_c CLIMATE_j + \xi_j$ (4)

where $\hat{I}_{ij}$ is the predicted income for each household head i possibly living in each of the 283 MSAs j; The details on how to obtain $\hat{I}_{ij}$ is shown in the section 5.2; Θ_j is the MSA fixed effects (i.e. coefficients of alternative specific constants), which is interpreted as the mean indirect utility for each MSA; η_{ij} is the idiosyncratic error term. Previous studies have demonstrated the importance of including alternative (location) specific constants to recover mean indirect utility that captures unobservables (Bayer et al., 2009; Bayer and Timmins, 2007; Klaiber and Phaneuf, 2010). An inclusion of unobservables controls for location-specific omitted variables. It also contributes to a clean identification of heterogeneous parameters in the first stage of sorting model. This term Θ_j includes both MSA-specific observables and unobservable ξ_j as shown in equation (4). These observables include housing price index ρ_j, MSA-specific attributes Z_j (e.g. wage rates by sector, entertainment, and natural amenity, etc.), and climate extremes $CLIMATE_j$ that include extreme heat days, extreme cold days, extreme precipitation, and number of tornado watches. Other variables are the same as those listed below equation (2).

According to the logic of the sorting model, individuals choose their locations where they maximize utility defined in equation (3). Assuming the idiosyncratic error term η_{ij} is independently and identically distributed (IID) type I extreme value, a multinomial logit specification is used to calculate the probability that household i chooses location j. The probability of choosing location j by individual i is:

$$P(\ln V_{ij} > \ln V_{ik} \;\; \forall j \neq k) = \frac{e^{\ln V_{ij}}}{\sum_{q=1}^{J} e^{\ln V_{iq}}} \tag{5}$$

The first-stage sorting model is estimated by maximizing the log likelihood function:

$$ll = \sum_j \sum_i Y_{ij} \ln(P(\ln V_{ij} > \ln V_{ik} \;\; \forall j \neq k)) \tag{6}$$

We recover MSA fixed effects (the coefficients of MSA specific constants) in the first stage sorting model. The MSA fixed effects can be interpreted as the mean indirect utility of residing in each MSA. From equation G.9 in appendix G, we derive the second stage sorting model, which is also shown in equation (4). As we believe the importance of MSA-specific unobservables that are observed by decision makers, but are not observed by researches (e.g. economic activity and high infrastructure), housing price index for each MSA is endogenous as it is likely to be correlated with these unobservables entered into the error term ξ_j. Following the methodology of Bayer et al. (2009), we move the housing price index ρ_j to the left hand side and include it in the dependent variable.

$$\hat{\Theta}_j + \beta_h \ln \rho_j = \beta_x \ln Z_j + \beta_c CLIMATE_j + \xi_j \tag{7}$$

where $\hat{\Theta}_j$ is the coefficient of MSA specific constants obtained from the first-stage sorting model; ξ_j represents the MSA-specific unobservables that are omitted and are included in the error term; other variables are listed below equation (4).

In the second stage sorting model, this indirect utility is decomposed into its MSA-specific attributes including climate extremes that are of our interest, and unobservables ξ_j.

5.2. Predicted Income

Bayer et al. (2009) argue that income estimation should be used to generate predicted income $\hat{I}_{ij}$ that is included in the first-stage sorting model as shown in equation (3). This is because a person's income is likely to vary across location. In order to obtain income for every individual possibly living in each of the 283 MSAs (i.e. $\hat{I}_{ij}$) rather than the observed income of each individual living in his/her residential location (i.e. I_i), we need to use income regression to predict income $\hat{I}_{ij}$.

Following the methodology of Bayer et al. (2009), we estimate the following income equation:

$$\begin{aligned} \ln I_i = {} & \alpha_j + \alpha_{WHITE,j} \cdot WHITE_i + \alpha_{MALE,j} \cdot MALE_i + \alpha_{AGE>60} \cdot AGE > 60_i \\ & + \alpha_{HSDROP,j} \cdot HSDROP_i + \alpha_{HSGRADi} HSGRAD_i + \alpha_{SOMECOLL,j} SOMECOLL_i \\ & + \alpha_{COLLGRAD,j} COLLGRAD_i + \varepsilon_{ij} \end{aligned} \tag{8}$$

where I_i represents income of the each household decision maker, $WHITE_i$ represents whether the head of the household is white or not (white = 1, non-white = 0), $MALE_i$ represents the gender of the household decision maker (MALE = 1, FEMALE = 0), $AGE > 60$ represents whether the head of the household is older than 60 years old, $HSDROP_i$ represents education level--high school dropout, $HSGRAD_i$ represents high school graduate, $SOMECOLL_i$ represents college degree (less than four years), $COLLGRAD_i$ represents college graduate (four years or more). $HSDROP_i$ is left out and is included in the constant term in the regression. (See Table B.2 in Appendix B).

Regression results from Table B.2 in Appendix B show that people less than 60 years old earn more than those over 60 years old. Males earn more than females. Whites have relatively higher incomes. People with higher education levels have higher incomes. This regression is used to predict an average income in each location for each individual in our sample. The mean value of predicted income is approximately $45,071. This estimated income is close to median income from the U.S. Census in the year 2000, where the median income of female household decision maker (no husband present) is $28,116, and male household decision maker (no wife present) is $42,129 (DeNavas-Walt et al., 2000).

5.3. Housing price index

A hedonic housing price model is used to obtain the housing price index (denoted as ρ_j) for each MSA that is included in the second stage sorting model as shown in equation (7). The hedonic housing price model is defined as:

$$\ln P_{ij} = \ln \rho_j + \beta_{ij} X_{ij} + e_{ij} \tag{9}$$

where P_{ij} is the housing price (only houses that are owned); X_{ij} are housing attributes (Table A.1 in Appendix A); j represents each of the 283 metropolitan statistical areas (MSAs) of the U.S.; $\ln \rho_j$ is an MSA fixed effects. We control for a bundle of housing attributes, such as the acreage of the house, the number of rooms of the property, the year then the house was built, etc.

These MSA fixed effects provide a consistent measurement of the estimated price of a homogeneous unit of housing services in a particular MSA, which serves as a housing price index for each MSA. The housing price index for each MSA is obtained through the hedonic housing price regression. By netting out the implicit values of housing attributes, housing price indices are comparable across MSAs. We take the exponential of the MSA fixed-effects from the results shown in Table A.2 of appendix A, and obtain the mean housing price index for each MSA, which is approximately $16,531. The scattered graph in Figure A (Appendix A) shows that California has a relatively high price index, which is consistent with our expectation.

5.4. Predictions of Population Changes

We use extreme temperature data (both extreme heat days and extreme cold days), respectively, in the base year 2000 and the projected 5-year period (2061-2065) to predict population changes between the year 2000 and 2065 under changes in climate. Due to the instability of a single-year projected data, we use the mean of five-year projected data from the year 2061 to 2065 instead of a single-year projected data. The projected data is from runs of the Canadian Regional Climate Model (CRCM), which is consistent with IPCC business-as-usual A2 scenario. The following probability equation based on multinomial logit specification is used to predict changes in population shares across regions under changes in extreme temperatures:

$$P_{ijt}(\ln V_{ijt} > \ln V_{ikt}\ \forall j \neq k) = \frac{e^{\beta_I \ln \hat{I}_{ijt} + \sum_{q=1}^{Q} \beta_q^T (HH_{qt}^i \times T_t^j) + \beta^{EduW}(EDU_t^i \times W_t^j) + \beta_m M_{ijt} + \Theta_{jt}}}{\sum_{l=1}^{J} e^{\beta_I \ln \hat{I}_{ilt} + \sum_{q=1}^{Q} \beta_q^T (HH_{qt}^i \times T_t^l) + \beta^{EduW}(EDU_t^i \times W_t^l) \beta_m M_{ilt} + \Theta_{lt}}} \tag{10}$$

$$\text{where } \Theta_{jt} = -\beta_h \ln \rho_{jt} + \beta_z \ln Z_{jt} + \beta_t CLIMATE_{jt} + \xi_{jt}$$

where i represents household i, j represents MSA j, t respectively represents respectively the starting point where t = 2000, and the ending point where t = 2065. T_t^j represents both extreme heat days and extreme cold days in MSA j. Other variables are the same as those described below equations (3) and (4).

In the simulation, housing price index ρ_{jt}, income measure $\hat{I}_{ijt}$, and wage rates W_t^j are assumed to change exogenously with a fixed yearly increase rate 2% (Maurer, 2008). We assume that new generation replaces the old generation, and demographic components in 2065 stay the same with those in the year 2000. In my future research, I will endogenize labor supply and wage rates in a computable general equilibrium (CGE) model. In addition, a few assumptions will be relaxed by using sensitivity analysis of changing educational attainment EDU_t^i and migration costs M_{ijt}.

The probability of choosing MSA j is aggregated to regional level—Northeast, Midwest, South, West, and California by adding up the weighted probabilities of choosing MSA j that belongs to region r.

$$P_{rt} = \sum_{j \in r}\left(weight_{jt} \times P_{jt}\right) = \sum_{j \in r}\left(\frac{pop_{jt}}{pop_{rt}} \times P_{jt}\right) = \sum_{j \in r}\left(\frac{pop_{jt}}{pop_{rt}} \times \frac{1}{N}\sum_{i=1}^{N} P_{ijt}\right) \quad (11)$$

where r represents one of the five regions in the U.S.; j represents one of the 283 MSAs; t respectively represents starting point in the year 2000 and ending point in the year 2065; P_r is the probability that region r is chosen; P_j is the probability that MSA j is chosen; P_{ijt} is the probability that the head of household i chooses MSA j as shown in equation (10); N is total number of individuals in the data sample; $weight_{jt} = \frac{pop_{jt}}{pop_{rt}}$, which represents the weight of each MSA j within region r based on population size in the year t; pop_{jt} is the total population in MSA j in the time period t, and pop_{rt} is the total population in region r in the time period t.

6. Empirical Results

6.1. Results from Two-Stage Sorting Model

Table 1 shows the parameter estimates from the first stage sorting model. Marginal utility of income is 1.00. This coefficient is used to calculate the coefficient of housing price index ρ_j. (See equation G.11 in Appendix G). Results from the same table show that people over 65 years old are more averse to extreme temperatures than younger people. College graduates are expected to be more mobile and have more options to move than people without college degrees. Highly mobile individuals are the more averse to temperature extremes than people that are less mobile. People born in cold regions (e.g. Northeast) are more sensitive to extreme heat than those born in the warm regions (e.g. South), while those born in California are more sensitive to extreme cold than people born in other regions. One reason may be that people find the weather that is similar to their hometowns more amenable. The migration dummy variable that indicates whether location j is out of an individual i's region is significant. The coefficient of this variable recovers migration costs in terms of utility. Specifically, there is a significant utility cost associated with leaving one's birth region, which is -2.0926. The mean indirect utility recovered from the 1st stage sorting model in terms of the coefficients of MSA specific constants are displayed in the scatter plot in Appendix C (selected MSAs). The mean indirect utility of residing in Los Angeles ranks top one, which indicates that quality of life in Los Angeles ranks the highest, and this utility comprises all of the MSA-specific attributes in Los Angeles that are common to all households.

Table 1: Parameter Estimates from First-Stage Sorting Model

Dependent variable: location choice (1 or 0) (multinomial logit)

Variable	Variable Description	Coefficient
Ln(predicted income)	Marginal Utility of Income	1.0000*** (0.0053)
Collgrad*Service_wage	College graduates*service wage	1.4378*** (0.0144)
M_Macro_Region	Migration dummy variable which indicates whether a specific MSA j is out of an individual i's birth macro region	-2.0926*** (0.0016)
Age_65_Hot	Age dummy variable which indicates whether a household head i is older than 65 years old (1 if >=65, 0 if <65)* Extreme Hot (mean number of days with maximum temp 90 degrees F or more/10)	-0.0076*** (0.0005)
Age_65_Cold	Age dummy variable which indicates whether a household head i is older than 65 years old (1 if >=65, 0 if <65)* Extreme Cold (mean number of days with maximum temp 32 degrees F or less/10)	-0.0316*** (0.0004)
Collgrad_Hot	Education dummy variable which indicates whether a household head i has four-year college degree or above (1 if college graduates, 0 otherwise)*Extreme Hot	-0.0268*** (0.0007)
Collgrad_Cold	Education dummy variable*Extreme Cold	-0.0305*** (0.0005)
Northeast*Hot	Whether a household head i was born in the Northeast macro-region (1 if yes)*Extreme Hot	-0.0286*** (0.0004)
South*Hot	Whether a household head i was	-0.0175***

	born in the South macro-region (1 if yes)*Extreme Hot	(0.0003)
West*Hot	Whether a household head i was born in the West macro-region (1 if yes)*Extreme Hot	-0.0494*** (0.0006)
CA*Hot	Whether a household head i was born in California	-0.0311*** (0.0007)
CA*Cold	Whether a household head i was born in California (1 if yes)*Extreme Cold	-0.0289*** (0.0006)
The size of matrices: 2,417,253 households(row)*283 MSAs(columns)		

Notes: MSA fixed effects, which are interpreted as the mean indirect utility for each of the 283 MSAs, are not listed in this table. A scatter plot is shown in Appendix C. Midwest is left out as a reference while interacting birth region with extreme heat days.

In the second stage sorting model, the mean indirect utility for each MSA is added to an additional term computing the housing price index for each MSA to form the dependent variable. (See equation (7) in section 5). The second-stage results in column (1) of Table 2 show that extreme cold is negatively significant, which is consistent with our expectation. The aggregate effects from both extreme heat and extreme cold are negative after we combine coefficients from both 1st and 2nd stages (Table 3). Wage rates by sector (tax inclusive) are used to measure the impacts of job opportunities. Service wage rate is positively significant, and job opportunity tends to be a significant driver in people's location decisions. The coefficient of precipitation extreme is negatively significant, which suggests that precipitation negatively affects household location choice. The area of the body of water is positively significant. One explanation is that people prefer to live near a body of water, such as lake, river, and ocean. Total establishments of arts, entertainment, and recreation per square mile are positively significant, and people generally value entertainment and recreation.

The first column in Table 2 reports OLS estimation results using robust standard errors. We do not use IV regression in our paper, since the main variables (i.e. temperature extremes) that we are interested in are exogenous. In order to address the unobservable effects across locations, a region fixed-effects model is used in the second stage.

Table 2 Parameter Estimates from Second-Stage Sorting Model

Dependent variable: mean indirect utility from 1st stage of model version (1) in Table 3 + β_h log(price) , where $\beta_h = \beta_I(\rho_j H_i / I_{ij}) = 1*(17{,}767/45{,}071) = 0.3942$

Variables	OLS (robust standard error) (1)	Regional fixed effects (5 macro regions) (2)
Extreme Hot (mean number of days with maximum temp 90 degrees F or more/10)	-0.0140	-0.0278
	(0.0229)	(0.0211)
Extreme Cold (mean number of days with minimum temp 32 degrees F or less/10)	-0.0375*	-0.0278*
	(0.0185)	(0.0101)
Ln(Construction wage) ($000s)	0.0749	0.03500
	(0.4090)	(0.4772)
Ln(Production wage) ($000s)	-0.0830	0.1270
	(0.1964)	(0.2378)
Ln(Service wage) ($000s)	2.9635***	2.6279***
	(0.7418)	(0.4778)
Annual days of precipitation with daily maximum over 1 inch	-0.0438*	-0.0298*
	(0.0164)	(0.0178)
Annual snowfall (inches)	-0.00079	0.0035
	(0.0024)	(0.0029)
Annual # of tornado watches	-0.0136	-0.0019
	(0.0139)	(0.0114)
Water area (square miles) (00s)	0.0420**	0.0362***
	(0.0167)	(0.0113)
Total establishments of arts, entertainment, and recreation per square mile	0.5736**	0.6592**
	(0.2895)	(0.113)
R-square	0.3329	0.4147
Observations: 283		

The marginal willingness to pay (MWTP) to reduce additional extreme day is calculated by multiplying the regression coefficients (the ratio of coefficients of extreme temp and income, which is called WTP elasticity) by mean household income $45,071. One example is shown in Appendix H. Since extreme temperature days are scaled in 10 days, MWTP to reduce one extreme temperature day is then divided by 10.

Table 3 Estimated Marginal Willingness to Pay (MWTP) for Temperature Extremes

Measures	OLS (robust std. err.) (1)			Region fixed effects (2)		
	Extreme heat	Extreme cold	Extreme precipitation (daily precipitation over 1 inch)	Extreme heat	Extreme cold	Extreme precipitation (daily precipitation over 1 inch)
Coefficients of extreme weather	-0.0376	-0.0512	-0.0438	-0.0514	-0.0415	-0.0298
MWTP to reduce additional extreme weather day ($)	$169	$231	$1,970	$232	$187	$1,340

Notes: The marginal willingness to pay (MWTP) to reduce one extreme day is calculated by multiplying the regression coefficients (the ratio of coefficients of extreme temp and income, which is called WTP elasticity) by mean household income $45,071. One example is shown in Appendix H. Since extreme heat days and extreme cold days are scaled in 10 days, MWTP to reduce one extreme temperature day is then divided by 10.

6.2 Prediction in Population Shares

The aggregated probability by region based on equation (10) in section 5 represents the predicted population share in one of the five regions. Column (5)-(7) of Table 4 shows changes in predicted population shares across five macro-regions by comparing population shares calculated between the base scenario without climate change and the one with climate change. Column (1) of Table 4 presents the base scenario. Three climate change scenarios are listed in column (2)-(4). These three scenarios, respectively, represent the scenario that changes only the extreme cold matrix, the one that changes only the extreme heat matrix, and the one with changes in both extreme cold and extreme heat matrices.

Table 4 Changes in Predicted Population Shares by region in Response to Changes in Temperature Extremes

Regions	Probability of choosing a specific macro-region by different scenarios				Probability Change by comparing the projected probability (2061-2065) and probability from the empirical model (2000)		
	Base scenario (2000)	Only change extreme cold matrix (2061-2065)	Only change extreme heat matrix (2061-2065)	Change both extreme cold and extreme heat matrices (2061-2065)	Only change extreme cold matrix (2061-2065)	Only change extreme heat matrix (2061-2065)	Change both extreme cold and extreme heat matrices (2061-2065)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Northeast	0.2674	0.3175	0.2679	0.3186	0.0501	0.0005	0.0512
Midwest	0.1491	0.1753	0.1477	0.174	0.0262	-0.0014	0.0249
South	0.3315	0.3015	0.3309	0.3007	-0.030	-0.0006	-0.0308
West	0.0799	0.0704	0.0882	0.0771	0.0095	0.0083	-0.0028
California	0.1718	0.1348	0.1653	0.1294	-0.0370	-0.0065	-0.0424

Results from Table 4 show that population share decreases in South and California, while population share increases in Northeast under changes in climate extremes. My next-step research is to input the wage responses to changes in population shares back into the probability equation (10) to re-predict changes in population shares. Wage responses will be predicted from a computable general equilibrium (CGE) model. Population shares are likely to increase in California and Southern region due to a higher wage rates in these regions.

7. Conclusion

This paper uses a RUM sorting model that incorporates migration costs and allows for preference heterogeneity in temperature extremes. Results show that people born in different regions have different preferences towards temperature extremes. For example, people born in cold regions such as the Northeast and West are more averse to extreme heat than people born in warm regions such as South, while people born in California find extreme cold less amenable. This makes sense in terms of people's preferences for climates that are similar to their places of birth. Besides the climate of an individual's place of birth, other demographic characteristics also have significant impacts on individuals' location decisions. We find that highly educated people (e.g. college graduates) are more averse to extreme temperature than individuals without college degrees. This finding potentially reflects that college graduates have more job opportunities than those without college degrees, and these highly educated individuals become more mobile than people with low education levels. People over 65 years old are more averse to extreme temperatures. One reason might be that older people after retirement relocate to new places for the sake of pleasant amenities, and it is possible that extreme temperatures have higher impacts on their location decisions. We find that migration costs are significant. If migration costs are high, people are not willing to relocate to the place for the sake of a moderate change in climate.

Besides climate, other factors such as wage rates, natural amenities (e.g. water area), arts and entertainment are significant factors in household location choice. Service wage rates are positively significant in one's location choice. In particular, college graduates have stronger preferences for higher service wages. College graduates may have a higher probability to pursue a business job with higher wages, and business jobs are categorized into the service sector. Water area as an index of natural amenity is positively related to household location choice. The total establishments of arts, entertainment, and recreation per square mile as a measurement of abundance in recreational opportunities have a positive effect on residential location choice.

One contribution of this paper is that it captures preference heterogeneity, which allows us to better understand climate change impacts on migration and household location choice by considering preference heterogeneity across individuals. This paper shows that it is not the case that all individuals have homogenous preferences, and they do not have the same preferences for weather extremes. In contrast, our results show that highly mobile people are more averse to extreme temperatures. People over 65 years old are more averse to extreme temperatures. Individuals born in cold regions are more sensitive to extreme heat, while those born in warm regions are more sensitive to extreme cold.

In addition, we find that population shares in the Southern region and California drop, while population share in the Northeastern region gains under simulations in the climate change scenario. In the future research, we will bring wage responses to changes in regional labor supply caused by climate change-induced migration, however, population shares in California and Southern region are likely to rise due to higher wage rates. In the next step, we will input climate change-induced migration (the change in total population and population by education type) predicted from the empirical model into the computable general equilibrium (CGE) model. This CGE model will produce economic parameters (e.g. wage rates) in response to this population changes. Wage rates produced by the CGE model will be input back into the empirical RUM. Iterations will continue between the CGE and empirical RUM models until a locational equilibrium is achieved in the RUM sorting model.

References

Anselin, L. and A. K. Bera. (1998). Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In Handbook of applied Economic Statistics. Edited by A. Ullah and D. E. Giles New York: Marcel Dekker: 237-89

Bayer, P., Keohane, N., and Timmins, C. (2009). Migration and Hedonic Valuation: The Case of Air Quality. Journal of Environmental Economics and Management. 58: 1-14

Bayer, P.J., McMillan, R, and Rueben, K.S.. (2005). An Equilibrium Model of Sorting in an Urban Housing Market. NBER Working Paper Series, No. 10865.

Bayer, P.J. , Timmins, C.(2007). Estimating Equilibrium Models of Sorting Across Locations. The Economic Journal. 117: 353-374

Berry, S.T., Linton, O., and Pakes, A. (2004). Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems. Review of Economic Studies. 71: 613-654

Brown, Charles. (1980). Equalizing differences in the labor market. Quarterly Journal of Economics. 94: 113-34

Deschenes, O., Moretti, E. (2007). Extreme Weather Events, Mortality and Migration. NBER Working Paper

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

Denavas-Walt, C., Cleveland, R. W., and Roemer, M.I. (2001). Money Income in the United States: 2000. Current Population Reports: Consumer Income. U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau. Available online at http://www.census.gov/prod/2001pubs/p60-213.pdf

Epple, D., Romer, T., and Sieg, H. (2001). Interjurisdictional Sorting and Majority Rule: An Empirical Analysis. Econometrica. 69 (6): 1437-1465

Irwin, E.G. (2002). The Effects Open Space on Residential Property Values. Land Economics 78:465–80.

IPCC Workshop on Sea Level Rise and Ice Sheet Instabilities. (June 2010). Intergovernmental Panel on Climate Change (IPCC). Available online at https://www.ipcc-wg1.unibe.ch/publications/supportingmaterial/SLW_WorkshopReport.pdf

Kuminoff, N.V. (2009). Decomposing the Structural Identification of Nonmarket Values. Journal of Environmental Economics and Management. 57: 123-139

Klaiber, A., and Phaneuf, D. (2010). Valuing Open Space in a Residential Sorting Model of the Twin Cities. Journal of Environmental Economics and Management. 60: 57-77.

Maurer, E.P., A.W. Wood, J.C. Adam, D.P. Lettenmaier, and B. Nijssen. (2002). A Long-Term Hydrologically-Based Data Set of Land Surface Fluxes and States for the Conterminous United States, J. Climate 15(22), 3237-3251

Mansur, E.T., Mendelsohn, R., and Morrison, W. (2008). Climate Change Adaptation: A Study of Fuel Choice and Consumption in the US Energy Sector. Journal of Environmental Economics and Management. 55: 175-193

Poston, D.L. et al. (2009). The effect of climate change on migration: United States, 1995-2000 in Social Science Research, 38743-5

Parmeter, C. F. and Pope, J. C. (2009). Quasi-Experiments and Hedonic Property Value Methods. Prepared for the Handbook on Experimental Economics and the Environment.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy. 82: 34-55

Smith, K. and Gilbert, C. (1985). Valuation of Environmental Risks Using Hedonic Wage Models. Working Paper from the National Bureau of Economic Research. P.359-392. Available online at http://ideas.repec.org/h/nbr/nberch/6156.html

Tra, C. (2010). A Discrete Choice Equilibrium Approach to Valuing Large Environmental Changes. Journal of Public Economics. 94: 183-196

Timmins, C. (2007). If You Cannot Take the Heat, Get Out of the Cerrado…Recovering the Equilibrium Amenity Cost of Nonmarginal Climate Change in Brazil. Journal of Regional Science. 47: 1-25

Walsh, R. (2007). Endogenous Open Space Amenities in a Locational Equilibrium. Journal of Urban Economics. 61: 319-344

Appendix A

Table A.1 Data Summary for Hedonic Housing Price Regression		
	Mean	Description
acre_9	0.1417	Acreage of property 1-9 acreages
acre_10	0.02629	Acreage of property 10+ acreages
room2	0.0127	2 rooms in dwelling
room3	0.0412	3 rooms in dwelling
room4	0.0850	4 rooms in dwelling
room5	0.2026	5 rooms in dwelling
room6	0.2350	6 rooms in dwelling
room7	0.1739	7 rooms in dwelling
room8	0.1243	8 rooms in dwelling
room9	0.1230	9 rooms in dwelling
bed2	0.0386	1 bedroom dwelling
bed3	0.2054	2 bedroom dwelling
bed4	0.4880	3 bedroom dwelling
bed5	0.2131	4 bedroom dwelling
bed6	0.0479	5 or more bedroom dwelling
unit2	0.0011	Boat, tent, van, other
unit3	0.7819	1 family house, detached
unit4	0.0633	1 family house, attached
unit5	0.0199	2 family building
unit6	0.0112	3-4 family building
unit7	0.0084	5-9 family building
unit8	0.0064	10-19 family building
unit9	0.0072	20-49 family building
unit10	0.0127	50+ family building
Noplumb	0.0037	Dwelling does not contain complete kitchen facilities
Nokitch	0.0027	Dwelling does not contain complete plumbing facilities
yr1	0.0239	0-1 year-old dwelling
yr2	0.0794	2-5 year-old dwelling
yr3	0.0793	6-10 year-old dwelling
yr4	0.1548	11-20 year-old dwelling
yr5	0.1690	21-30 year-old dwelling
yr6	0.1375	31-40 year-old dwelling
yr7	01474	41-60 year-old dwelling

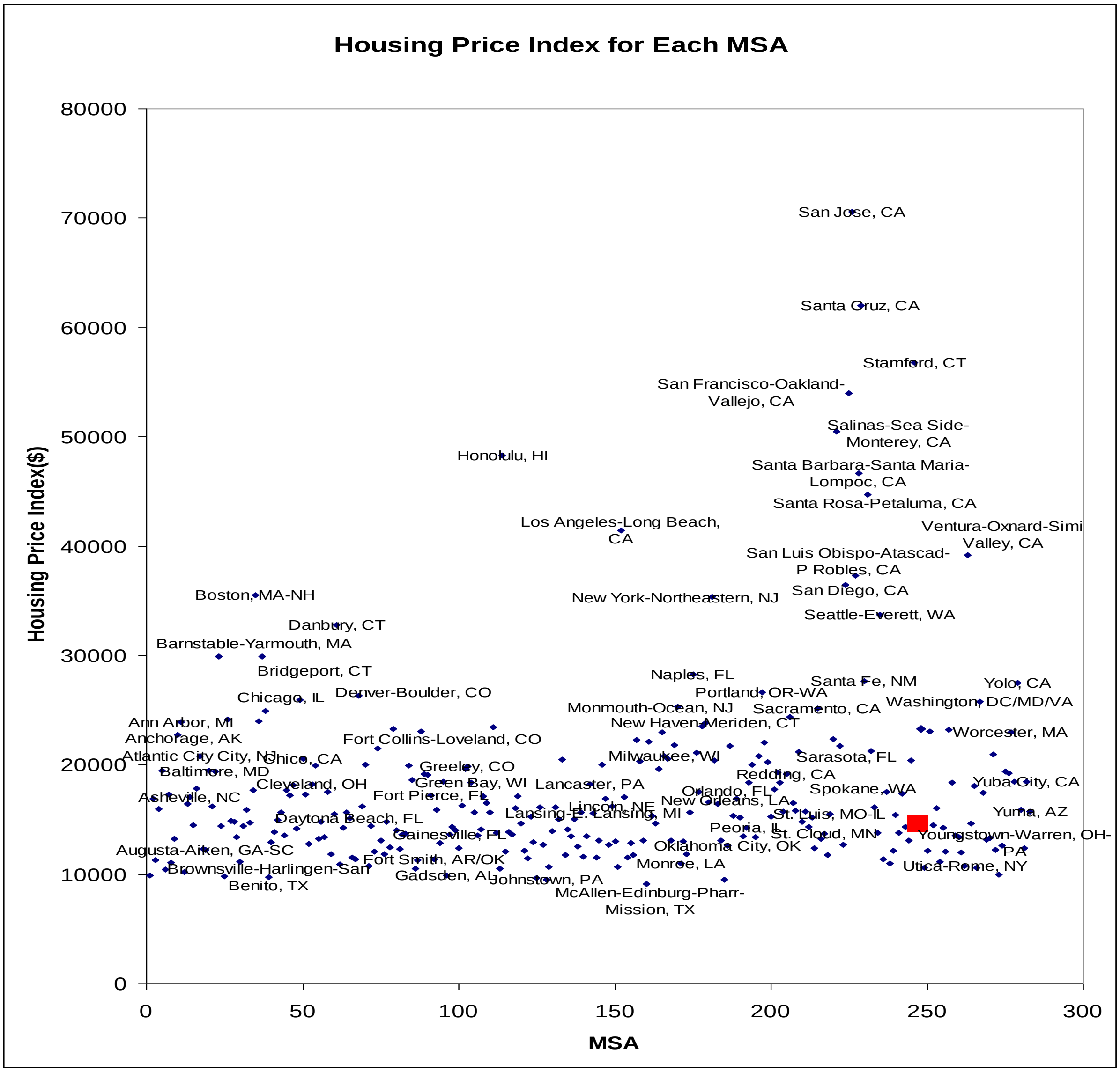
The following Table A.2 shows results from hedonic housing price regression.

Table A.2 Results from Hedonic Housing Price Regression
Dependent Variable: log(housing price)

	Coef.	Std.	Err.	T	P>|t|	[95% Conf. Interval]
acre_9	0.238574	0.001115	213.9	0	0.236388	0.24076
acre_10	0.504254	0.002377	212.12	0	0.499594	0.508913
room2	0.316785	0.009854	32.15	0	0.297471	0.336099
room3	0.495023	0.009708	50.99	0	0.475995	0.514051
room4	0.446755	0.009965	44.83	0	0.427225	0.466285
room5	0.64063	0.009986	64.15	0	0.621057	0.660203
room6	0.798261	0.010019	79.68	0	0.778625	0.817897
room7	0.948742	0.010045	94.45	0	0.929055	0.968429
room8	1.086049	0.010081	107.74	0	1.066292	1.105807
room9	1.309055	0.010111	129.46	0	1.289237	1.328872
Bed2	-0.1169	0.005831	-20.05	0	-0.12833	-0.10548
Bed3	-0.02974	0.006033	-4.93	0	-0.04156	-0.01791
Bed4	0.056016	0.006135	9.13	0	0.043991	0.068041
Bed5	0.126727	0.006225	20.36	0	0.114526	0.138929
Bed6	0.195024	0.006471	30.14	0	0.182341	0.207707
Unit2	-0.35472	0.011706	-30.3	0	-0.37766	-0.33178
Unit3	0.802728	0.00139	577.43	0	0.800004	0.805453
Unit4	0.67791	0.00202	335.61	0	0.673951	0.681869
Unit5	0.860843	0.003036	283.58	0	0.854893	0.866792
Unit6	0.880736	0.003779	233.06	0	0.873329	0.888142
Unit7	0.760116	0.004295	176.99	0	0.751698	0.768534
Unit8	0.727957	0.004882	149.11	0	0.718389	0.737526
Unit9	0.826935	0.004644	178.08	0	0.817834	0.836036
Unit10	0.981787	0.003685	266.41	0	0.974564	0.98901
noplumb	-0.17694	0.006934	-25.52	0	-0.19053	-0.16335
nokitch	-0.17015	0.008067	-21.09	0	-0.18597	-0.15434
yr1	0.470302	0.002587	181.78	0	0.465231	0.475373
yr2	0.416239	0.001625	256.11	0	0.413054	0.419425
yr3	0.338295	0.001613	209.71	0	0.335134	0.341457
yr4	0.210377	0.001327	158.57	0	0.207777	0.212978
yr5	0.074306	0.001278	58.14	0	0.071801	0.076811
yr6	0.062975	0.001325	47.54	0	0.060379	0.065571
yr7	0.061517	0.001288	47.76	0	0.058992	0.064041
Constant	9.7127	0.012807	803.0272	0	9.687627	9.73783
R-square: 0.9976 Observations: 2,417,253						

The following graph shows housing price index in each MSA (obtained from hedonic housing price regression)

Figure A Housing Price Index for Each MSA



Note: the red plot is where State College, PA locates

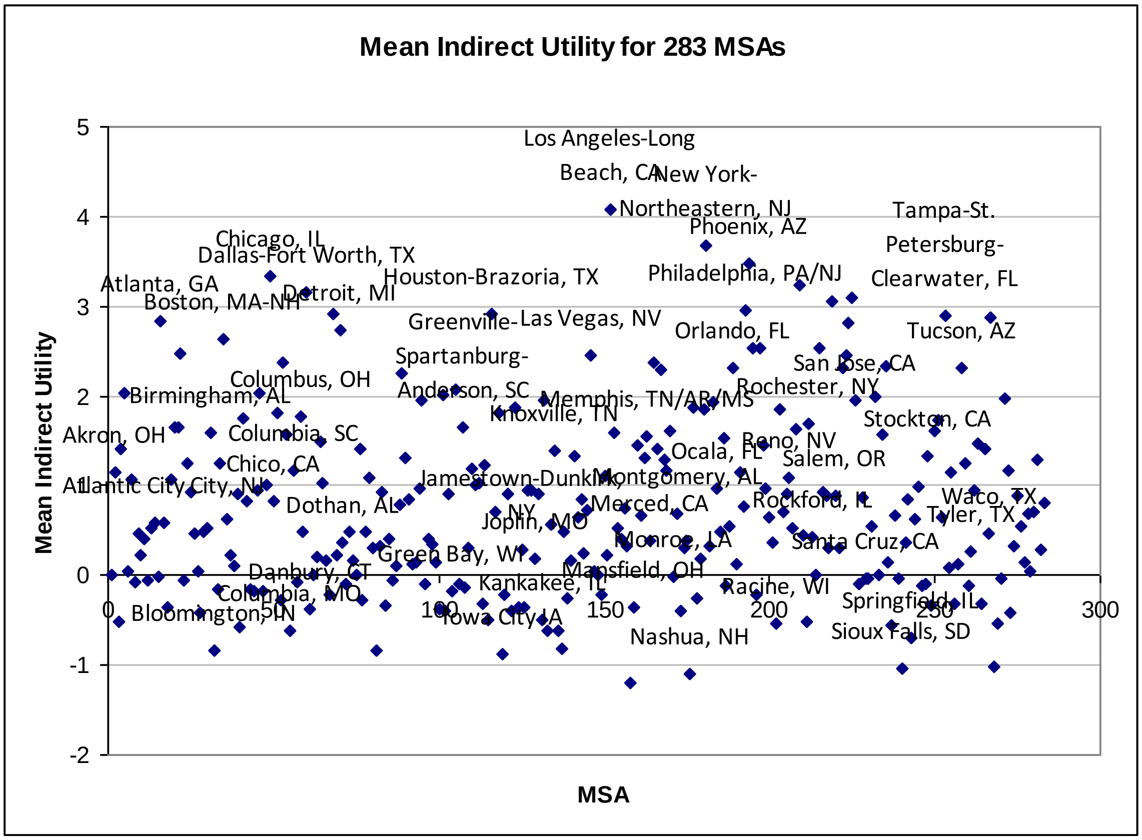
Appendix B

Table B.1 Demographic Variable Description		
Variable	Mean	Description
WHITE	0.837	White = 1; Non-white = 0
MALE	0.706	Male = 1; Female = 0
AGE>60	0.304	Age>60 = 1; Age <=60 = 0
HSDROP	0.0539	High school dropout
HSGRAD	0.419	High school graduate
SOMECOLL	0.3998	Completed some college (not four year degree)
COLLGRAD	0.127	College graduate
Lntotinc	10.82	Log(total personal income $)

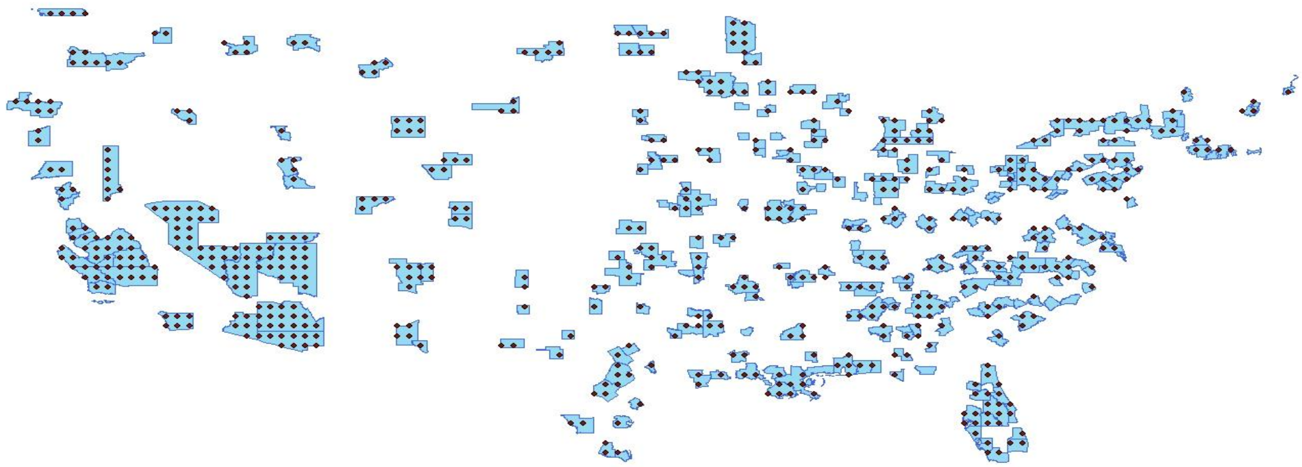
Table B.2 Results from Income Regression						
Lninctot	Coef.	Std. Err.	T	P>|t|	[95% Conf. Interval]	
Age_g_60	-0.32602	0.001126	-289.64	0	-0.32823	-0.32381
Male	0.483654	0.001099	440.05	0	0.4815	0.485808
White	0.169664	0.001403	120.97	0	0.166915	0.172413
Hsgrad	0.279948	0.002316	120.87	0	0.275408	0.284487
Coll	0.632171	0.002361	267.78	0	0.627544	0.636798
Collgrad	0.995492	0.002628	378.79	0	0.990341	1.000642
Constant	9.835	0.0131	926.36	0	9.8086	9.86
Observations: 2,417,253						

Notes: HSDROP is left out and is included in the constant as a reference of other education types.

Appendix C Mean Indirect Utility for 283 MSAs



Appendix D Intersect Gridded Temperature Data with MSAs

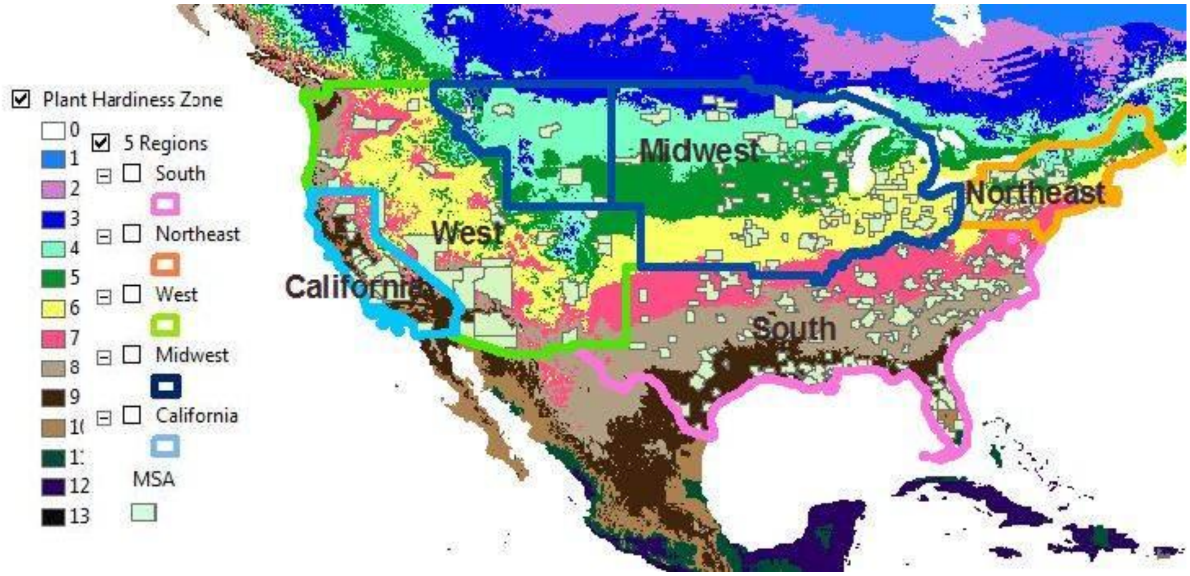


On the map, polygons represent 283 MSAs. Dots represent temperature data with exceedance days. The annual number of days with maximum daily temperature above 90F, annual number of days with minimum daily temperature below 32F, and annual number of days with maximum daily precipitation over 1 inch are calculated based on the arithmetic mean of extreme days in each MSA.

Appendix E Descriptive Statistics for Site-Specific Attributes

Variable	Obs.	Mean	Std. Dev.	Min	Max	Description
Extreme Hot	283	4.0	4.3542	0	17.5	mean annual number of days/10 with maximum temp 90 degrees F or more (Ed Maurer's downscaled data)
Extreme Cold	283	9.1	5.1862	0	21.3	mean annual number of days/10 with minimum temp 32 degrees F or less (Ed Maurer's downscaled data)
Ln (Construction wage) ($000s)	283	3.4633	0.1915	2.8707	3.9522	Natural log of construction wage (wage rates data from the U.S. Bureau of Labor Statistics. ($000s))
Ln(production wage) ($000s)	283	3.2353	0.2479	0.8671	3.7746	Natural log of production wage ($000s)
Ln(service wage) $000s	283	3.4375	0.1242	2.9684	3.9174	Natural log of service wage ($000s)
Annual snowfall (inches)	283	17.9694	23.5865	0	115.6	Annual snowfall (inches) (NCDC)
Extreme precipitation	283	7.6	2.9231	1	22	Annual days of precipitation with daily maximum over 1 inch (Ed Maurer's downscaled data)
Annual number of tornado watches	283	8.5018	5.3438	0	40	Annual number of tornado watches (NCDC)
Total establishments of arts, entertainment, and recreation per square mile	283	0.1419	0.3148	0.004	4.227	Total establishments of arts, entertainment& recreation/land are (in square miles) (U.S. Census)
Water area	283	2.47	5.13	0.0073	39.55	Water area (area in square miles/100) (U.S. Census)

Appendix F Five Regions Defined



Regions defined by coordinating economic regions with USDA plant hardiness zones: 1) Northeast (CT, ME, MA, NH, RI, VT, NJ, NY, PA); 2) Midwest (IA, MN, NE, SD, ND, MT, WY, IL, IN, MI, OH, WI); 3) South (FL, GA, AR, MD, NC, SC, VA, WV, AL, KY, MS, TN, LA, KS, MO, OK, AR, TX); 4) West (NV, AZ, CO, NM, UT, OR, WA, ID); 5) California

*AK and HI are excluded due to the unavailability of projected temperature data in these two states

Table F Summary Statistics of Temperature Extremes by Regions (2000 vs. 2061-2065)

Regions	Description	Time Period 2000				Projected Time Period 2060-2065			
		Mean	Min	Max	Std. Dev.	Mean	Min	Max	Std. Dev.
Northeast	Days above 90F	17	0	77	19.89	42	17	70	14.75
(46 MSAs)	Days below 32F	153	79	177	20.05	132	66	165	19.62
Midwest	Days above 90F	12	0	143	23.91	72	1	181	28.83
(72 MSAs)	Days below 32F	147	28	200	28.41	136	84	205	24.84
West	Days above 90F	12	0	143	23.91	61	1	210	62.05
(30 MSAs)	Days below 32F	128	42	194	21.8	149	32	238	59.51
South	Days above 90F	68	1	166	39.96	117	40	185	27.43
(111 MSAs)	Days below 32F	54	0	167	38.31	49	1	169	49.41
California	Days above 90F	58	4	129	33.3	118	67	165	24.26
(22 MSAs)	Days below 32F	47	6	136	42.02	66	31	151	37.98

Source: Data is provided by Rob Nicholas in the Department of Geosciences at Penn State University
Data is obtained from NARCCAP, Canadian Regional Climate Model (CRCM)

Appendix G
Derive the Second-Stage Sorting Model
and the Coefficient of Housing Price Coefficient

Maximize utility subject to budget constraint, set up the Lagrangian expression

$$\underset{C_i,H_i,X_j}{Max\ell} = C_i^{\beta_c} H_i^{\beta_h} Z_j^{\beta_x} e^{\sum_{q=1}^{Q} \beta_q^T (HH_q^i \times T^j) + \beta_{edu}^w (EDU_i \times W_j) + \beta_m M_{ij} + \xi_j + \eta_{ij}} + \lambda(I_{ij} - C_i + \rho_j H_i)$$

Individuals choose their location j, along with consumption of C_i and H_i to maximize their utility subject to a budget constraint.

F.O.C. with respect to C_i and H_i

$$\begin{cases} \dfrac{\partial \ell}{\partial C_i} = \beta_c C_i^{\beta_c - 1} H_i^{\beta_h} Z_j^{\beta_x} e^{\sum_{q=1}^{Q} \beta_q^T (HH_q^i \times T^j) + \beta_{edu}^w (EDU_i \times W_j) + \beta_m M_{ij} + \xi_j + \eta_{ij}} - \lambda = 0 & (G.1) \\ \dfrac{\partial \ell}{\partial H_i} = \beta_h H_i^{\beta_h - 1} C_i^{\beta_c} Z_j^{\beta_x} e^{\sum_{q=1}^{Q} \beta_q^T (HH_q^i \times T^j) + \beta_{edu}^w (EDU_i \times W_j) + \beta_m M_{ij} + \xi_j + \eta_{ij}} + \lambda \rho_j = 0 & (G.2) \\ \dfrac{\partial \ell}{\partial \lambda} = I_{ij} - C_i + \rho_j H_i = 0 & (G.3) \end{cases}$$

In equilibrium, individuals must be indifferent among locations. If not, they would prefer to move. Hence, I can write H_i, C_i, and I_{ij} as H_{ij}, C_{ij}, and I_{ij}.

From (G.1)/(G.2) $\quad \dfrac{\beta_c}{\beta_h} \dfrac{H_{ij}}{C_{ij}} = \dfrac{1}{\rho_j}$

From (G.3) $\quad C_{ij} + \rho_j H_{ij} = I_{ij}$

$$H_{ij} = \frac{\beta_h C_{ij}}{\beta_c \rho_j} = \frac{\beta_h (I_{ij} - \rho_j H_{ij})}{\beta_c \rho_j} = \frac{\beta_h I_{ij} - \rho_j \beta_h H_{ij}}{\beta_c \rho_j}$$

$$(\beta_c \rho_j + \beta_h \rho_j) H_{ij} = \beta_h I_{ij}$$

$$H_{ij} = \frac{\beta_h}{\beta_c + \beta_h} \frac{I_{ij}}{\rho_j} \qquad \text{(G.4)}$$

Substitute H_{ij} into equation (G.3), $C_{ij} = \dfrac{\beta_c}{\beta_h + \beta_c} I_{ij}$ (G.5)

Plugging (G.4) and (G.5) into utility function, the indirect utility function is obtained:

$$\begin{aligned} V_{ij} &= \left(\frac{\beta_c}{\beta_c + \beta_h} I_{ij}\right)^{\beta_c} \cdot \left(\frac{\beta_h}{\beta_c + \beta_h} \frac{I_{ij}}{\rho_j}\right)^{\beta_h} \cdot Z_j^{\beta_x} e^{\sum_{q=1}^{Q} \beta_q^T (HH_q^i \times T^j) + \beta_{edu}^w (EDU_i \times W_j) + \beta_m M_{ij} + \xi_j + \eta_{ij}} \\ &= \left(\frac{\beta_c}{\beta_c + \beta_h}\right)^{\beta_c} \cdot \left(\frac{\beta_h}{\beta_c + \beta_h}\right)^{\beta_h} \cdot I_{ij}^{\beta_c + \beta_h} \cdot \rho_j^{-\beta_h} Z_j^{\beta_x} e^{\sum_{q=1}^{Q} \beta_q^T (HH_q^i \times T^j) + \beta_{edu}^w (EDU_i \times W_j) + \beta_m M_{ij} + \xi_j + \eta_{ij}} \qquad \text{(G.6)} \\ &\cong I_{ij}^{\beta_c + \beta_h} \cdot e^{\sum_{q=1}^{Q} \beta_q^T (HH_q^i \times T^j) + \beta_{edu}^w (EDU_i \times W_j) + \beta_m M_{ij} + \xi_j + \eta_{ij}} \end{aligned}$$

Let $\beta_I = \beta_c + \beta_h$, $\Theta_j = -\beta_h \ln \rho_j + \beta_x \ln Z_j + \xi_j$, $I_{ij} = \hat{I}_{ij} + \varepsilon_{ij}^I$, and $v_{ij} = \beta_I \varepsilon_{ij} + \eta_{ij}$ and take the log of indirect utility, equation (G.6) becomes the following

$$\begin{aligned}\ln V_{ij} &= \beta_I \ln I_{ij} + \sum_{q=1}^{Q} \beta_q^T (HH_q^i \times T^j) + \beta_{edu}^w (EDU_i \times W_j) + \beta_m M_{ij} + \Theta_j + v_{ij} \\ &= \beta_I \ln \hat{I}_{ij} + \sum_{q=1}^{Q} \beta_q^T (HH_q^i \times T^j) + \beta_{edu}^w (EDU_i \times W_j) + \beta_m M_{ij} + \Theta_j + \eta_{ij}\end{aligned} \tag{G.7}$$

Recall $\Theta_j = -\beta_h \ln \rho_j + \beta_x \ln Z_j + \xi_j$, in the second stage sorting model, MSA fixed effects Θ_j can be decomposed according to this equation. In this case, predicted income for every location j is entered into indirect utility function as a standalone measure.

In the second stage, the regression equation is:

$$\Theta_j = -\beta_h \ln \rho_j + \beta_x \ln Z_j + \xi_j \tag{G.8}$$

Now move $-\beta_h \ln \rho_j$ to the LHS of equation (G.7), regression equation becomes the following ($CLIMATE_j$ is included in Z_j):

$$\Theta_j + \beta_h \ln \rho_j = \beta_x \ln Z_j + \xi_j \tag{G.9}$$

From equation (4) $H_{ij} = \dfrac{\beta_h}{\beta_c + \beta_h} \dfrac{I_{ij}}{\rho_j}$

$$\beta_h = \beta_I (\rho_j H_i / I_{ij}) \tag{G.10}$$

The parameter β_I is estimated in the first stage of sorting model, and set $\rho_j H_i / I_{ij}$ (the share of housing expenditure in income) equal to its median value in the sample.

From our regression results, $\beta_I = 1.00$, and the mean values $\rho_j = 17{,}767$ $I_{ij} = 45{,}071$,

$$\beta_h = \beta_I (\rho_j H_i / I_{ij}) = 1.00 * (17{,}767 * 1) / 45{,}071 = 0.3942 \tag{G.11}$$

Appendix H

The regression coefficient for extreme heat days is calculated as the following by combining results from both 1st and 2nd stage (example from OLS sorting model (1) in Table 2):

1) Coefficient of extreme heat (overall effect):

$$\beta_{heat} = \frac{\partial \ln V}{\partial HEAT} = \beta_{heat\_1st\_stage} + \beta_{heat\_2nd\_stage}$$
$$= \beta_{age\_g\_65\_heat} \times \overline{AGE} + \beta_{collgrad\_heat} \times \overline{COLLGRAD} + \beta_{northeast\_heat} \times \overline{NE}$$
$$+ \beta_{south\_heat} \times \overline{SO} + \beta_{west\_heat} \times \overline{WE} + \beta_{CA\_heat} \times \overline{CA} + \beta_{cold\_2nd\_stage}$$
$$= -0.0076 * 0.2465 - 0.0268 * 0.1269 - 0.0286 * 0.2731 - 0.0175 * 0.3037 - 0.0494 * 0.0604 - 0.0311 * 0.0712 - 0.0140$$
$$= -0.0376$$

2) Coefficient of extreme cold (overall effect):

$$\beta_{cold} = \frac{\partial \ln V}{\partial cold} = \beta_{cold\_1st\_stage} + \beta_{cold\_2nd\_stage}$$
$$= \beta_{age\_g\_65\_cold} \times \overline{AGE} + \beta_{collgrad\_cold} \times \overline{COLLGRAD} + \beta_{CA\_cold} \times \overline{CA} + \beta_{cold\_2nd\_stage}$$
$$= -0.0316 * 0.2465 - 0.0305 * 0.1269 - 0.0286 * 0.0712 - 0.0375$$
$$= -0.0512$$

Mean value of household head's income is $45,071, and mean values of extreme heat days and extreme cold days are measure in 10 days. Coefficient of marginal utility of income is $\beta_I = 1.00$. MWTP to reduce additional extreme heat day = (0.0376/1)*45,071/10 = $169. MWTP to reduce additional extreme cold day = (0.0512/1)*45,071/10 = $231