ABSTRACT

Title of Document: THE VALUE OF CLIMATE AMENITIES: A DISEQUILIBRIUM APPROACH

Paramita Sinha, Ph.D., 2008

Directed By: Professor Maureen Cropper, Department of Economics

Formulating efficient climate policies requires estimates of the impacts of climate change. An important category of impacts are climate amenities—the value people attach to temperature and precipitation. There is a large literature that attempts to value climate amenities using the fact that climate amenities will be capitalized into wages and property values. Many of these estimates assume that people are perfectly mobile and are based on estimates of national hedonic wage and property value functions. These functions will yield biased estimates of consumers' willingness-to-pay if consumers are not in locational equilibrium due to search or migration costs or if markets take time to adjust.

I provide estimates of the value of climate amenities in the US using a discrete model of residential location choice. I model the location choices of over 400,000 households who changed metropolitan statistical areas between 1995 and 2000 using

the 5% PUMS data from the Census. To avoid making equilibrium assumptions, I face the migrants with the market conditions in each MSA.

The empirical model is motivated by a Random Utility Model framework, which posits that the utility that a household derives from living in an MSA depends on climate amenities along with earnings potential, housing costs and location-specific amenities. Households choose the MSA where they derive the maximum utility. The model is estimated using a two step procedure (Bayer, Keohane and Timmins, 2006). In the first stage, location-specific constants are estimated together with other parameters of the utility function. In the second stage, these location-specific intercepts are regressed on location-specific amenities to estimate the average utility attached to these amenities.

The dissertation estimates the marginal rate of substitution between climate variables and income. The results show that households facing an average winter temperature of 37 degrees Fahrenheit are willing to pay approximately about 3% of their income for an increase in average winter temperature by one degree. Willingness to pay to raise summer precipitation by an inch from a level of about 11 inches is roughly 3% of their income. The study also provides estimates of the quality of life in 297 Metropolitan Areas.

THE VALUE OF CLIMATE AMENITIES: A DISEQUILIBRIUM APPROACH

By

Paramita Sinha

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2008

Advisory Committee:
Professor Maureen L. Cropper, Chair
Professor Kenneth E. McConnell
Professor Lars Olson
Associate Professor John K. Horowitz
Professor Mahlon Straszheim

Dedication

To Dia and Dadi – I am truly blessed to have such grandparents.

Acknowledgements

First and foremost, I would like to thank my advisor Maureen Cropper for her invaluable guidance. Her comments and insights into my research were indispensable. She is an ideal advisor who not just expects a very high quality of work from her students but also strives to instill such standards in them. It is impossible to express in a few words the extent of support and encouragement she has given me and no student could hope for more.

I also thank my committee members John Horowitz, Ted McConnell, Lars Olson and Mahlon Straszheim for taking the time to serve on my committee. I thank Ted McConnell and Nancy Bockstael for their comments during the initial phase of my work. Comments from all of them have helped to improve the quality of my dissertation. I also take this opportunity to thank John List, Howard Leathers and Ken Leonard for their support and encouragement during my graduate school years. I would like to acknowledge the staff members of our department - Chris Aggour, Barbara Burdick, Jeff Cunningham, Jane Doyle, Curtis Henry, Liesl Koch, Katherine Faulkner, Chuck McCormick and Lynda Wilkinson.

I thank Partha Lahiri for all the time and help he has given me with data issues. I also thank Don McCubbin from Abt Associates, Paul MacKun from the Census Bureau, Mike Squires and Scott Stephens from the National Oceanic and Atmospheric Administration for help with obtaining data. I am grateful to Michael Greenstone and

Olivier Deschenes for generously sharing some of their data with me. The IPUMS and the SAS technical support group have also been of immense help.

I must express my gratitude towards Jyotsna Jalan for always believing in me and supporting me when I was applying for graduate school – I cannot thank her enough. I would like to thank Steve Smith for being such a great mentor. I also thank Marshall Wise and everyone at the Joint Global Change Research Institute for making my time at JGCRI such a rewarding experience. I learnt a lot during my time at JGCRI and believe that this experience will be greatly beneficial to me in my career.

This work would not have been possible without the support and encouragement of my family and friends. I am greatly indebted to my parents for the countless sacrifices and years of hard work they have put in to help me achieve my goals. I honestly have no words that will suffice to thank my maternal grandparents and uncle for their blessings and unwavering support throughout the years. I would especially like to thank my grandmother for being such an inspiration to me. I must express my gratitude towards Ela mashi for always believing in me and being there for me. All these people have been the pillars of my strength and I would not be where I am if it had not been for them. I cannot forget my late grandfather for helping make my childhood years memorable. I owe a special thank you to my mother –in –law who has always been extremely supportive and more of a friend than anything else and also to all my other in- laws for their understanding and support.

No acknowledgement would be complete without mentioning those who made this entire experience a bright, enjoyable and memorable one. My childhood friends Renata, Shabnam and Dolly have stood by me all these years. I am truly lucky to have them. A thank you to my friends from my days at the Indian Statistical Institute - Aniruddha Bagchi, Sudip Gupta and Arijit Mukherjee for their support and encouragement. I must say a special thank you to Sandipan Ray for always being there for me when I have needed a friend. I would like to thank Khusi and Gopal Dhar for making me smile even during the roughest times. I must express my gratitude to Moumita and Souvik Gupta, Barnali Das and Nilanjan Chatterjee for their endless patience and support through the rough times. A special mention must also be made of Soma Choudhury and Satarupa Nandy for being such great listeners. I am very grateful to Subrata Kundu and Sudeshna Basu. I cannot forget Mishi, Pokai, Ponto, Rimi and Swaha who have been a constant source of joy and have showered me with unconditional love and affection.

Last, but not the least, my heartfelt thanks to my husband, Dipan for his incredible patience and understanding, without which I could not have done this. He has always been there for me, through the good times and bad. Thank you, Dipan, for always believing in me.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Emissions of greenhouse gases contribute to the growing problem of climate change; however, the costs of abating such gases are high. Efficiently assessing policies to reduce emissions therefore requires valuing the impacts of climate change. These impacts include both market and non-market effects. Market impacts are likely to occur in agriculture, forestry, coastal infrastructure, energy use, and recreation. Non-market impacts of climate change include effects on climate amenities, human health and species loss. The focus of this dissertation is to provide information about the value of one category of non-market effects—climate amenities in the United States.

The value of climate amenities is measured by what people are willing to pay to obtain them. The goal of this study is to use data on migrants in the 2000 PUMS dataset to estimate the marginal rate of substitution between income and climate amenities for migrant households[1]. I use a discrete choice approach to model the location decisions of households in the US who moved between 1995 and 2000. The utility that a household derives from living in a location is assumed to depend on potential earnings, housing costs, climate amenities and other location-specific amenities. Households are assumed to choose the location from which they derive maximum utility. Under standard assumptions, the probability that the household chooses a location is given by the conditional logit model.

[1] Exact welfare measures are not computed in a discrete choice framework.

The model is estimated in two stages, following Bayer, Keohane and Timmins (2006). In the first stage, MSA-specific constants are estimated together with other utility function parameters to explain the location choices of migrants. In the second stage, the MSA-specific constants, which may be interpreted as Quality-of-Life indices, are regressed on amenities that vary by MSA to estimate the average utility attached to these amenities. This procedure allows me to identify the parameters of consumers' utility functions and in particular, the preference parameters for climate variables.

There are few recent estimates of the value of climate amenities in the US. Most estimates in the literature are based on hedonic wage and property value functions, following the approach of Rosen (1974) and Roback (1982). These studies, including Blomquist et al. (1988), Gyourko and Tracy (1991) and Smith (1983), assume that households and firms can migrate costlessly from one location to another and that, as a result, national labor and housing markets are in equilibrium. The continuous hedonic approach assumes that households (and firms) move immediately in response to exogenous shocks, such as federal pollution regulations and technical change. However, in reality, adjustments may not occur instantaneously. This could be due to several factors. For consumers, impediments to migration include transportation costs, search costs (for jobs and housing), and personal and family considerations. Firms may face barriers to entry into a region. Examples of these include permit requirements, long term contracts and transportation costs. If, as a result, national housing and labor markets are not in equilibrium, the partial derivatives of national hedonic wage and property value functions will not measure marginal willingness to pay for amenities.

Cragg and Kahn (1997) overcome some of these difficulties by valuing climate amenities using a discrete model of location choice. They focus on households who migrated between 1975 and 1980, who are more likely to be in locational equilibrium than all households in the population. When households choose the state in which to live, their earnings opportunities are described by state-specific hedonic wage equations. Cragg and Kahn thus avoid the assumption of a national labor market.

My analysis builds on and extends the work of Cragg and Kahn. Like Cragg and Kahn, I focus on migrant households. It is reasonable to assume that these households are in locational equilibrium and also reasonable to treat conditions in the housing and labor markets in each MSA as exogenous to migrants. I extend Cragg and Kahn by explicitly include moving costs in my model, which increase as the migrant moves to a different state or a different region of the country. These costs significantly affect estimates of the value of climate amenities. Further, by employing the two-stage estimation approach of Bayer et al. (2006) I am able to estimate Quality of Life indices for each MSA in 2000, in addition to valuing climate amenities.

The results indicate that households facing an average winter temperature of 37 degrees Fahrenheit are willing to pay approximately about 3% of their income for an increase in average winter temperature by one degree. However, marginal willingness to pay decreases as average temperature increases reaching 0 at a temperature of about 58 degrees Fahrenheit. This is due to the fact that households prefer higher temperatures but there exists an "optimum temperature" beyond which higher temperatures reduce utilities. Willingness to pay to raise summer precipitation from a level of about 11 inches throughout the season by one inch equals roughly 3% of household income. Households are also willing to pay about 2% of their income to lower winter precipitation by an inch from its mean level of 9 inches. There is no evidence that average summer temperature plays a significant role in household location decision.

This dissertation is organized as follows. I briefly review the literature on valuing climate amenities in Chapter 2. In Chapter 3 I present the household's location decision and the econometric models to be estimated. Chapter 4 presents stylized facts about migration patterns and spatial variation in wages and housing costs in the US. It also describes the data used in my analysis. Results are presented in Chapter 5. Chapter 6 concludes the study.

Chapter 2: Literature Review and Motivation

This chapter describes different approaches to valuing climate amenities in the literature, beginning with the hedonic approach (Section 2.1.1). Section 2.1.2 discusses the limitations of this approach. The following section briefly describes the studies that have used discrete models of location choice to value amenities. As the migration literature is very closely related to the topic of location choice of households, Section 2.1.4 provides a brief overview of these models. Section 2.2 provides a summary of which climate variables have been found to be amenities or disamenities in past empirical studies.

2.1 Different Approaches to Valuing Climate Amenities

There are three strands of the economics literature that value climate amenities. The first uses hedonic wage and property value functions to compute marginal amenity values, following Rosen (1974) and Roback (1982). The second uses discrete models of consumer location choice to estimate the parameters of household utility functions. The migration literature quantifies the role that climate amenities play in migration decisions, although it does not provide estimates of willingness to pay.

2.1.1 Hedonic Approach

The continuous hedonic approach, developed by Rosen (1974) is based on the notion that location-specific amenities, such as climate, should be reflected in household location decisions, and, hence, be capitalized into wages and land values: Other things equal, workers should accept lower wages to live in more pleasant climates and should be willing to pay more for housing in more desirable climates. In the labor market, for example, if we hold other location attributes constant, an individual must receive a positive compensating differential for living in a less agreeable climate to keep his utility constant. Workers moving out of cities with less desirable climates will reduce the supply of labor, putting upward pressure on wages. This will also reduce the demand for land, putting downward pressure on rents. The assumptions that drive these results are that consumers are perfectly mobile and that firm location decisions are unaffected by climate amenities.

This approach was refined by Roback (1982) who emphasized that the implicit price of attributes obtained from hedonic studies reflects not only their marginal value to consumers but also their marginal cost to firms. In her model, Roback incorporates firm behavior and allows amenities to influence firm productivity. She shows that if amenities directly affect firms' costs, the results in the previous paragraph do not necessarily hold. For example, clean air would be an "unproductive amenity" as firms have to use a nonpolluting technology to produce it and this would raise costs. Workers would be willing to accept lower wages and firms would also pay lower wages to keep costs down. The effect on property values would be ambiguous as firms would want to move away but households would migrate into the area. An example of a productive amenity would be “lack of heavy snowfall.” In this case, workers are willing to work for less in a milder climate, but firms are willing to pay more due to reduced absenteeism. Property values should, however, be unambiguously higher: both firms and workers would like to purchase land in cities with milder climates, which should drive up land prices.

The Roback Model (1982)

Roback (1982) assumes that the amount of land in each city is fixed. Individuals are identical in preferences and skills and each individual supplies one unit of labor independent of wage rate.[2] The utility function of a representative individual is a function of the land consumed (H), consumption of a composite commodity (or Hicksian bundle, denoted by C) and the amenities of a given location (s). The rental payment for land is denoted by r.

Maximizing this function subject to the budget constraint,

$$\text{Max } U(C, H;s) \text{ subject to } w = C + rH \qquad (2.1)$$

yields the indirect utility (denoted by V(.)) which is a function of the wages (w), rents (r) and the level of amenities (s). Equilibrium for households implies that the wages and

[2] Roback (1982) claims that the major conclusions are unchanged when this assumption is relaxed and in fact, she does extend her model to include two types of workers in a later paper (Roback, 1988).

rents in an economy must adjust to equalize utility across locations. Locational equilibrium for consumers thus requires

$$V(w,r;s)=k, \quad \text{where k is a constant.} \tag{2.2}$$

Roback assumes a constant returns to scale technology used by a representative firm to produce the only commodity in the economy. Thus, the equilibrium condition for the firm is that the unit cost of production must equal the product price (assumed to be unity) in all locations. Firm equilibrium is thus is represented by the following equation:

$$C(w, r; s) =1 \tag{2.3}$$

In equilibrium, the value of a marginal amenity change to a household may be obtained by totally differentiating equation (2.2). This yields

$$f_s = q\,(dr/ds) - (dw/ds), \quad \text{where q is the quantity land consumed ,} \tag{2.4}$$

Therefore, the marginal willingness to pay (WTP) for an amenity equals the sum of the slope of the hedonic wage function with respect to the amenity plus the slope of the hedonic property value function evaluated at the chosen amenity vector (adjusted for quantity of housing consumed). This is also known as "the full implicit price" of amenities. The value of locational amenities is thus inferred from hedonic wage and property value functions. Note that the signs of (dr/ds) and (dw/ds) and consequently, the sign of the full implicit price, will depend on both the productivity effects (C_s) and amenity effects (V_s).

Blomquist expands Roback's model to allow for amenity variation both within and across urban areas. The linkage between counties within an urban area occurs through an agglomeration effect: The population of an entire urban area affects production costs of firms regardless of which county they are located in. Incorporating agglomeration effects, implies that in general, the signs on both equations are ambiguous. The signs of (dr/ds) and (dw/ds) thus depend not only amenity and productivity effects but also on agglomeration effects. Therefore, in general, an amenity may have "unexpected" signs in both equations. A positive (negative) full implicit price represents a marginal net amenity (disamenity).

The main issue addressed in the Blomquist paper is that there is amenity variation within a city. This suggests that it is appropriate to use counties (as opposed to cities) as the unit of location choice to value amenities. For some amenities such as crime and quality of schools it is true that amenities vary significantly within a city. Some climate variables also vary across counties within an MSA. However, it is reasonable to assume that these differences are actually not large enough that households actually take these into account when choosing between two neighboring counties.[3] Thus, for valuing climate amenities, using Metropolitan Statistical Areas as the unit of location choice seems reasonable.[4]

---

[3] For example, households would not be likely to perceive the climate in PG county to be that different from Montgomery County!

[4] There are other issues with using county as the unit of location choice of course, such as data considerations -most of the data is not available by county but by MSA (5% sample of the 2000 PUMS and the Almanac data). Also households may choose to live in one county but commute to another. Thus they would consider the amenities associated with the entire area, not just the county.

Empirical Applications of Hedonic Models

Early attempts to estimate how much consumers will pay for more desirable climates relied on estimating hedonic wage and property value functions. Assuming a national labor market, wages in different cities should reflect differences in climate amenities, holding job and worker characteristics constant. The hedonic wage function relates the equilibrium wage to job characteristics, including occupation and industry; worker characteristics (e.g., education and years of experience); and locational amenities—climate, crime, air quality, proximity to the ocean. The hedonic property value function describes how, in equilibrium, housing prices vary across cities as a function of housing characteristics and locational amenities.

The following studies have employed the continuous hedonic approach, estimating either national hedonic wage and/or property value functions to value location-specific amenities. Hedonic wage and property value models have been estimated by Hoch and Drake (1974); Cropper and Arriaga-Salinas (1981); Cropper (1982); Roback (1982); Smith (1983); Blomquist et al. (1988); and Gyourko and Tracy (1991). The first three studies estimate only hedonic wage functions, while the last four estimate both wage and property value equations. As Moore (1998) and Gyourko and Tracy (1991) note, this literature suggests that climate amenities are capitalized to a greater extent in wages than in property values.[5] Roback (1982), Smith (1983) and Blomquist et al. (1988) all find sunshine to be capitalized in wages as an amenity, while

---

[5] The effect of climate variables on property values is mixed, with Blomquist et al. (1988) finding property values to be negatively correlated with precipitation, humidity and heating and cooling degree days, but Roback (1982) finding property values positively correlated with heating degree days. Gyourko and Tracy (1991) find heating and cooling degree days negatively correlated with housing expenditures, but humidity positively correlated.

heating degree days are capitalized as a disamenity (Gyourko and Tracy,1991; Roback,1982, 1988). Though the objective of Gyourko and Tracy (1991) study is to estimate Quality of Life Indices for cities in the US, it does provide estimate of WTP for climate amenities. [6]

2.1.2 Problems with using National Hedonic Models to Value Amenities

Unfortunately, hedonic wage and property value studies have limitations that have caused them to be replaced by alternate approaches to analyzing data on location choices. One drawback of the hedonic approach is that it assumes that national labor and housing markets exist and are in equilibrium. This is due to the underlying assumption that it is costless for households (and firms) to migrate to a different location. Moving costs could affect location decision problems in two ways.

First, they may drive a wedge between MWTP for an amenity and the slopes of the hedonic wage and property value equations. As Bayer, Keohane and Timmins (2006) (henceforth, BKT) show, if moving costs (denoted by M) matter to households, the indirect utility function and hence the equilibrium condition in (2.2) would be modified as follows:

$$V(w,r;s, M)=k \qquad (2.5)$$

The corresponding implicit price (compare with 1.4) is given by

[6] This paper incorporates the idea that city characteristics may include government services that are not pure amenities but have explicit tax prices. State and local taxes are included in the wage and rental regressions in order to estimate the full price of these services.

$$f_s = q\,(dr/ds) - (dw/ds) - (V_M/V_w)(dM/dX)\,, \quad (2.6)$$

Thus, the Roback model is a special case of this model if moving is costless (V_M=0) or mobility costs are constant (dM=0). If amenity increases (or decreases) with distance and moving is costly, housing and labor markets will undervalue (overvalue) amenities. An example of an amenity that increases with distance would be temperature – it becomes hotter as one moves from the North to the South. Thus, for climate variables, it is unlikely that dM =0. If moving costs could be measured exactly, then dM/ds could be estimated in the same way as the (dr/ds) and (dw/ds). Since moving costs are typically unobservable, it would be difficult to estimate equation (2.6). Thus a different empirical strategy is required.

Moving costs also affect the location that households (and firms) ultimately choose. For example, personal and family considerations may cause households to locate close to their birthplace. It is difficult to incorporate such costs into national hedonic models.

Secondly, moving costs would also affect the speed of adjustment in response to any exogenous shock. The fact that national labor and housing markets exist and are in equilibrium implies that households (and firms) move immediately in response to exogenous shocks and markets clear.[7] Realistically, such instantaneous adjustments may not occur due to impediments to migration such as transportation costs, search costs (for jobs and housing) and lack of availability of perfect information[8]. For firms, certain barriers to entry into a region may exist. Some examples of this would be permit

[7] Some examples of such shocks would be policies to lower pollution, changes in technology and immigration of people into the country.
[8] For example, information about changing conditions may not circulate instantaneously or it may take time for a person to find out the socio-economic characteristics of regions where he can potentially move.

requirements, long term contracts and transportation costs. Thus, adjustments are typically lagged and the economy at a particular point in time is not, in general, in equilibrium.

The Appendix describes in detail a simple model of consumer location in a two-city world with labor markets. A partial adjustment model is used to describe the process of adjustment of individuals. Population in the current period is a linear combination of population in the previous period and the equilibrium population. Equilibrium population is attained when differences in utility across regions are zero and thus no further migration is induced. The equilibrium values will be a function of the parameters of the system and also the value of the exogenous amenity.

Simulating a simple lagged adjustment model,[9] I obtain the following results. If I use the data from a period immediately after a shock and estimate a national hedonic model (while individuals are still in the process of adjusting) to obtain estimates of WTP for an amenity, the results are biased. The closer we are to the period of shock (or the further away from equilibrium), the larger the bias is. The sign and magnitude of the bias will also depend on the nature (positive or negative) and extent of the shock. If the shock is sufficiently large, the sign of willingness to pay may be wrong.

An additional problem plaguing the hedonic approach is omitted variable bias. Typically hedonic wage and property value regressions are estimated using a single cross section of data. Variables that are correlated with climate (e.g., the availability of recreational facilities) may be difficult to measure; hence, climate variables may pick up their effects. In hedonic property value studies, for example, the use of heating and cooling degree days to measure climate amenities is problematic because their

[9] A brief description of the model and simulation results are given in Appendix A.2.

coefficients may capture differences in construction and energy costs as well as climate amenities per se.

Another type of omitted variables problem in hedonic wage equations is that more able workers may locate in areas with more desirable climates. If ability is not adequately captured in the hedonic wage equation, the coefficients of climate amenities will reflect worker ability as well as the value of climate.

2.1.3 Discrete Choice Models

Cragg and Kahn (1997) were the first to relax the national land and labor market equilibrium assumption by estimating a discrete location choice model. Using Census data, they model the location decisions of people in the U.S. who moved between 1975 and 1980. Movers compare the utility they would receive from living in different states—which depends on the wage they would earn and on the cost of housing, as well as on climate amenities—and are assumed to choose the state that yields the highest utility. Formally, consumer i chooses his consumption bundle in location k (denoted by C_{ik}) to maximize his utility subject to a budget constraint,

Maximize $U(s_k, C_{ik} \mid X_i)$ subject to

$$C_{ik} = (1-T_k) * y_k (X_i, s_k) * Z_k (X_i) - r_k (X_i, s_k) \quad (2.7)$$

The unit of location choice used is states. Here T denotes state taxes, y weekly wages and z total weeks worked in the year. Thus (y*Z) is total wages earned in the year. s_k, as defined above represents the vector of amenities (or disamenities). Solving this problem yields an indirect utility function given by $V(S_k, y_k, Z_k, r_k, X_i)$. The location that yields the highest indirect utility is chosen.

The empirical estimation is motivated by a random utility model. Let

$$V_{ik} = V(s_k, y_{ik}, Z_{ik}, r_{ik}, X_i) + \varepsilon_{ik} \qquad (2.8)$$

Location k is chosen such that $V_{ik} > V_{ik}$ for all j and this is given by the conditional logit model (assuming that ε_{ik} are drawn from a Weibull distribution). Estimation of this model allows Cragg and Kahn to obtain estimates of the parameters of individuals' utility functions and thus infer the rate at which they trade income for climate amenities.

Unfortunately, the empirical estimates in this study are extremely large: The authors estimate, for example, that a non-college graduate between 50 and 60 would pay over $67,000 per year for a one standard deviation increase in mean February temperature!

One potential problem could be the fact that states are used as the unit of location choice in this paper. Though climate varies widely across states, it also varies within states. Other amenities (including education, pollution variables and crime) are omitted from their model since they are using states as the unit of choice. This may have resulted in omitted variables bias.

An alternate approach to modeling the location decisions of migrants is to acknowledge that moving is costly and to explain the location decisions of all households, assuming that all households are in equilibrium, given moving costs. Bayer, Keohane and Timmins (2006) use this approach to value air pollution.

They estimate a random utility model in which the indirect utility of a consumer i in location k is represented by $V(w_{ik},r_k,s_k, M_{ik})$[10]. A two-step approach is used to estimate the parameters of the utility function. The form of the utility function assumed implies that the log of the indirect utility can be written as

$$\ln V_{ik} = \beta_w \ln w_{ik} + M_{ik} + \theta_k \qquad (2.9)$$

$$\theta_k = -\beta_H \ln r_k + \beta_s \ln s_k \qquad (2.10)$$

Here, the β_s are the parameters to be estimated. In the first stage, location-specific constants (θ_k) are estimated together with other parameters of the utility function. These parameters include the coefficients corresponding to the wages and moving costs (which vary by consumer and location).

BKT use Census data from 1990 and 2000 to estimate this model. For each year, random samples of 10,000 household heads who are under the age of 35 and who live in one of 242 MSAs are used. The first stage model is estimated for the two years. Equation (2.10) is then estimated in first differences. Thus the change in the θ_k between 1990 and 2000 is regressed on the change in $\ln r_k$ and $\ln s_k$ to estimate the average utility attached to location-specific amenities.

---

[10] Note that the indirect utility function is different from the Cragg and Kahn model in two ways. First, moving costs have been included. The other distinction is that r_k, interpreted as the cost of housing here does not vary over consumers. It is assumed that expenditure on housing is the product of the cost of housing (r_k) in the area and the amount of housing (i.e., a vector of housing characteristics). A national housing hedonic is estimated to control for dwelling characteristics and the cost of living in each area is estimated by using location specific dummy variables.

Their study does not, however, include climate amenities as a determinant of location choice. This is because BKT estimate a differences model based on data from two consecutive decades. As it is hard to find significant variation in climate over such a short span of time, climate variables are not included. It thus appears that the discrete choice literature has yet to provide reliable estimates of the value of climate amenities in the U.S.

2.1.4 Migration Models

Another very closely related strand of economic literature involves studies of migration. The migration literature studies the importance of climate variables in migration. The purpose of these studies is not to value climate amenities per se but rather to explain who moves and why. Migration studies typically model population flows into regions (state, cities or counties). Specifically, this literature examines the roles of economic opportunities and amenities as determinants of migration rates in each region. Thus, in a typical model estimated in this literature, the migration rate into a region is regressed on wage rates and amenities in the region.

Some examples of migration studies are papers by Greenwood et al. (1991) and Mueser and Graves (1993). The Greenwood et al. (1991) study does not include climate amenities. However, the model in this paper does demonstrate that inferring the value of amenities from hedonic wage and property value studies can potentially lead to biased results. This is demonstrated as follows:

Greenwood et al. regress log of migration rates of the 50 states over the period 1971-1988 for each location on the logs of relative expected income (RY) of the location and a location specific constant (λ_a) . Relative expected income wage rates is defined as the ratio of wages rate in that area and the average wage rate in the economy. Mathematically

$$Ln(\text{migration rate into a location})_{a,t} = \ln\lambda_a + \lambda_1 RY_{a,t} + \text{error} \qquad (2.11)$$

λ_1 is expected to be positive. λ_a represents the effects on amenities in area a relative to other states. ln λ_a is negative for amenity poor states and positive for amenity rich states. This estimated λ_a is used to calculate estimates of the relative income that generates zero migration (RY^*). This is the value of relative income that compensates for the impact of the estimated λ_a. $RY^* < 1$ would imply that individuals are willing to accept lower wages to live in amenity rich states and vice versa. The gap between actual and estimated RY is the "disequilibrium gap" in this model. In amenity rich states, amenity valuations assuming equilibrium will overstate (understate) compensating differential when the actual RY is less than (greater than) RY^*. The reverse is true for amenity poor states. The paper also shows that almost all states were in disequilibrium during the time period considered in studies using national hedonic models (e.g. Blomquist et al, 1988). [11]

In Mueser and Graves, levels of migration are explained as a function of factors affecting migrant labor demand (e.g., "economic opportunity") and migrant labor supply (e.g., "residential amenities"). Climate variables are part of these residential amenities.[12]

[11] Another interesting result of the model is the classification of "attractiveness" of the states. Western (12 out of 13) and Southern (10 of 17) states seem to be more "amenity rich" than other regions of the country (4 of 9 in Northeast and 5 of 12 in Midwest).

[12] The paper concludes that there is no definitive answer as to whether economic opportunity is more important than amenities in determining migration levels. The answer depends on the time period being considered.

2.2 Different Measures of Climate Used in the Literature

Since the objective of this study is to estimate the value of climate amenities, it is useful to review which climate variables have been used in past empirical studies. First, however, it is important note the distinction between weather and climate. Climate is a long term phenomenon as opposed to weather, which may fluctuate substantially from year to year. To account for the fact that the weather in a particular area during a specific year may differ from "usual trends," the literature uses "climate normals." A climate normal is defined as the arithmetic mean of a climatological element computed over three consecutive decades (U.S. Climate Normals, 1971-2000, September 8, 2003).

Some measures of climate that are commonly used are average temperatures in a representative summer and winter month, precipitation, sunshine, humidity and wind speed. As an alternative to temperature, some studies use heating (HDD) and cooling degree days (CDD). Heating and cooling degree days are computed by the National Climatic Data Center as follows. First, the average of the high and low temperatures for the day is computed. If this is greater than 65 F, then the day is associated with (Average temperature - 65) cooling degree days. If the average temperature is less than 65 degrees, then the day is associated with (65 - Average temperature) heating degree days. HDD and CDD are thus functions of temperature. Both annual and seasonal measures of degree days have been used in the literature.

Several issues typically arise when trying to infer the value that households place on climate. First, it is essential to control for other amenities in the city to avoid omitted variables bias. Table 2.1 summarizes the climate variables and other amenities used in some of these studies. Second, estimates of the value of amenities can be potentially sensitive to specifications. Third, some amenities may be highly correlated and it may be difficult to interpret the results for such variables. Below, I describe what results have been obtained for climate variables in the studies described in the previous section.

Roback (1982) estimates a national hedonic wage equation using data for the 98 largest US cities and a housing hedonic equation based on 83 of those 98 cities. She finds that the values of amenities are very sensitive to specification. She notes that due to a small number of observations and high degree of multicollinearity between the variables, only a limited number of amenities could be used. Climate variables do not seem to be significant in housing regressions but are significant in the wage regressions.[13] Implicit prices also indicate that HDD, snowfall and cloudy days are disamenities and that number of clear days is an amenity.

In the paper by Blomquist et al. (1988), precipitation and sunshine have positive full implicit prices and are thus amenities and HDD, CDD, humidity and windspeed are disamenities. However, HDD and windspeed are not statistically significant.

In Gyourko (1991) and Tracy, the only climate variable that is significant at the 5% level is HDD, which is a disamenity. Precipitation, CDD and relative humidity are disamenities while sunshine and windspeed are amenities. However, these variables are not statistically significant.

Cragg and Kahn (1997) estimate a the probability of choosing a location for 3 age groups (30-40, 40-50, 50-60) and for two different education groups (at least a college

[13] Note, however, that what matters is the full implicit price. This is because Roback (1982) and Blomquist et al (1988) show that the signs on these coefficients are ambiguous if an amenity affects productivity and there are agglomeration effects.

degree, completed high school but did not complete college). They also use two specifications for the utility function – one that is linear in consumption and one that is quadratic in consumption of the Hicksian bundle. Higher February temperature and lower July temperature are amenities and are both statistically significant. They obtain mixed results for sunshine, and this variable is not always significant. Humidity is a disamenity for all models and is significant.

Mueser and Graves (1993) estimate net migration rates for metropolitan areas for three decades (1950 -1960, 1960 -1970 and 1970-1980) and they find that higher January temperatures induce migration while higher July temperatures reduce migration. This result is statistically significant for all three decades.[14]

Cushing (1987) attempts to figure out which climate variables prove to be most effective in explaining population movements in the US. This paper uses temperature variables to explain interstate migration flows in the US. In this paper, three alternative temperature measures are considered: average annual temperatures, heating and cooling degree days and average temperatures during the hottest and coldest month. Results indicate that the last one is the most important in explaining population movements in the US while the first performs the worst.

[14] They also note that due to high collinearity between variables (including the climate variables) standard errors were high. Thus, only a subset of variables has been used in each model.

Table 2.1. List of Climate Variables and Other Amenities in Other Studies

	Author	Dependent Variable	Climate Variables	Other Amenities	Type of Study
1	Roback (1982)	Log weekly earnings (N =12,001) Log of average residential site price per square foot (N=83)	Either one of • HDD • total snowfall • # of clear days • # of cloudy days	• Total crime rate • Particulate matter • Population • Population density • Percentage growth in population • % of person's neighborhood below the poverty line	National Wage and Housing Hedonic Equations
2	Blomquist et al 1988	Monthly housing expenditures (N=34,414) Average hourly earnings (N=46,004)	• Precipitation • Humidity • HDD • CDD • Wind speed • Sunshine	• Coast • Violent Crime • Teacher-Pupil Ratio • Visibility • TSP • NPDES Effluent Discharges • Landfill Waste • Superfund Sites • Treatment Storage and Disposal sites • Central City	National Wage and Housing Hedonic Equations
3	Gyourko and Tracy (1991)	Average weekly wages (N=38870) Annual housing expenditure (N=5263)	• Precipitation • CDD • HDD • Relative humidity • Sunshine • Wind speed	• Particulate matter • Coast • Non-land cost of living • SMSA population • % working in other SMSA • Violent crime rate • Student/teacher ratio • Rating of fire department quality • Hospital beds • Effective property tax rate • State and local income tax rate • State corporate income tax rate • % public union organized	National Wage and Housing Hedonic Equations

	Author	Dependent Variable	Climate Variables	Other Amenities	Type of Study
4	Cragg and Kahn	Probability that location j is chosen (N=26,988)	• Average rainfall in 1980's • Average February temperature • Average July temperature • Sunshine • Humidity	• Significant coastal beach • Number of cities	Discrete Choice Model
5	Mueser and Graves (1993)	Net migration rate	• Average January and July temperatures • Average January and July humidity • Sunlight • Annual precipitation[15]	• % of area covered by lakes • Metropolitan dummy • Population Density minus metropolitan mean • Population Density minus nonmetropolitan mean • Distance to major urban area • % black • Median education • Census divisions • Local departure propensity • Measures of Industrial composition • State capital dummy	Migration Model

[15] The authors note that they use a subset of the independent variables listed above in each model since, due to high collinearity, standard errors were high when all of them were used together.

Chapter 3: Model and Empirical Specification

The goal of this dissertation is to estimate the value placed on climate amenities using a discrete model of location choice. I model households as selecting their preferred metropolitan area (MSA) from a set of 297 MSAs in the United States in 2000. Household utility depends on housing, on location-specific amenities, and on expenditure on all other goods (income minus the cost of housing). In the econometric model, households select among locations based on the indirect utility they receive from each location.

First, I estimate this model using data on households who changed locations between 1995 and 2000. I focus on migrants because the assumption that households are in locational equilibrium is more reasonable for migrants than for all households. It is reasonable to assume that these households have overcome the issue of moving costs and after their move are in equilibrium. [16]

However, movers differ from stayers with respect to a number of characteristics.[17] So, it might be argued that the preferences of the movers may not be representative of the entire population. Therefore, I also estimate this model for both movers and stayers for the purposes of comparison.

I begin by describing the household utility maximization problem in Section 3.1. Section 3.2 presents the empirical specification.

[16] Also, a problem of endogeneity between the location decision of households and wages and rents may exist. It may be reasonable to assume that migrants are too "small" to affect the wages and rents. Thus, the problem of endogeneity, while it may not have been taken care of completely, is at least reduced.

[17] These are presented in the next chapter in Table 4.1.

3.1 Household Utility Maximization

Each household decides how to optimally allocate its income between housing expenditure and all other goods, and also chooses the location in which to live that yields the highest possible utility. I first describe the budget allocation problem in Section 3.1.1 and then the location decision problem in Section 3.1.2.

3.1.1 Budget Allocation Problem

Each household chooses the quantity of consumption of a numeraire good and housing to maximize its utility subject to a budget constraint. Mathematically, the utility maximization problem of household i living in location j is given as follows:

Choose $\{C_{ij}, H_{ij}\}$ to maximize $U(C_{ij}, H_{ij}, MC_{ij}, E_j)$

subject to the budget constraint $C_{ij} + R_j H_{ij} = \sum_{m=1}^{Ni} w_{mj}$ (3.1)

where

m ≡ individual

i ≡ household

j ≡ location

N_i ≡ number of household members in household i

C_{ij} ≡ Consumption of a numeraire good by household i living in location j

H_{ij} ≡ Quantity of housing consumed by household i living in location j

R_j ≡ Cost of housing in location j

MC_{ij} ≡ Moving costs of household i to location j

w_{mj} ≡ Wages earned by an individual m when living in location j

E_j ≡ Vector of Amenities (e.g. climate) and disamenities (e.g. pollution, crime, etc) in location j

Solving the utility maximization problem in (1), I obtain the optimal values of consumption and housing expenditure, C_{ij}^* and H_{ij}^*. Substituting these values into the utility function yields i's utility from MSA j,

$$V_{ij} = V(W_{ij}, MC_{ij}, R_j, E_j) \quad (3.2)$$

Here, W_{ij} represents the total household wages of household i in location j.

i.e. $W_{ij} = \sum_{m=1}^{Ni} w_{mj}$

3.1.2 Random Utility Model (RUM) and Migration Equation

I now turn to the location decision of households. To model this, I use the Random Utility Model (McFadden, 1973) which assumes that the indirect utility of a household is known up to an error term. Mathematically, the random utility of household i from living in location j is given by:

$$V_{ij} = V(W_{ij}, MC_{ij}, R_j, E_j) + \varepsilon_{ij} \quad (3.3)$$

where

ε_{ij} ≡ idiosyncratic error

V(.) $\equiv$ deterministic component of the utility function

Assuming that the idiosyncratic errors (ε_{ij}) are i.i.d. Type I Extreme Value, the probability of household i migrating to region j is given by the Conditional Logit Model[18]:

$$\Pr(V_{ij} \geq V_{ik}, \forall k \neq j) = \frac{e^{V(W_{ij}, MC_{ij}, R_j, E_j)}}{\sum_{k=1}^{K} e^{V(W_{ik}, MC_{ik}, R_k, E_k)}} \tag{3.4}$$

where K = number of alternatives. This is the household migration equation.

3.2 Empirical Specification

I begin by defining the functional form of the utility function in Section 3.2.1. Next, I define the moving costs and the migration equation. In Section 3.2.2, I describe the functional form of the wage and housing hedonic.

3.2.1 Utility Function

I assume that the form of the utility function is Cobb Douglas. The utility function can therefore be written as

$$U_{ij} = C_{ij}^{\alpha_C} H_{ij}^{\alpha_H} e^{MC_{ij}} e^{g(E_j)} \tag{3.5}$$

[18] This does impose the assumption of the Independence of Irrelevant Alternatives (or IIA). However, using a more general error structure (e.g by using a nested logit model or a random parameters model) would involve huge computational costs.

The form of the function g(.) depends on what is assumed about the preferences for the amenities. For example, it might be reasonable to assume that there is an optimal temperature that households prefer. A quadratic form for g(.) would capture this. In the empirical implementation of this model, I present results using different functional forms for g(.).

Maximizing the utility subject to the budget constraint yields the indirect utility function,

$$V(W_{ij}, MC_{ij}, R_j, E_j) = \left(\frac{\alpha_C}{\alpha_C+\alpha_H}\right)^{\alpha_C} \left(\frac{\alpha_H}{\alpha_C+\alpha_H}\right)^{\alpha_H} W_{ij}^{(\alpha_C+\alpha_H)} e^{MC_{ij}} \left(\frac{1}{R_j}\right)^{\alpha_H} e^{g(E_j)} \tag{3.6}$$

Marginal willingness to pay for an amenity by a household is given by the marginal rate of substitution between the amenity and income. For example, if we assume that $g(E_j) = \alpha_E \ln E_j$ then the MWTP of a household i for climate amenity E is $(\alpha_E / \alpha_C + \alpha_H)^* \ W_{ij} / E_j$. In the remainder of the dissertation, I focus on estimating this marginal rate of substitution. Calculating a complete welfare measure would entail looking at the impact of a change in the vector of amenities on expected household utility, as is usually done in a random utility framework (Freeman, 1993). However, this is not the focus of this study.

Moving Costs

Moving costs are defined, following Bayer et al. (2006), as follows:

$$MC_{ij} = \alpha_{M0} d_{ij}^{State} + \alpha_{M1} d_{ij}^{Division} + \alpha_{M2} d_{ij}^{Region} \tag{3.7}$$

where d_{ij}^{State} denotes a dummy variable that equals one if location j differs from the state in which household i lived in 1995. $d_{ij}^{Division}$ = 1 if location j is outside of the Census Division in which household j lived in 1995, and $d_{ij}^{\mathrm{Re}gion}$ = 1 if location j lies in a different Census Region than the one is which household i lived in 1995. The use of the moving cost dummies captures both physical moving costs and the psychology and information costs of moving.[19]

Migration Equation

The logarithm of the systematic portion of the indirect utility function can be written as

$$\ln(V(W_{ij}, MC_{ij}, R_j, E_j))$$

$$= \alpha_C \ln(\frac{\alpha_C}{\alpha_C + \alpha_H}) + \alpha_H \ln(\frac{\alpha_H}{\alpha_C + \alpha_H}) + (\alpha_C + \alpha_H)\ln W_{ij} + MC_{ij} - \alpha_H \ln(R_j) + g(E_j) \tag{3.8}$$

implying that the migration equation in log form is as follows:

$$\Pr(\ln V_{ij} \geq \ln V_{ik}, \forall k \neq j) = \frac{e^{(\alpha_C + \alpha_H)\ln W_{ij} + MC_{ij} - \alpha_H \ln(R_j) + g(E_j)}}{\sum_{k=1}^{K} e^{(\alpha_C + \alpha_H)\ln W_{ik} + MC_{ik} - \alpha_H \ln(R_k) g(E_k)}} \tag{3.9}$$

[19] An alternative specification would be to allow moving costs to be a function of distance. However, my goal is to capture more than transportation costs.

I replace all the variables that vary only by MSA by a location specific intercept A_j,

$$A_j = -\alpha_H \ln(R_j) + g(E_j) \tag{3.10}$$

Therefore, $\Pr(\ln V_{ij} \geq \ln V_{ik}, \forall k \neq j)$ reduces to

$$= \frac{e^{(\alpha_C+\alpha_H)\ln W_{ij} + MC_{ij} + A_j}}{\sum_{k=1}^{K} e^{(\alpha_C+\alpha_H)\ln W_{ij} + MC_{ij} + A_j}} . \tag{3.11}$$

To estimate the migration equation requires information on the wages that a household would earn and the cost of housing in all possible locations; however, wages and housing costs are observed only in the household's chosen location. I therefore estimate these for all possible locations. Having replaced the estimated values of W_{ij} and R_j in migration equation, I estimate the migration equation (given by equation 3.11) using maximum likelihood techniques. This gives estimates of the location specific intercepts A_j.

In the second stage, the goal is to regress the MSA-specific fixed effects, A_j on R_j and location specific amenities to obtain the parameters of equation (3.10). The left hand side of this equation represents the average indirect utility from MSA j after controlling for household income and moving costs. Because living costs are likely to be correlated with the error term[20], α_H is set equal to 0.25 (which is the median share of income spent on housing in my sample) and $\alpha_H R_j$ is added to the dependent variable. Thus, in the second stage, I estimate the following equation:

$$(A_j + 0.25 R_j) = g(E_j) + \eta_j \tag{3.12}$$

In reality, α_H varies across MSAs. However, the share of income spent on housing is a function of prices in the MSA. To incorporate this idea, a more flexible functional form than the Cobb Douglas is required. This is beyond the scope of this dissertation.

3.2.2 Predicting Wages and Housing Costs

Wages

To estimate $W_{ij} \forall j$, I estimate a hedonic wage equation for each of the 297 MSAs in my sample. In this way I avoid making one of the standard but restrictive assumptions of traditional hedonic models, viz., the existence of a national labor market. The hedonic wage functions are of the following form:

$$\begin{aligned}\ln(WageRate_i) &= \beta_0 + \beta_D DemographicCharacteristics + \beta_{ED} EducationVariables \\ &+ \beta_{EXP} ExperienceVariables + error\end{aligned} \tag{3.13}$$

The dependent variable in this equation is the log of the hourly wage rate of each individual. I use the coefficients from these hedonic equations to calculate the wage rates for each individual in each location. I then use the product of these estimated wage rates

[20] The correlations between all the location-specific variables are presented in Table 4.2.

and the total hours each individual works in a year to estimate the individual's wages for all locations. Summing these over all individuals in each household, I obtain household wages for all locations. In doing this, I abstract from modeling labor-leisure choice decisions and make the simplifying assumption that individuals work the same number of hours and number of weeks in any location.

To estimate the hedonic wage equations, I use the following exclusion criteria. Self employed individuals or those who report working in agriculture, farming, fishing or forestry are not included in the sample. These individuals would likely have different considerations when making location choices (and consequently, different moving costs) than an "average" household. I also exclude military personnel (those who were enrolled from 1995 to 2000) and handicapped individuals (defined as persons having difficulty working). Following common practice in labor economics, I do not include part time and part year workers in the sample. This is because hourly wage rates for such workers are often very noisy. I define full time workers as those who work at least 30 hours per week. I use two different definitions of full year workers: those who work more than 30 weeks and those who work more than 40 weeks. I also delete observations if hours worked are more than 60. This is to avoid including individuals who have multiple jobs.[21]

Housing Costs

[21] I do this because the Census asks respondents to describe the job at which the person worked the most hours. As the answer to this question forms the basis for the occupation and industry dummies, if anyone worked at multiple jobs I would incorrectly attribute earnings from the second job to the first one and thus bias the occupation and/or industry dummies.

In order to impute the housing costs that each household would face in each location, I estimate an index for the cost of housing in each MSA (i.e. $R_j \forall j$). This is accomplished by estimating an hedonic housing equation controlling for dwelling characteristics and using dummy variables for each of the MSAs. If I were to estimate a separate equation for each metropolitan area, I would have to make an assumption about the housing bundle consumed by each household in each area to predict housing expenditure for a household in each city. The housing price index approach is much cleaner.

Ideally I would like to estimate separate equations for owner and rental markets since supply conditions in these two markets differ. I would then need to predict the probability that a person would buy or rent when moving to a new location. I therefore ignore the rent/own distinction and estimate a national hedonic housing market equation that pools observations from the owner and rental markets. The MSA dummies from this equation constitute the $\{R_j\}$.

I first calculate the user costs of owning a house in order to make them comparable with rents. User cost is calculated as the sum of mortgage payments, property taxes and insurance. Assuming a 30 year Fixed Rate Mortgage (FRM), mortgage payments are calculated using the following amortization formula:

$$\text{Monthly Payment} = \frac{\text{P}*(\text{r}/12)}{1-(1+(r/12))^{(-n)}} \tag{3.14}$$

Here,

P is the value of the house reported by the household[22]

[22] The housing price literature shows that shows that biases from self appraisals of value of a home are small (e.g. Follain and Malpezzi (1981) and Goodman and Ittner (1992))

r is the annual interest rate (or Fixed Rate Mortgage)[23]

n = the number of periods over which the housing loan is paid

User costs should include expected house price appreciation. However, this was excluded because it was difficult to calculate. Utility costs were added to user costs because a major portion of utility costs are due to heating and cooling requirements. Such costs need to be separated from climate amenities. To address this issue, I add utilities to user costs. Utility costs were calculated as the sum of amount paid for electricity, gas, water and other fuels by households. This is fuel for household purposes, not for transportation. To calculate insurance costs I use the insurance variable in the Census data, which is the amount that households pay annually for insurance against fire, hazard and flood damage. The amount of property taxes reported by households in the Census was also added to user costs.

The dependent variable in the housing model is the logarithm of user cost - the sum of the monthly mortgage payment or rent, utilities, taxes and insurance.

$$\ln(user\cos t_i) = R_0 + R_{OWN} OwnershipDummies + R_{BR} BedroomDummies$$

$$+ R_R RoomDummies + R_{KIT} KitchenDummy + R_P PlumbingDummy$$
$$+ R_{ACRE} Acreagedummies + R_{AGE} AgeofStructureDummies$$
$$+ R_{UNITS} UnitsStructureDummies + R_{MSA} MSADummies + error \qquad (3.15)$$

[23] I am using the FRM from Freddie Mac's Primary Mortgage Market Survey. These are weekly rates, so I calculate the average annual value. The average for the 30 year FRM is 7.72% and for 15 Year FRM is 8.05%. I have used 8% in the model results reported here. Here n =30*12.

The sample used to estimate this equation consists of all houses excluding farms, mobile homes and boats occupied by households in the PUMS. The key objective here is to obtain unbiased estimates of { R_{MSA} }.

The characteristics of the house used in the hedonic regression are very basic and are by no means an exhaustive list of variables that affect the value of a home. However, Malpezzi et al. (1998) demonstrates that while parsimony (omission of relevant variables) will be a problem if the goal is to estimate implicit prices, [24] it is not a problem if the goal is to predict house prices. Thus, using this specification should result in unbiased estimates of the costs of living across different MSAs.

It should be noted that amenities are not a part of the right hand side variables in either the wage or the housing hedonic equations as I do not want climate to be capitalized into wage and property values. A more detailed list of the RHS variables in the wage and housing equations is provided in the next chapter.

[24] This also emphasizes to need to include more detailed characteristics of the house if one were to estimate hedonic housing models to value amenities.

Chapter 4: Data and Stylized Facts

The data used to estimate the migration model come from the 5% Public Use Microdata sample (PUMS) of the 2000 Census[25], as well as other publicly available data sources. The PUMS contains data on the locations of households in 2000 and 1995 as well as data on household and individual characteristics. Section 4.1 discusses the Census data and presents some stylized facts about migration patterns over the 1995 to 2000 period, as well as information about spatial variation in wages and housing prices.

Data on location-specific amenities, including climate, air pollution, crime and quality of transportation, education, recreation, arts and healthcare services, come from a variety of sources. Section 4.2 briefly discusses these data and presents some summary statistics.

4.1 Census Data

In this study, I model the location decisions of U.S. households who moved between 1995 and 2000 and who lived in one of 297 Metropolitan Statistical Areas (MSAs) in 2000. Figure 4.1 illustrates the proportion of the US population residing in metropolitan areas over the past century.[26] This figure shows that this proportion has been growing over time and that over 80% of the US population lived in MSAs in 2000. Sections 4.1.1 and 4.1.2 describe the variables used for the wage and housing hedonic

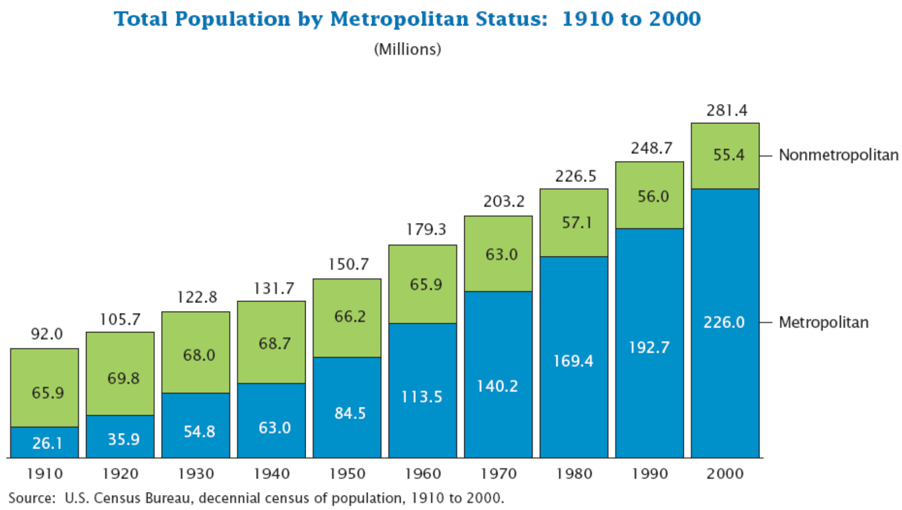
---

[25] The 5% PUMS data are publicly available from the U.S. Census bureau (www.Census.gov), or at http://usa.ipums.org/usa/index.shtml .

[26] Figure A4.1 in the Appendix shows the proportion of population living in metropolitan areas by state.

price functions. Section 4.1.3 describes migration patterns of households, while Section 4.1.4 compares the characteristics of migrant households with those of non-migrants.

Figure 4.1 ***(Source: US Census Bureau)***



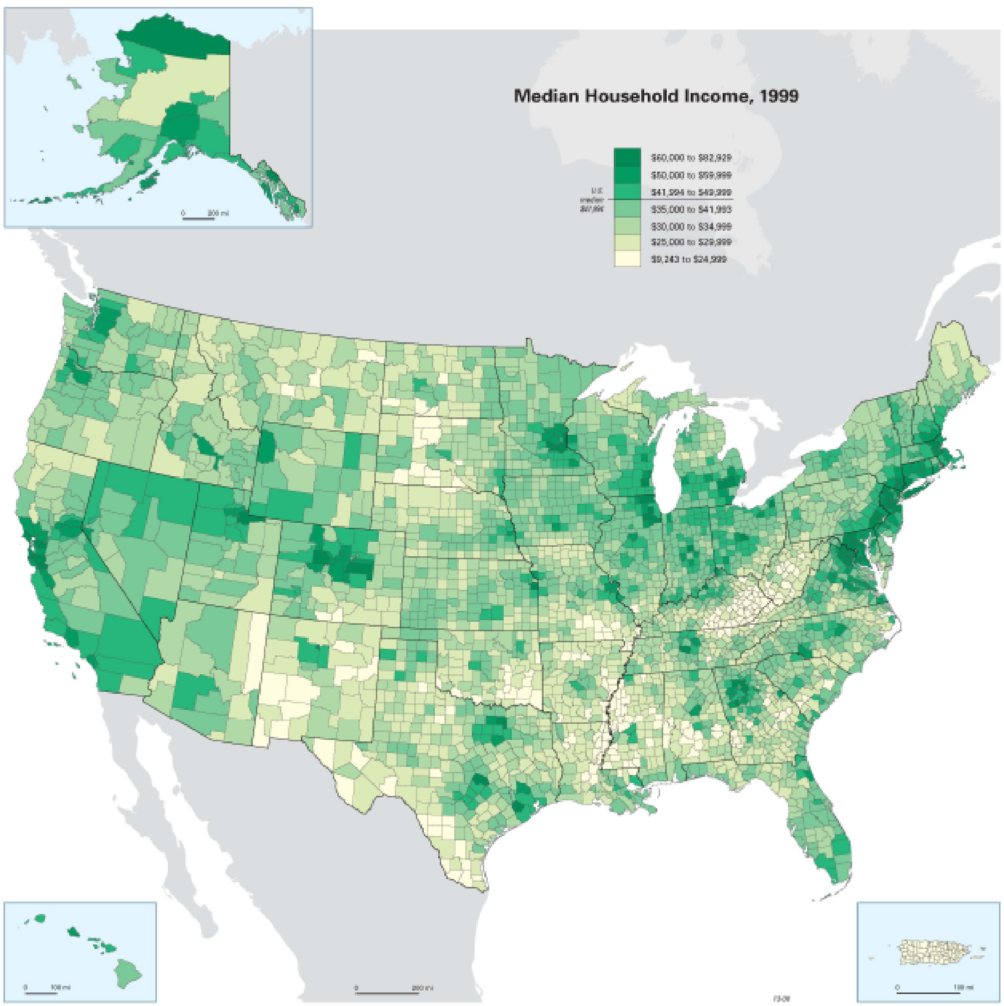
4.1.1 Wages

The PUMS data also contain information on the labor force participation, hours and earnings of individuals, as well as their occupation and the industry in which they worked. This dataset also includes variables describing the demographic characteristics of individuals such as race, age, gender and marital status. A complete list of these variables and their means is included in the Appendix (Table A4.2). Variables measuring the quality of human capital, such as education and whether an individual speaks English are also part of this dataset. These data are used to estimate hedonic wage equations for 297 MSAs. Figure 4.2 presents the median household income by counties.[27] On an average, wage earnings are roughly 75% of a household's income. Throughout this

[27] Summary statistics of the wage income variable of individuals by MSA are presented in Table A4.1 of the Appendix and Figure A4.2 (in the Appendix) shows the median household income by state.

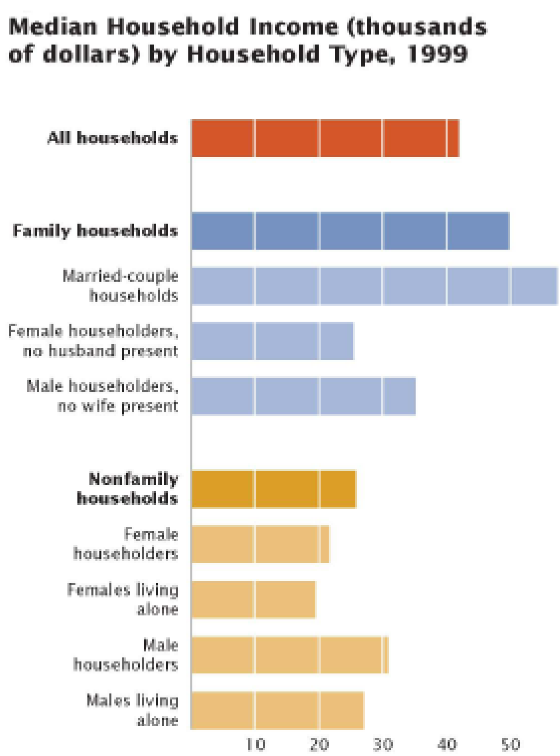
section, I present figures showing income since figures showing just wage earnings are not easily available. Figure 4.2 shows a huge variation in income across the counties. It should also be noted that median household income was on average, higher in metropolitan areas than in nonmetropolitan counties.

Figure 4.2 ***(Source: US Census Bureau)***



Since I am modeling the migration decisions of households (as opposed to single individuals) and estimating a household utility function, it is appropriate to use total household earnings in determining the budget constraint faced by households. These household earnings vary by type of the household. For example, married couple households earn more on average due to the presence of multiple earners and male householders earn more. Figure 4.3 shows the median household earnings by type of household.

Figure 4.3 ***(Source: US Census Bureau)***

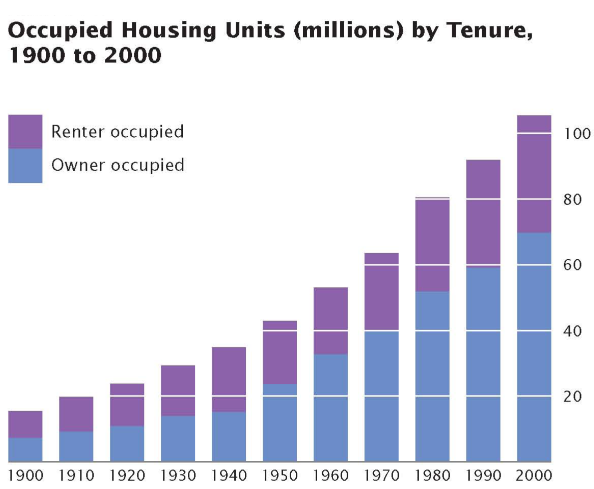


4.1.2 Housing

Information on housing costs (such as the value of owner occupied housing and rents[28], costs of utilities, insurance and property taxes) and characteristics (such as number of rooms and bedrooms), for estimation of the housing hedonic equation, are also taken from the PUMS. A complete list of these variables and their means is included in the Appendix (Table A4.3).

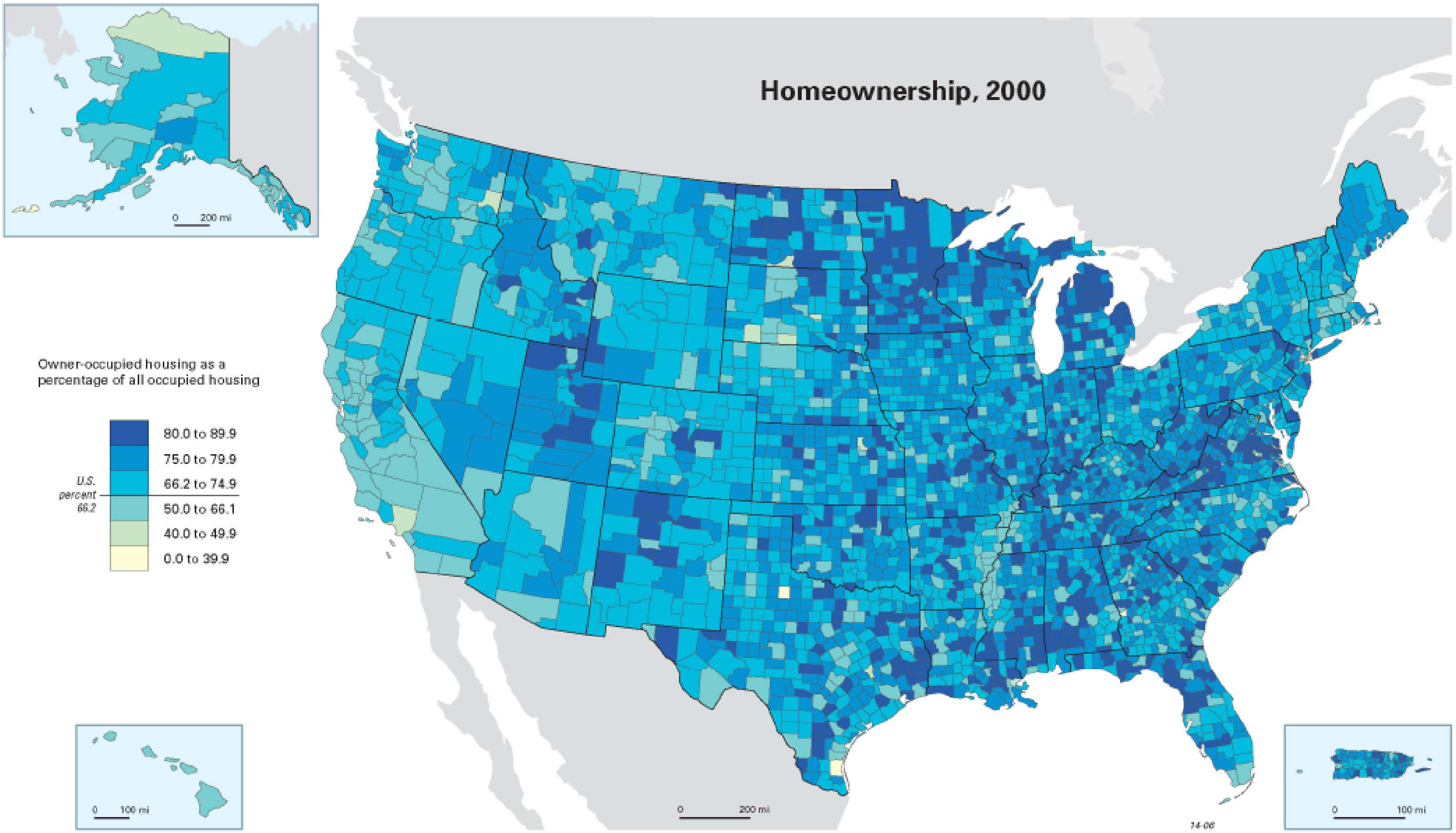
A variable describing whether a household owns or rents their dwelling is part of the Census dataset. The average proportion of owner-occupied to renter-occupied housing has been increasing (shown in Figure 4.4). This proportion also varies widely across the country (shown in Figure 4.5). [29] This could be due to differences in monthly costs between renters and owners (shown in Figure 4.6) and/or due to differences in income (Figure 4.2).[30]

Figure 4.4 ***(Source: US Census Bureau)***



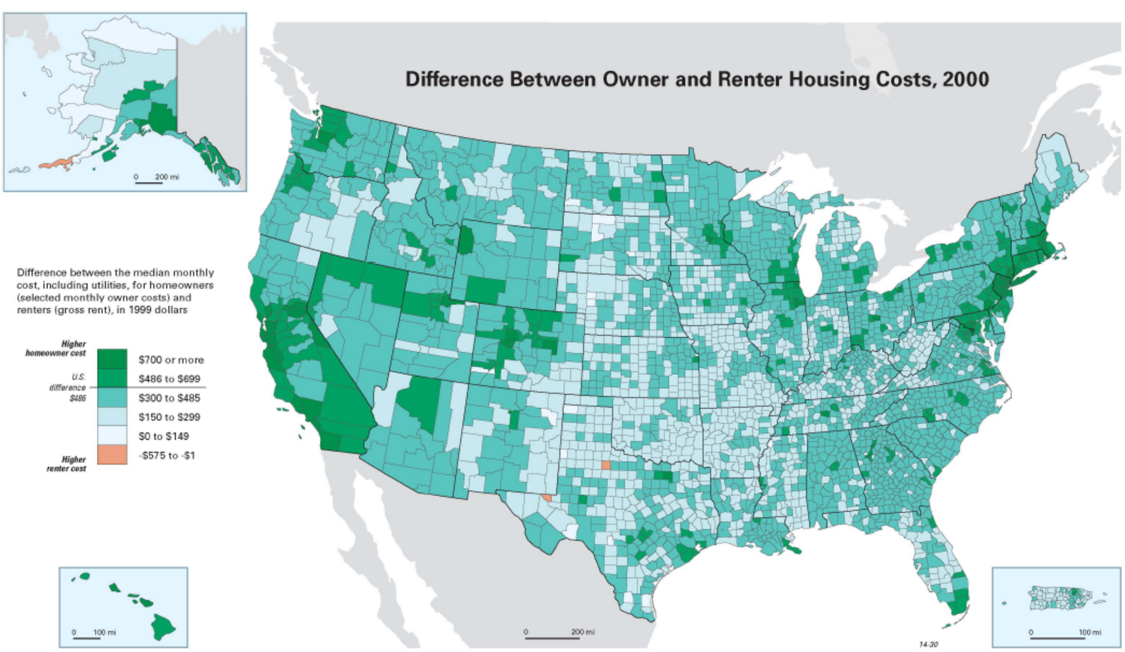
[28] Figure A45, A4.6 and A4.7 (in the Appendix) display these.
[29] Figure A4.3 (in the Appendix) shows homeownership by state.
[30] The value of the house relative to income varies across the country as Figure A4.4 (in the Appendix) illustrates. However, this does not seem to be perfectly correlated with the ownership rates (compare A4. 3 with A4.4).

Figure 4.5 ***(Source: US Census Bureau)***



If there were information that allowed an econometrician to predict the probability that a household would own or rent in different locations, it would be appropriate to estimate separate housing hedonic equations for owners and renters. In that case one could calculate different cost of living indices for owners and renters and multiply each by the probability that a household would own or rent, respectively. In the absence of such information, however, estimating separate hedonic equations for owners and renters would force one to assume that a household would always rent (or own) regardless of what MSA they chose to locate in. As shown by the variation in ownership rates (in Figure 4.5), this assumption would not be very reasonable. I therefore use a cost of living index that does not distinguish between owners and renters.

Figure 4.6 ***(Source: US Census Bureau)***



4.1.3 Migration Patterns

The PUMS contains information on over 5.6 million households. Tables 4.1 and 4.2 below describe the households who changed MSAs between 1995 and 2000, for whom both the origin and destination MSA can be identified.[31] Of these 441,393 households, 60.8% moved to a different state and 46.9% moved to a different Census[32] division. Thirty-six percent moved to a different Census region. Below, population movements are presented at these three geographical levels – movements by region, division and state.

Table 4.1: Origin and Destination of Migrants by Census Region

Region (1995)	Region (2000)				Total
	Midwest	Northeast	South	West	
Midwest	38865 (8.81%)	5129 (1.16%)	20905 (4.74%)	13176 (2.99%)	**78075 (17.69%)**
Northeast	5230 (1.18%)	55513 (12.58%)	29725 (6.73%)	10223 (2.32%)	**100691 (22.81%)**
South	11712 (2.65%)	11906 (2.7%)	99761 (22.6%)	17259 (3.91%)	**140638 (31.86%)**
West	7707 (1.75%)	5831 (1.32%)	18663 (4.23%)	89788 (20.34%)	**121989 (27.64%)**
Total	**63514 (14.39%)**	**78379 (17.76%)**	**169054 (38.3%)**	**130446 (29.55%)**	**441393**

[31] Of the 5.66 million households in the PUMS, 1.53 million lived in named MSAs in both 1995 and 2000. Twenty-eight percent of these households changed location between 1995 and 2000. A household was considered to have moved if the head of household moved.
[32] Figure 4.12 shows the Census regions and divisions on a map.

Table 4.1 shows the origin and destination of households by Census region. Over 60% of these households moved within the region in which they lived in 1995. Twenty three percent of the households who moved between 1995 and 2000 remained in the South; 20% were living in the West. In contrast, only about 10% of the movers who lived in the Northeast or Midwest in 1995 remained in their region of origin. On net, household left the Northeast and Midwest for the South and West. This has been the general trend in population movement over the century as illustrated by Figure 4.7. Migration patterns during the period 1995 to 2000 are shown on a map of the US in Figure 4.8.

Figure 4.7 ***(Source: US Census Bureau)***

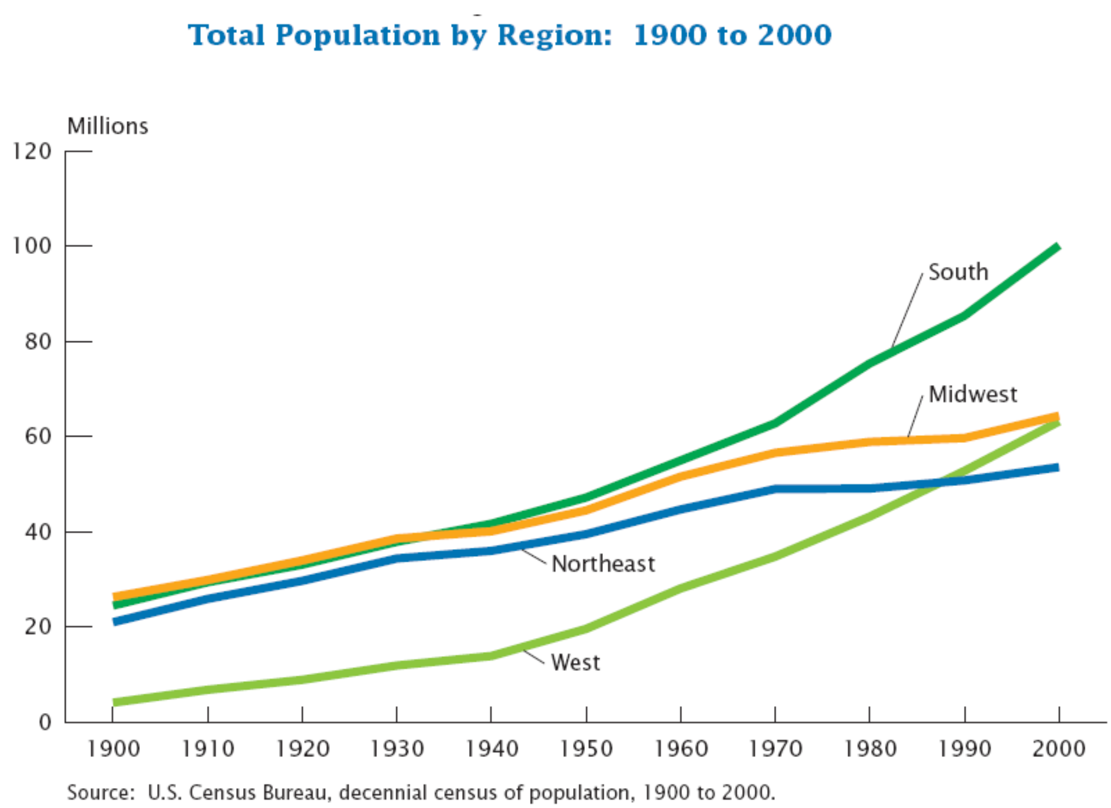


Figure 4.8 ***(Source: US Census Bureau)***

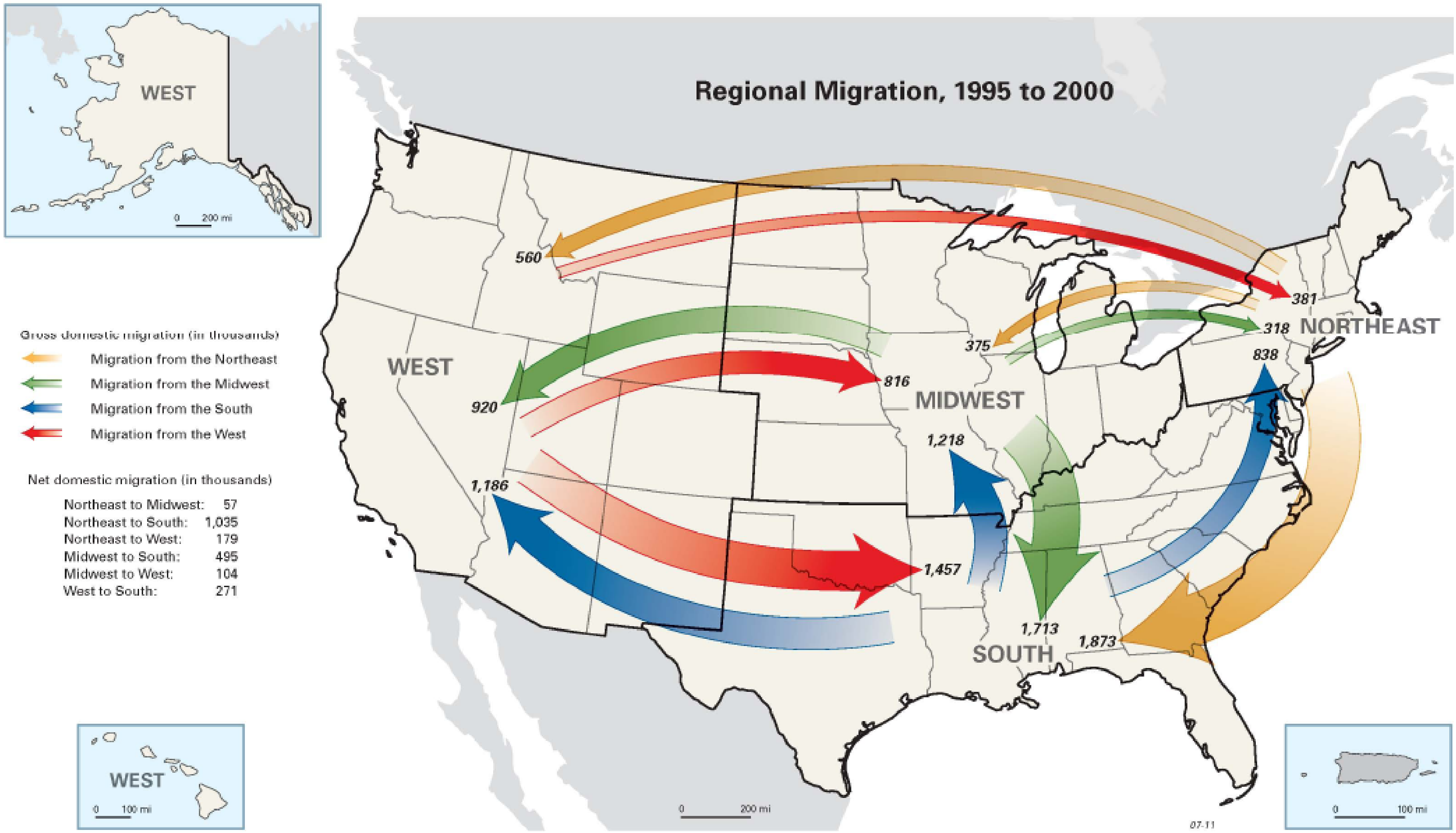
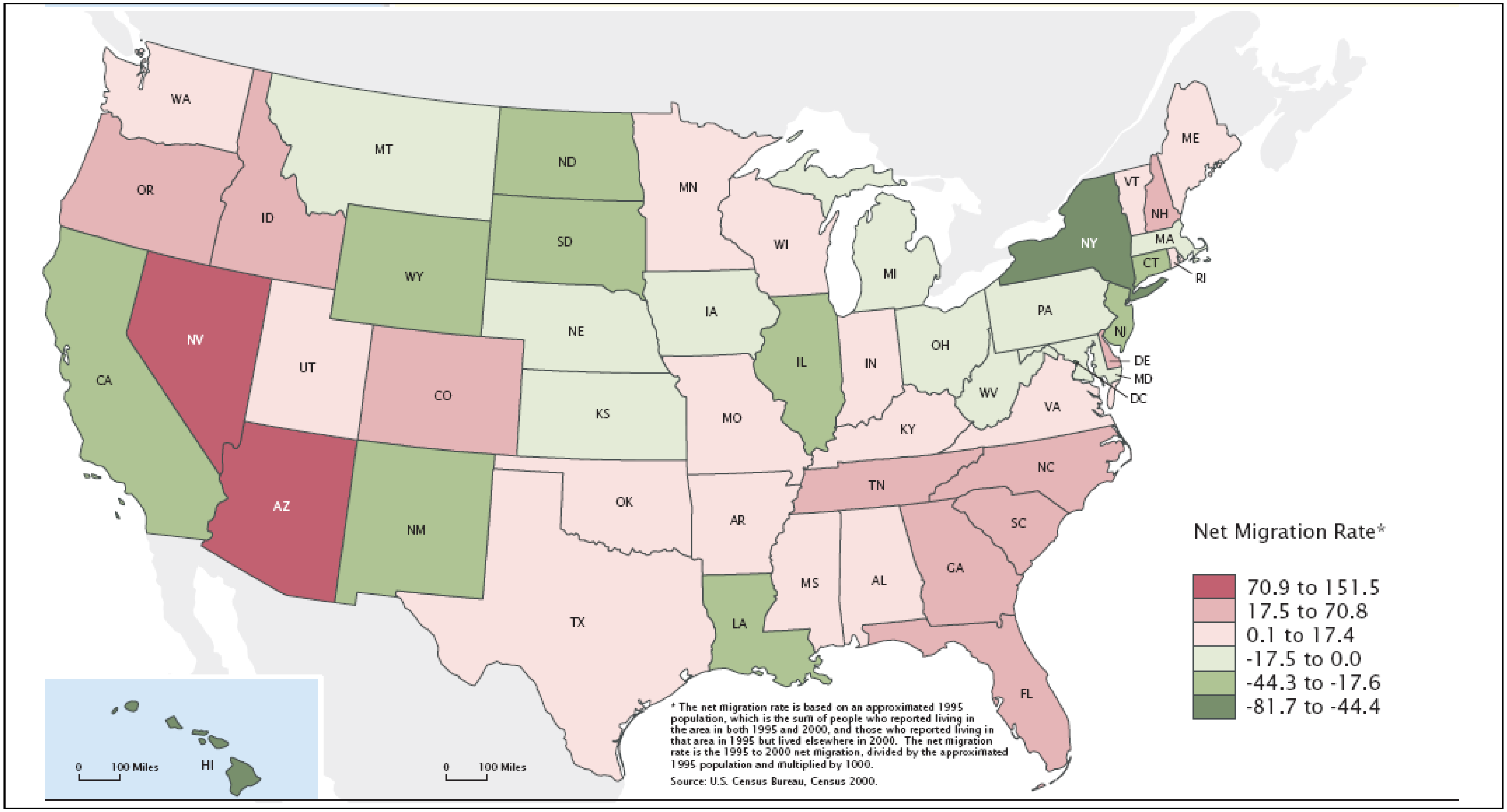


Table 4.2: Origin and Destination of Migrants by Census Division

Division (1995)	Division (2000)									Total
	East North Central	East South Central	Middle Atlantic	Mountain	New England	Pacific	South Atlantic	West North Central	West South Central	
East North Central	29374 (6.65%)	2109 (0.48%)	3108 (0.7%)	4137 (0.94%)	1046 (0.24%)	5093 (1.15%)	11037 (2.5%)	2370 (0.54%)	3070 (0.7%)	**61344 (13.9%)**
East South Central	1357 (0.31%)	5830 (1.32%)	530 (0.12%)	545 (0.12%)	169 (0.04%)	963 (0.22%)	4077 (0.92%)	323 (0.07%)	1421 (0.32%)	**15215 (3.45%)**
Middle Atlantic	3556 (0.81%)	967 (0.22%)	39258 (8.89%)	2673 (0.61%)	4731 (1.07%)	5062 (1.15%)	21589 (4.89%)	649 (0.15%)	2058 (0.47%)	**80543 (18.25%)**
Mountain	1338 (0.3%)	364 (0.08%)	848 (0.19%)	6074 (1.38%)	295 (0.07%)	5253 (1.19%)	2154 (0.49%)	673 (0.15%)	1973 (0.45%)	**18972 (4.3%)**
New England	849 (0.19%)	220 (0.05%)	2696 (0.61%)	694 (0.16%)	8828 (2%)	1794 (0.41%)	4331 (0.98%)	176 (0.04%)	560 (0.13%)	**20148 (4.56%)**
Pacific	3899 (0.88%)	1215 (0.28%)	3157 (0.72%)	11264 (2.55%)	1531 (0.35%)	67197 (15.22%)	8023 (1.82%)	1797 (0.41%)	4934 (1.12%)	**103017 (23.34%)**
South Atlantic	5552 (1.26%)	4043 (0.92%)	7286 (1.65%)	3057 (0.69%)	2205 (0.5%)	6176 (1.4%)	49156 (11.14%)	1341 (0.3%)	4591 (1.04%)	**83407 (18.9%)**
West North Central	2319 (0.53%)	512 (0.12%)	681 (0.15%)	1962 (0.44%)	294 (0.07%)	1984 (0.45%)	2412 (0.55%)	4802 (1.09%)	1765 (0.4%)	**16731 (3.79%)**
West South Central	2016 (0.46%)	1510 (0.34%)	1232 (0.28%)	2680 (0.61%)	484 (0.11%)	3838 (0.87%)	5444 (1.23%)	1123 (0.25%)	23689 (5.37%)	**42016 (9.52%)**
Total	**50260 (11.39%)**	**16770 (3.8%)**	**58796 (13.32%)**	**33086 (7.5%)**	**19583 (4.44%)**	**97360 (22.06%)**	**108223 (24.52%)**	**13254 (3%)**	**44061 (9.98%)**	**441393**

Figure 4.9 Net Domestic Migration Rates by State, 1995 to 2000 ***(Source: US Census Bureau)***



Movements of the population are broken down in more detail in Table 4.2, which shows the origin and destination of households by Census division. The populations in the Mountain and South Atlantic divisions have increased the most. The populations in the East North Central and Middle Atlantic divisions have fallen between 1995 and 2000. Net domestic migration rates by state for 1995 to 2000 are shown in Figure 4.9. The maximum out-migration has been from the state of New York while the maximum in-migration has been in Nevada and Arizona. Figure A4.8 in the Appendix also shows the percentage of households residing in their state of birth in 2000.

4.1.4 Characteristics of Migrant and Non-Migrant Households

Because the migration equation is estimated using data on movers,[33] it is interesting to see how their characteristics differ from those of "stayers." Table 4.3 compares the characteristics of movers and stayers. Households who moved are, on average, smaller and have fewer children than households who did not move. A higher proportion of households who moved are male-headed, and the head of household is better educated than is the case for households that did not move. The average wage per person in the household is also higher for movers.

[33] The main results presented in Chapter 5 pertain to movers; however, I also estimate the model for movers and stayers. This is discussed in the next chapter.

Table 4.3[34]: Descriptive Statistics of Migrants vs. Non Migrants

	Variable	Movers (N=441393)	Stayers (N=1083986)
Gender of head of household (proportions)	Male	64.13	60.39
Race of head of household (proportions)	White Black Other	75.9 11.03 13.07	73.04 14.95 12.01
Marital Status of head of household (proportions)	Married	46.36	47.12
Education of head of household (proportions)	No high school High school Some college College graduate Postgraduate education	10.71 17.83 34.19 23.35 13.92	19.25 25.49 30.41 16.39 8.46
Age of head of household (Mean)	Age	38.44	42.88
Household Wage Earnings (Mean)	Sum of the wage earnings of all household members	44870.83	43863.29
Total Household Income (Mean)	Sum of wage ,business and farm incomes and income from other sources [35] of all household members	63578.73	56857.41

[34] There are 5,663,214 households in the PUMS data. We know the MSAs that households lived in 1995 and 2000 for 26.9% of these households (1,525,379 households). For the remaining households, we do not have values for the MSA variable. This may be because these were households who did not live in MSAs in either of the two years, migrated to the US from abroad or we have missing values for either of the two years.

[35] Income from other sources would include Social Security income, welfare (public assistance) income, Supplementary Security income, interest, dividend, and rental income, retirement income and other income.

Size of household	1 member	39.85	30.06
	2 members	27.37	26.59
	3 members	13.22	16.87
	4 members	11.41	14.68
	More than 4 members	8.15	11.8
Number of children in the household	0 children	68.4	54.67
	1 child	13.31	18.98
	2 children	11.68	16.41
	3 children	4.68	6.85
	4 children	1.39	2.14
	>4 children	0.54	0.95

4.2 Amenities

4.2.1 Climate

Previous studies have used a variety of climate variables, including mean January temperature, mean July temperature, average January precipitation and average July precipitation, heating and cooling degree days, wind speed and percent possible sunshine.

In this study, I estimate models using the following different climate variables. I use seasonal variables for temperature and precipitation. Winter variables are calculated using the months of December, January and February; spring variables using March, April and May; summer variables using the months of June, July and August; fall using

September, October and November.[36] However, the high correlation between the four seasons (shown in Table 4.4) prevents me from using all four together. Instead, I use winter and summer variables.

Keeping in mind that past studies have used annual or January and July values of degree days, I have used these alternative measures too. Heating and cooling degree days are computed by the National Climatic Data Center as follows. First, the average of the high and low temperature for the day is computed. If this is greater than 65 F, it results in (Average temperature - 65) cooling degree days. If the average temperature is less than 65 degrees it results in (65 - Average temperature) heating degree days. [37] Degree days are likely to be highly correlated with average temperature. For example, January HDD = 2015 – (31*Average January Temperature) provided average temperature is less than 65 degrees for all days in January. Thus, in theory, there should not be any significant difference in using degree days instead of temperature. Average cooling degree days in July should likewise be correlated with average July temperature. The correlation between degree days and temperature are shown in Table 4.4. Figures 4.10 and 4.11 show temperature and precipitation climate zones of the US and average annual precipitation (though the period of record of the data for the maps is 1961-1990).

[36] Correlations between the temperature variables for the 12 months are shown in the Section A4.2.1 of the Appendix.
[37] For detailed methodology, please see United States Climate Normals, 1971-2000; Degree Day Computation Methodology; National Climatic Data Center/NESDIS/NOAA ; January 15, 2003 . (http://lwf.ncdc.noaa.gov/oa/climate/normals/normdegdmeth.pdf)

As explained in Chapter 2, the temperature, precipitation and degree day variables are climate normals, i.e., the arithmetic mean of a climatological element computed over three consecutive decades. I use data for the period 1971-2000.[38]

Following Deschenes and Greenstone (2007), I also experimented with "bin data."[39] That is, I used the number of days a county faces temperatures in 5 degree Fahrenheit intervals or "bins." Using the bin data or annual degree days, however, leads to loss of seasonality. In the results reported below, I use this data to create number of days below 35 degrees and above 75 degrees to capture extreme temperatures. In some of the runs, these are interacted with precipitation to proxy days with snow and humidity.

Other aspects of climate which are potentially relevant to households are average wind speed and the amount of possible sunshine. The latter is defined as the total time that sunshine reaches the surface of the earth, expressed as the as the percentage of the maximum amount possible from sunrise to sunset with clear sky conditions. Unfortunately, data on wind speed and sunshine are available for fewer than half of the MSAs in the dataset. Humidity is another aspect of climate that households care about but humidity data are unavailable for 136[40] of the 297 MSAs. Thus, including either of these variables would result in a loss of too many observations.

[38] These numbers are weighted by county population shares to get average values for each MSA. Please see Section A4.2.1 of the Appendix for details.

[39] This was generously provided by Olivier Deschenes and Michael Greenstone.

[40] These are not randomly missing across the country. Thus certain Census divisions will be over represented in the data and some will be under represented. Another problem with this data is that these are humidity levels at weather stations averaged over several years. Data from different number of years are used for different weather stations and thus the data are not comparable across MSAs.

Figure 4.10 Climate zones of the US: 1961 to 1990

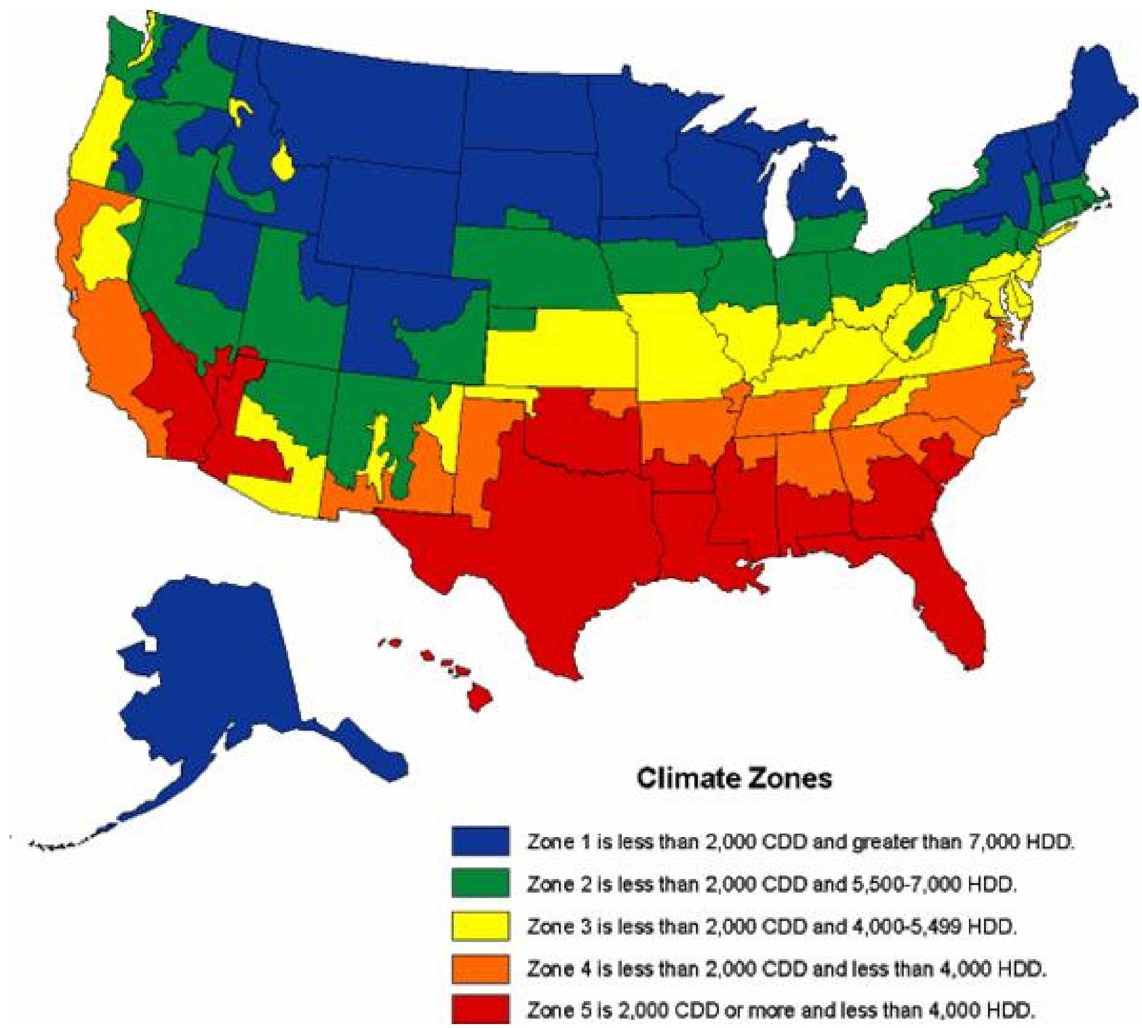


Figure 4.11 Average Annual Precipitation: 1961 to 1990

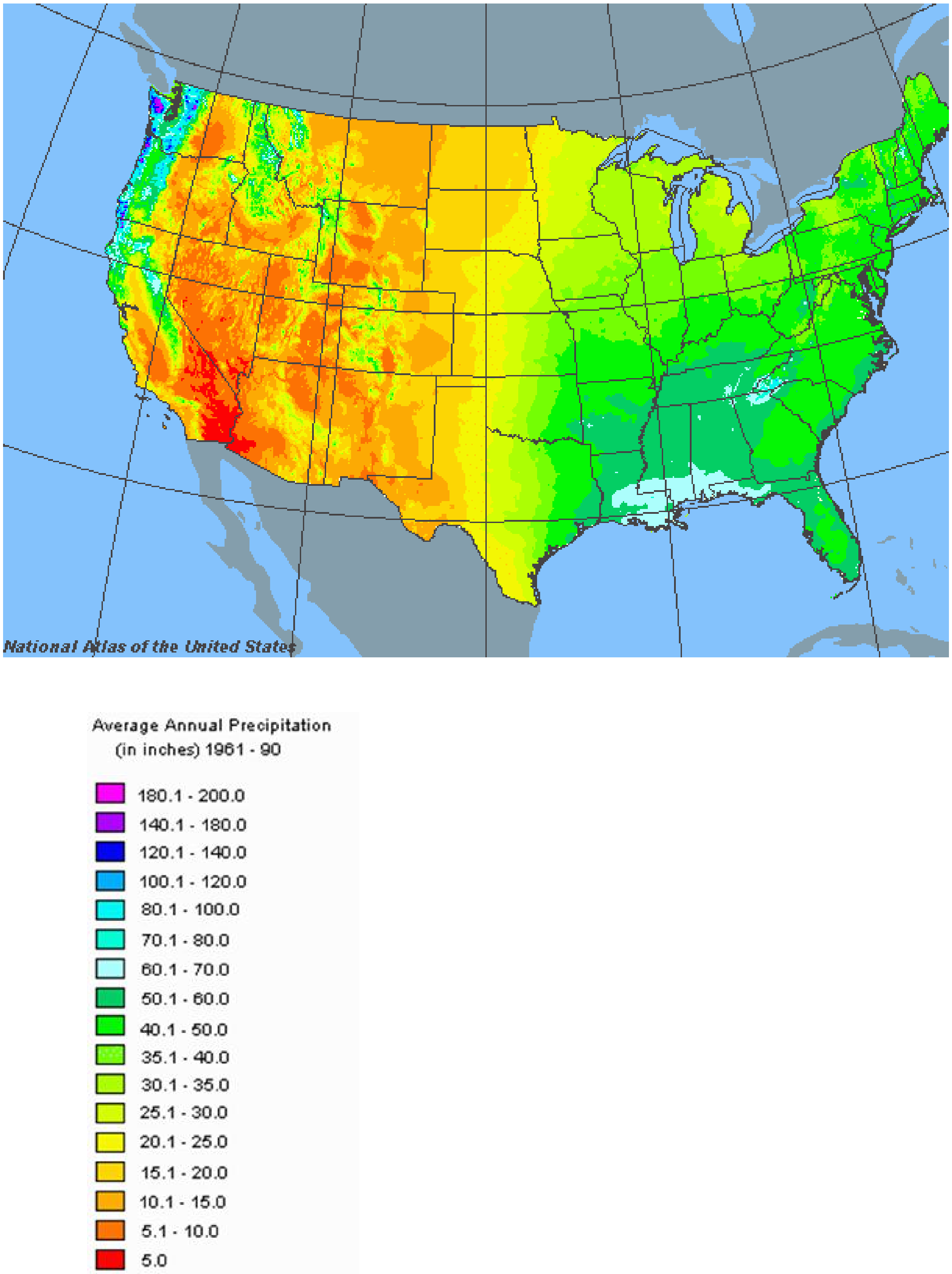


Table 4.4 Summary Statistics and Correlation of Seasonal Climate Variables

VARIABLE	WINTER_TEMP	SUMMER_TEMP	SPRING_TEMP	FALL_TEMP	WINTER_PRCP	SUMMER_PRCP	SPRING_PRCP	FALL_PRCP	WINTER_HDD	SUMMER_HDD	SPRING_HDD	FALL_HDD	WINTER_CDD	SUMMER_CDD	SPRING_CDD	FALL_CDD
MEAN	**37.2**	**73.3**	**54.6**	**57.3**	**3.1**	**3.7**	**3.5**	**3.3**	**844.5**	**23.3**	**372.9**	**312.6**	**8.3**	**280.0**	**54.4**	**77.7**
STD	**12.1**	**5.7**	**8.6**	**8.1**	**1.7**	**1.7**	**1.2**	**1.1**	**346.8**	**30.5**	**205.0**	**168.0**	**23.9**	**151.6**	**67.9**	**86.9**
N	**295**	**295**	**295**	**295**	**295**	**295**	**295**	**295**	**295**	**295**	**295**	**295**	**295**	**295**	**295**	**295**
WINTER_TEMP	1.00	0.76	0.94	0.96	0.33	0.19	0.09	0.19	-1.00	-0.50	-0.94	-0.96	0.65	0.77	0.82	0.87
SUMMER_TEMP		1.00	0.91	0.89	-0.02	0.41	0.18	0.17	-0.75	-0.82	-0.90	-0.87	0.50	0.99	0.83	0.85
SPRING_TEMP			1.00	0.99	0.19	0.34	0.15	0.22	-0.94	-0.66	-0.99	-0.97	0.64	0.93	0.90	0.93
FALL_TEMP				1.00	0.20	0.34	0.14	0.22	-0.95	-0.67	-0.98	-0.98	0.67	0.90	0.90	0.94
WINTER_PRCP					1.00	0.03	0.67	0.61	-0.35	0.10	-0.25	-0.28	-0.05	0.00	-0.01	0.03
SUMMER_PRCP						1.00	0.56	0.65	-0.16	-0.40	-0.28	-0.28	0.53	0.40	0.45	0.44
SPRING_PRCP							1.00	0.85	-0.09	-0.21	-0.18	-0.18	0.00	0.16	0.05	0.05
FALL_PRCP								1.00	-0.18	-0.07	-0.20	-0.21	0.28	0.18	0.25	0.23
WINTER_HDD									1.00	0.50	0.95	0.96	-0.61	-0.77	-0.80	-0.85
SUMMER_HDD										1.00	0.68	0.70	-0.27	-0.75	-0.50	-0.53
SPRING_HDD											1.00	0.98	-0.55	-0.91	-0.83	-0.87
FALL_HDD												1.00	-0.56	-0.87	-0.81	-0.86
WINTER_CDD													1.00	0.52	0.83	0.83
SUMMER_CDD														1.00	0.86	0.88
SPRING_CDD															1.00	0.99
FALL_CDD																1.00

Table 4.5 Descriptive Statistics of Amenities Used in the Second Stage

Variable	N	Mean	Std Dev	Minimum	Maximum	Median
WINTER_TEMP	295	37.177	12.066	9.442	67.922	34.805
SUMMER_TEMP	295	73.346	5.729	60.848	89.733	72.547
SUMMER_CDD	295	279.999	151.559	32.005	760.000	245.987
WINTER_HDD	295	844.536	346.832	68.333	1670.550	908.591
WINTERPR	295	9.402	4.971	1.500	28.084	9.206
SUMMERPR	295	11.029	4.981	0.440	23.300	11.954
ANNUAL_CDD	295	1261.240	939.873	111.783	4171.000	931.000
ANNUAL_HDD	295	4660.010	2188.000	240.667	9863.630	5017.000
ANNUAL_PRCP	295	40.723	13.592	5.080	66.747	43.218
DAYS WITH AVG TEMPERATURE <35	293	48.212	40.715	0.000	146.629	45.329
DAYS WITH AVG TEMPERATURE > 75	293	67.708	53.774	0.600	237.273	50.314
TRANSPORTATION	295	50.354	29.199	0	100	50.420
EDUCATION	295	51.015	29.182	0	100	50.990
ARTS	295	51.021	28.825	0	100	51.000
HEALTHCARE	295	48.418	28.696	0	98.3	48.440
RECREATION	295	52.586	28.658	0	100	53.540
TOTAL CRIME RATE	290	0.043	0.015	0.0019744	0.0890493	0.0417917
MSA OUT OF ATTAINMENT WITH NAAQS	297	0.199	0.400	0	1	0
MEAN PM 2.5	295	12.914	2.879	5.382	19.535	12.947
95TH PERCENTILE OF PM 2.5	295	27.063	6.547	9.389	58.177	27.261
PM 10 (MEAN)	295	23.503	4.647	10.930	44.384	23.315
PM 10(95TH PERCENTILE)	295	44.824	10.038	19.124	96.148	43.680
POPULATION DENSITY PER SQUARE MILE	297	471.266	970.289	5.400	13043.600	255.100
POPULATION	297	747077.67	1191629.06	101541	9519338	341851
MSA ON COAST	297	0.313	0.465	0	1	0
MSA ON GREATLAKES	297	0.064	0.245	0	1	0
MSA ON PACIFIC	297	0.067	0.251	0	1	0
MSA ON ATLANTICGULF	297	0.182	0.386	0	1	0
NORTHEAST	297	0.178	0.384	0	1	0
MIDWEST	297	0.246	0.431	0	1	0
WEST	297	0.199	0.400	0	1	0
SOUTH	297	0.377	0.485	0	1	0
NEW ENGLAND	297	0.064	0.245	0	1	0
MIDDLE ATLANTIC	297	0.114	0.319	0	1	0
EAST NORTH CENTRAL	297	0.175	0.381	0	1	0
WEST NORTH CENTRAL	297	0.071	0.257	0	1	0
SOUTH ATLANTIC	297	0.185	0.389	0	1	0
EAST SOUTH CENTRAL	297	0.067	0.251	0	1	0
WEST SOUTH CENTRAL	297	0.125	0.331	0	1	0
MOUNTAIN	297	0.067	0.251	0	1	0
PACIFICDIV	297	0.131	0.338	0	1	0

4.2.2 Crime

Bearing in mind that people may react differently to violent as opposed to property crime, I attempted to use two different crime variables in the analysis. Property crimes include burglaries, larcenies, motor vehicle thefts, and arsons. Violent crimes include murders, rapes, robberies, and aggravated assaults. These are expressed as rates by dividing by population. The source of these data is the Uniform Crime Reporting (UCR) Program of the Federal Bureau of Investigation. High correlation between the two measures, however, led me to combine them in the models below. As these data are available by county, crime rates in a county are weighted by the shares of county population, yielding a population-weighted average for each MSA. Table 4.5 provides summary statistics of this variable.

4.2.3 Air Quality

Average annual PM10 data from the US Environmental Protection Agency is used to measure air pollution. PM10 is visible to the human eye and has deleterious health effects. The BenMAP tool (Abt 2005)[41] from the US EPA was used to covert monitor level data to air quality grids for each MSA. From these grids, it is possible to compute population-weighted annual average PM10. Unfortunately, as Table 4.5 indicates, there is little variation in average annual PM10 across MSAs. I thus use the 95th percentile of annual values as well.

[41] Available online at http://www.epa.gov/air/benmap/

PM2.5 is also used as an alternate pollution variable as it is believed to have more severe health effects than PM10. I also experiment with both the mean and the 95th percentile of annual values for this variable. These pollution measures are also weighted using county population shares.

An alternative measure of air quality is also used in second stage regressions. This is a dummy variable that indicates whether an MSA is in violation of the Environmental Protection Agency's National Ambient Air Quality Standards (NAAQS).[42] An MSA is defined to be "out of status or attainment" if at least one of its component counties is in violation of the NAAQS with respect to any of the criteria pollutants in 2000. This is constructed using data (available online) from the EPA which has information on the non-attainment status for each county by year.[43]

4.2.4 Data from the Places Rated Almanac

A key difficulty in valuing climate amenities is to separate their effects from endogenous amenities that are likely to be correlated with climate: Recreational opportunities, for example, are likely to be more numerous in cities with milder climates. To capture other amenities that may be correlated with climate I use the *Places Rated Almanac*. This publication contains indices of the quality of education, transportation, recreation and health services and the arts for all MSAs in 2000. For transportation, the factors used to rate each MSA are its supply of public transit, average commute time, and connectivity with other metro areas via national highways, scheduled air service, and

[42] These standards are available at http://www.epa.gov/air/criteria.html .
[43] Source: http://www.epa.gov/air/oaqps/greenbk/anay.html

passenger rail service, as well as proximity to all other metropolitan areas. The education index reflects School Support (measured by the average pupil-teacher ratio and percent of funding received from local sources), Library Popularity (the circulation rate added to number of volumes divided by population), College Enrollment and College Options (the variety of higher education institutions in the MSA). A more detailed description of these data may be found in Section A4.2.3 of the Appendix.

4.2.5 Population Density and Population

Population density is included to capture amenities not specifically captured by the *Places Rated Almanac*. Population in an MSA also captures the effects of city size. Households may be attracted to "big" cities. This would proxy for unmeasured amenities in such cities. Furthermore, in a multinomial logit model without covariates, MSA-specific constants would reflect the proportion of choosers in the sample going to each city. The source of data for these two variables is the Census Bureau. Figure 4.10 shows the population density by state in 2000 while Figure 4.11 shows the population in millions.

Figure 4.10 Population Density by State, 2000 ***(Source: US Census Bureau)***

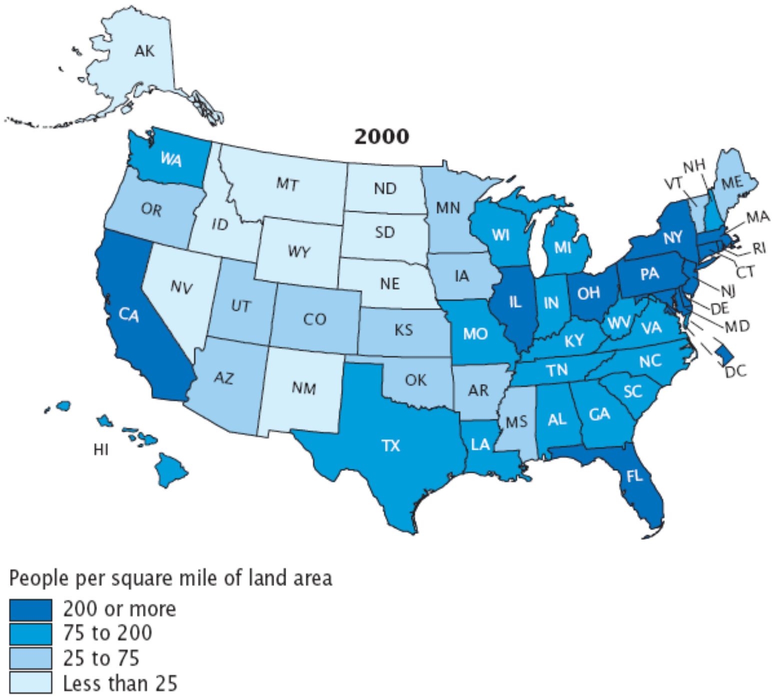
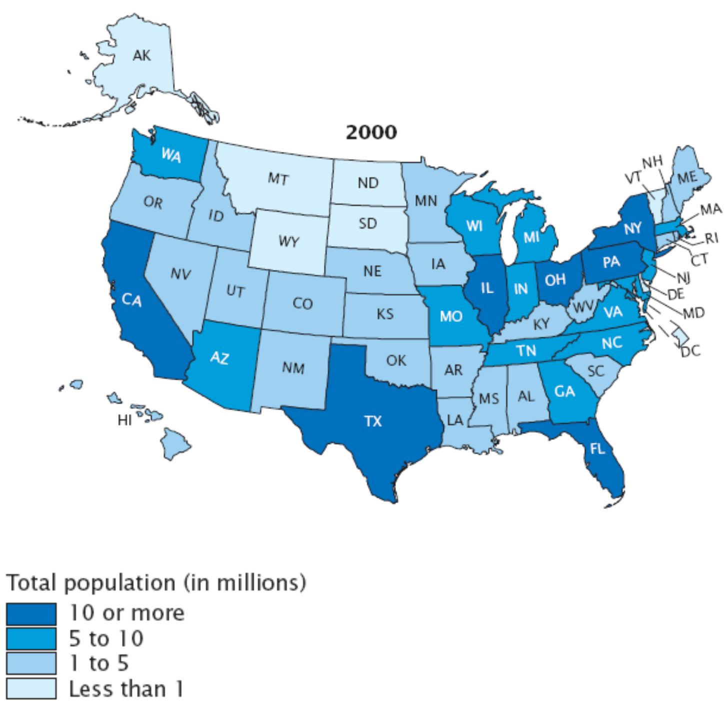


Figure 4.11 Population by State, 2000 ***(Source: US Census Bureau)***

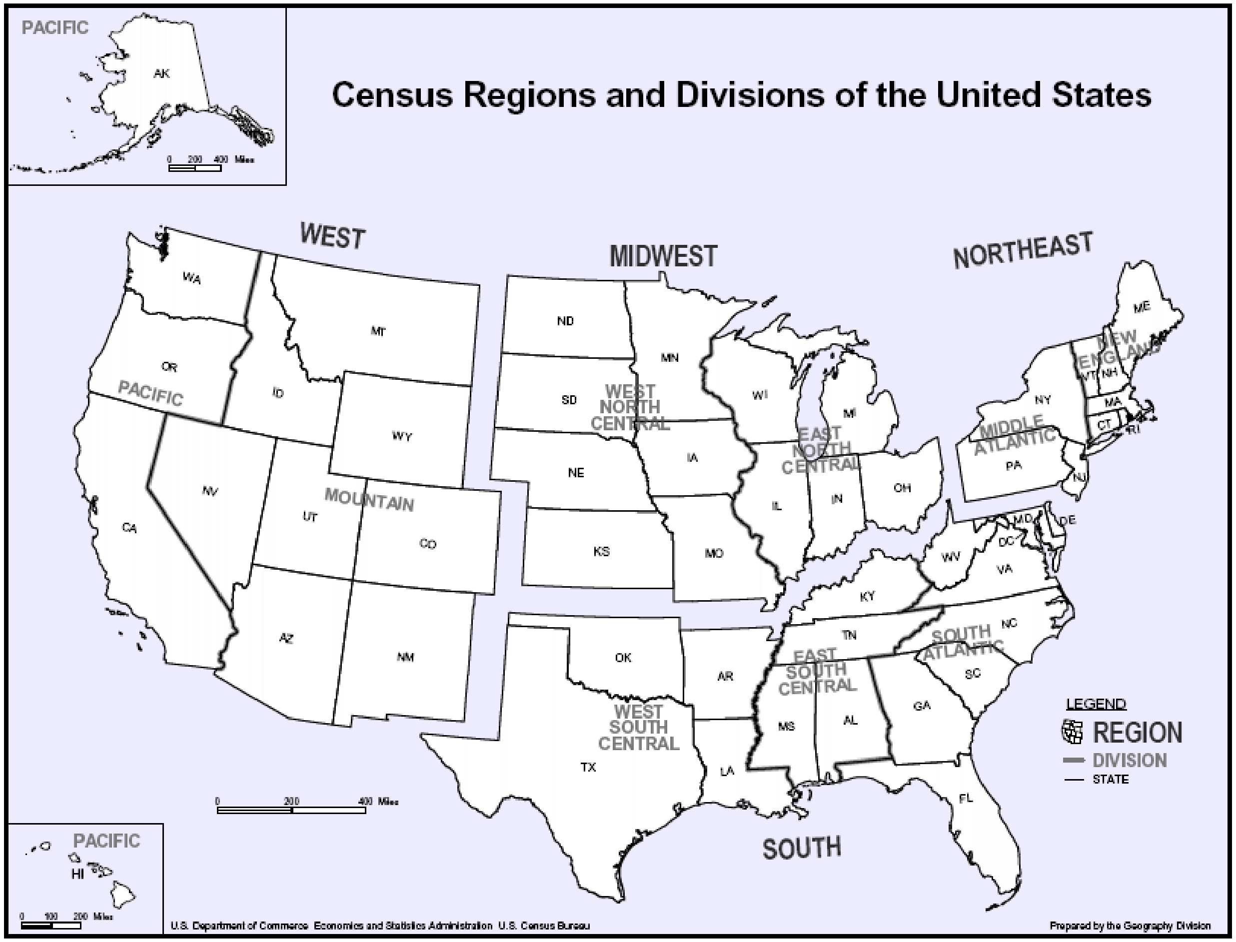


4.2.6 Proximity to the Coast, Regional and Divisional Dummies

The coastal dummy indicates that the MSA is located on the Pacific Ocean, the eastern coast (including the Atlantic Ocean and the Gulf of Mexico) or the Great Lakes. I also use dummies for these three coasts separately. This is to capture different preferences for the three coasts. For example, households may value the Great Lakes differently from the California coast.

Regional or divisional dummies are included to reduce regional variation in amenities not explicitly controlled for as well as differences in the cost of non-housing goods. A map of the US showing the different Census Regions and Divisions is shown in Figure 4.12. To show the variation in the amenities within each region (or division), summary statistics for the amenities are shown by region and division in Section A4.2.4 and A4.2.5of the Appendix.

Figure 4.12 ***(Source*[44]*: US Census Bureau)***



[44] Available online at http://www.census.gov/geo/www/

Chapter 5 Estimation Results

This chapter begins in section 5.1 by presenting results from the hedonic wage and housing equations that are used to predict wages and housing costs in all MSAs in the universal choice set. The results from the migration equation are presented in Section 5.2 Second stage estimates are presented in Section 5.3.

5.1 Hedonic Price Functions

Hedonic models are estimated for the labor and housing markets to predict wages households would earn and living costs in all locations. As noted above, separate hedonic wage equations are estimated for each MSA. A single housing market equation is estimated to obtain housing price indexes for each MSA. The results for the wage and housing hedonic functions are summarized in Section 5.1.1 and Section 5.1.2 respectively.

5.1.1 Hedonic Wage Functions

The results of the hedonic wage functions are presented in the Table A5.1 of Appendix. Since the wage regressions are estimated separately for each MSA, the mean and standard deviation of the 297 coefficients for each explanatory variable are presented in the table. Most variables are significant at the 5% level for all MSAs. Older workers earn more, but the premium on age declines with age, as expected. Married individuals and males earn more. Good English-speakers earn more than people who have difficulty with the language and Hispanics earn less than non-Hispanics. There are positive returns to education. Occupation dummies also have the expected signs, i.e., occupations requiring more education and/or white collar occupations earn more. What is notable, however, is that the returns to different occupations and industries vary significantly across MSAs, suggesting that the assumption of a national labor market, made in earlier hedonic studies, is inappropriate.

As a sensitivity test, I use two different definitions of full year workers: those who work more than 30 weeks and those who work more than 40 weeks. The results obtained are extremely similar. The means of the 297 coefficients for each explanatory variable have a correlation coefficient of 0.9996.The results from the sample using those who work more than 40 weeks is used in the final estimation.

5.1.2 Hedonic Housing Functions

The results of the hedonic housing equation are presented in Table A5.2 of the Appendix. An owner-occupied house carries a premium. Houses with greater numbers of rooms and bedrooms are worth more. Older houses have lower value than newer houses. These variables are all statistically significant at the 5% level. Ninety three percent of the MSA specific dummy variables are statistically significant at the 5% level. The MSA specific dummies, which reflect cost of living indices in an MSA after controlling for housing specific characteristics, seem reasonable. For example, Boston has a higher index than Seattle, which is in turn more expensive than Washington DC. The MSAs in

California, New York and New Jersey have very high costs of living. The 20 most expensive and 20 least expensive MSAs are listed in Table 5.1.

To take into account the fact that the marginal value of dwelling characteristics (such as number of bedrooms and number of rooms) might differ between owners and renters, I interact the ownership dummy with these characteristics. Most of these interaction terms are significant. However, the cost of living indices are very similar to what I obtained without the interaction terms. The correlation between the living costs indices obtained from these two regressions is 0.9955.

Table 5.1 Most Expensive and Least Expensive Cities

20 Most Expensive MSAs		20 Least Expensive MSAs	
Ranking	**Name of MSA**	**Ranking**	**Name of MSA**
1	San Francisco-Oakland-Vallejo, CA	297	McAllen-Edinburg-Pharr-Mission, TX
2	San Jose, CA	296	Johnstown, PA
3	Stamford, CT	295	Gadsden, AL
4	Santa Cruz, CA	294	Anniston, AL
5	Nassau Co, NY	293	Brownsville-Harlingen-San Benito, TX
6	Oakland, CA	292	Dothan, AL
7	Santa Barbara-Santa Maria-Lompoc, CA	291	Joplin, MO
8	Bergen-Passaic, NJ	290	Alexandria, LA
9	Salinas-Sea Side-Monterey, CA	289	Sumter, SC
10	Orange County, CA	288	Danville, VA
11	Santa Rosa-Petaluma, CA	287	Florence, AL
12	Danbury, CT	286	Hattiesburg, MS
13	Honolulu, HI	285	Laredo, TX
14	Ventura-Oxnard-Simi Valley, CA	284	Fort Smith, AR/OK
15	New York-Northeastern NJ	283	Terre Haute, IN
16	Boston, MA	282	Monroe, LA
17	Los Angeles-Long Beach, CA	281	Beaumont-Port Arthur-Orange,TX
18	Newark, NJ	280	Shreveport, LA
19	Middlesex-Somerset-Hunterdon, NJ	279	Decatur, AL
20	San Diego, CA	278	Houma-Thibodoux, LA

5.2 Results from the First Stage Estimation (Migration Equation)

This section summarizes the results from the first stage estimation. A household is considered to have moved if the head of the household has moved. Households with the head serving in the military were deleted from the sample used to estimate the migration equation. This is because their location choices were not likely to have been voluntary. Also, those working in farming, fishing and forestry as well as those who were self-employed were deleted as it was difficult to predict their wages in each MSA. If households reported some members to have been in the labor force but reported zero household wages, they were deleted from the sample due to likely reporting errors.

To make the analysis computationally tractable, I chose a 20% random sample of households when estimating the migration equation, yielding 75,293 households. Following McFadden (1978) the choice set for each household included the MSA the household chose and 19 other randomly selected MSAs. This random sampling of alternatives has been shown to produce consistent estimates when the uniform conditioning property holds (McFadden 1978).

Table 5.2 presents three sets of results for the migration equation. In the first specification presented below (called Specification 0), the only independent variables are the log of household wages and the location specific dummies. Keeping in mind that households incur moving costs when moving away from their original location, the moving cost dummies are also included in the second specification (specification 1). The coefficient on the log of wage is 0.830 in specification 0 and 0.972 in specification 1. In specification 1, the moving cost dummies are statistically significant at conventional levels and have negative coefficients: changing states reduces utility as does changing Census divisions and Census regions. I use the results from specification 1 in the second stage. Henceforth, I refer to this case as the "base case."

To check whether the results are sensitive to the number of MSAs included in the choice set, I also estimate the model using the same set of households but defining the choice set as the MSA actually chosen and 39 other MSAs (Specification 2). This is presented in Table 5.2. The estimated parameters for the log of wages and the moving costs are extremely similar. The correlation between the estimated location specific parameters of specifications 1 and 2 is 0.9993. Thus, the results do not seem to be sensitive to the number of MSA included in the choice set.

Comparing Movers and Stayers

As Table 4.3 illustrates, the characteristics of movers differ on various dimensions from stayers. This raises two questions: Do stayers have the same preferences as movers? Can the preferences of stayers be estimated based on their location choices? For the latter to be possible, it must be the case that stayers are in equilibrium. While it is difficult formally to test the hypothesis that stayers are in equilibrium, some information can be provided by estimating the migration equation using both movers and stayers and comparing the results with estimates using movers only.

The results of the estimated model for the movers and stayers are presented in table 5.3 under the column titled specification 3. For the pooled sample of movers and stayers, I use a 5% random sample and thus use 66,864 observations. Roughly 28% of the sample are movers. The coefficient on the wage is 0.76 as compared to 0.97 for the movers-only sample. The coefficients for the moving costs are also very different compared to the movers-only sample. In fact, they imply that a household would give up five dollars in wages in a new location for every dollar currently earned to avoid leaving the state in which they were located in 1995. This is a much larger rate of substitution than in specification 1 and suggests that stayers may not be in equilibrium.

Following Bayer, Keohane and Timmins (2006), I also define moving costs relative to the state of birth. Thus, in models 4 and 5 of Table 5.3 each moving cost dummy equals 1 if MSA j entails the chooser leaving left the state (Census division, or region) of his birth. In the movers-only sample (Specification 4) the coefficient on the log of the wage variable is 0.838. The coefficients on the moving costs are similar to those obtained from Specification 1. The correlation between the estimated location specification intercepts of the base case is 0.979. Thus, the two models using a sample of movers yield very similar results and suggest that the estimates are robust to specification.

The model estimated with a sample of movers and stayers and using moving cost dummies calculated from birthplace (Specification 5) yields a coefficient of 0.239 on the log of wages. Given that 28% of the sample consists of movers, this suggests that the coefficient on wages for stayers is approximately zero[45]. A comparison of specifications 3 and 5 reveals that the results for the movers and stayers sample are very sensitive to

[45] The coefficient on the log of wage for the sample using both movers and stayers can be interpreted as a weighted mean of the movers and stayers. Thus,
(0.28 * 0.97) + (0.72 * Coefficient from stayers) = 0.239. Thus the coefficient from the stayers is very close to zero.

specification. Though there is no clear test as to whether the stayer households are in equilibrium, it is reasonable to conclude that the first stage results for movers and stayers seem unstable and cannot be used to estimate preferences for amenities. I will therefore estimate the preferences of movers for locational amenities using specification 1.

MSA Dummies

The MSA dummies estimated in stage one (the $\{A_j\}$) can be interpreted as Quality of Life Indices: They represent the average utility obtained from location-specific amenities net of housing costs. Tables 5.4 and 5.5 list the top 20 and bottom 20 MSAs respectively from specifications 1 and 2. The top 20 and bottom 20 cities from the base model and the model with 40 MSAs are almost the same.

Table 5.2 First Stage Estimates

Variable	Specification 0		Specification 1 (Base Case)		Specification 2	
	Movers with 20 MSAs in Choice Set and Without Moving Costs		Movers with 20 MSAs in Choice Set and Moving Costs		Movers with 40 MSAs in Choice Set and Moving Costs	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Log(household wages) ($\alpha_C + \alpha_H$)	0.8299	17.34	0.9720	18.46	0.9952	19.94
State dummy (α_{Mo})			-1.9865	-134.08	-1.9385	-142.60
Division dummy (α_{M1})			-0.5239	-30.25	-0.5185	-31.60
Regional dummy (α_{M2})			-0.6895	-48.20	-0.6865	-50.09
Number of Observations	75293		75293		75293	
Log Likelihood	-183910		-143768		-190807	
Number of Iterations	56		100		100	

Table 5.3 First Stage Estimates (Comparison with Other Models)

Variable	Specification 1 (Base Case)		Specification 3		Specification 4		Specification 5	
	Movers with Moving Costs Calculated from Location in 1995		Movers and Stayers with Moving Costs Calculated from Location in 1995		Movers with Moving Costs Calculated from Birthplace		Movers and stayers with Moving Costs Calculated from Birthplace	
	Coefficient	**t-statistic**	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
Log(household wages) ($\alpha_C + \alpha_H$)	0.9720	18.46	0.7629	9.15	0.8384	16.38	0.2387	3.84
State dummy (α_{Mo})	-1.9865	-134.08	-4.0038	-154.7	-2.1125	-131.98	-2.9306	-154.60
Division dummy (α_{M1})	-0.5239	-30.25	-0.6155	-18.51	-0.6045	-32.07	-0.7745	-32.65
Regional dummy (α_{M2})	-0.6895	-48.20	-0.6385	-22.85	-0.5264	-33.43	-0.4296	-21.32
Number of Observations	75293		66864		75293		66864	
Log Likelihood	-143768		-57023		-150960		-104018	
Number of Iterations	100		92		103		123	

Table 5.4 Top 20 Cities According to the Quality of Life Indices[46]

Rank	Specification 1 (Base Case) (Movers w/ MC)	Specification 2 (Movers w/ MC) Choice Set =40 MSAs
1	Phoenix, AZ	Phoenix, AZ
2	Atlanta, GA	Atlanta, GA
3	Washington, DC/MD/VA	Washington, DC/MD/VA
4	Las Vegas, NV	Las Vegas, NV
5	Chicago-Gary-Lake, IL	Chicago-Gary-Lake, IL
6	Boston, MA	Boston, MA
7	Tampa-St. Petersburg-Clearwater, FL	Tampa-St. Petersburg-Clearwater, FL
8	Los Angeles-Long Beach, CA	Los Angeles-Long Beach, CA
9	New York-Northeastern NJ	Denver-Boulder-Longmont, CO
10	Denver-Boulder-Longmont, CO	Dallas-Fort Worth, TX
11	Dallas-Fort Worth, TX	New York-Northeastern NJ
12	Philadelphia, PA/NJ	Philadelphia, PA/NJ
13	Orlando, FL	Orlando, FL
14	Raleigh-Durham, NC	Seattle-Everett, WA
15	Seattle-Everett, WA	Raleigh-Durham, NC
16	Charlotte-Gastonia-Rock Hill, SC	Houston-Brazoria, TX
17	Houston-Brazoria, TX	Charlotte-Gastonia-Rock Hill, SC
18	Portland-Vancouver, OR	Portland-Vancouver, OR
19	Baltimore, MD	Baltimore, MD
20	West Palm Beach-Boca Raton-Delray Beach, FL	Riverside-San Bernadino, CA

[46] These indices are relative to Abilene, TX

Table 5.5 Bottom 20 Cities According to the Quality of Life Indices[47]

Rank	Specification 1 (Base Case) (Movers w/ MC)	Specification 2 (Movers w/ MC) Choice Set =40 MSAs
297	Houma-Thibodoux, LA	Houma-Thibodoux, LA
296	Laredo, TX	Laredo, TX
295	Kokomo, IN	Sioux Falls, SD
294	Altoona, PA	Kokomo, IN
293	Sioux Falls, SD	Altoona, PA
292	Mansfield, OH	Mansfield, OH
291	Wausau, WI	Sioux City, IA/NE
290	Gadsden, AL	Wausau, WI
289	Sioux City, IA/NE	Alexandria, LA
288	Alexandria, LA	Gadsden, AL
287	Flint, MI	Flint, MI
286	Wichita Falls, TX	Billings, MT
285	Danville, VA	Springfield, IL
284	St. Joseph, MO	St. Joseph, MO
283	Springfield, IL	Danville, VA
282	Billings, MT	Williamsport, PA
281	Williamsport, PA	Wichita Falls, TX
280	Jamestown-Dunkirk, NY	Sumter, SC
279	Decatur, IL	Jamestown-Dunkirk, NY
278	Sheboygan, WI	Yuba City, CA

[47] These indices are relative to Abilene, TX

5.3 Results from the Second Stage Estimation

In this stage, I regress estimated values of the MSA-specific fixed effects on housing costs and amenities.

$$A_j = -\alpha_H \ln(R_j) + g(E_j) + \eta_j$$

The left hand side of this equation represents the average indirect utility from MSA j after controlling for household income and moving costs. Because living costs are likely to be correlated with the error term η_j, α_H, the fraction of income spent on housing, is set equal to 0.25 (which is the median share of income spent on housing in my sample) and $\alpha_H R_j$ is added to the dependent variable.[48]

$$(A_j + 0.25 R_j) = g(E_j) + \eta_j$$

Specifications of climate variables other than temperature are described in Section 5.3.2. To check the sensitivity of the results to the choice of α_H, the value is alternately set to 0.2 and 0.3 and the dependent variable is calculated using these numbers. These results are presented in Section 5.3.3. Sensitivity of temperature and precipitation results to other equation specifications are also presented in this section. To see how much the results are driven by the effects of amenities on living costs, the living cost indices are

[48] Since I am using estimated numbers as the dependent variable in the second stage" estimation, the errors should be adjusted to reflect the correct standard errors. Generalized Least Squares could be used to estimate the second stage using the covariance matrix of the alternative specific constants estimated in the first stage as the transformation matrix. However, as the dependent variable is the sum of the MSA - specific constant and the cost of living index, it would be difficult to do this. To argue that the second stage estimation is consistent using sample size calculations requires that the sample size be > (# of MSAs squared) (Berry, Linton and Pakes (2004)). So a sample of over 90,000 households is required to argue this. So, I have also run the base case of the model using a 25%, which yields 93,737 households. The results are almost identical – the correlation between the estimated MSA coefficients is 0.998.

themselves regressed on amenities. These results are presented in Section 5.3.4. Estimates of willingness to pay for climate amenities are presented in Section 5.3.5.

5.3.1 Second Stage Results for Movers

Table 5.6 present two sets of specifications for the second stage, based on the model with movers only and moving costs calculated based on the MSA lived in 1995 (Specification 1 described in section 5.3) as dependent variables. The two different specifications include variables described in chapter 4 (Table 4.5). The first set of results (labeled as model 1 in the table) includes the Census region dummies while the second set of results (labeled as model 2 in the table) includes the Census division dummies.[49] Figure 4.12 shows the different Census regions and divisions. Results with division dummies generally show smaller impacts of climate on migration decisions and are more conservative than estimates with region dummies. This is because temperature and precipitation may pick up differences in non-housing costs of living when only regional dummies are included.

The second stage models fit well ($R^2 \approx 0.76$-0.77) and most variables are significant at conventional levels, with expected signs. Exceptions to this include air pollution, which has a positive sign and the health care index from the *Places Rated Almanac*, which also has the wrong sign, but is statistically insignificant. Pollution levels are likely to be correlated with local economic activity and thus instruments are needed in

---

[49] Variation of amenity values within regions and divisions is shown in Tables (A4.2.4) and (A4.2.5) of the Appendix.

order to get consistent estimates of the coefficient on particulate matter. This issue is however, not of focus of this study and is addressed by Bayer, Keohane and Timmins (2006). Note that the West regional dummy has a positive and significant coefficient which agrees with migration patterns presented in Figures 4.7 and 4.8.

The climate variables included in the second stage regressions are average winter and summer temperatures and precipitation. I also include the squares of these variables to allow for preferences consistent with an optimal value for each of the climate variables.[50] For example, it is reasonable to assume that individuals prefer higher winter temperatures, but only up to a point. Beyond an optimum point, higher temperatures reduce utility. Table 5.6 indicates that winter temperature increases utility up to 53 degrees Fahrenheit (model 1) and 59 degrees Fahrenheit (model 2). Figure 5.1 shows a plot of utility against average winter temperature. The slope of the utility function reflects the marginal effect at each temperature and this slope is zero at the optimum temperature. The corresponding numbers for summer are 87 and 74 degrees Fahrenheit. It should be kept be mind, however, that the summer variables are not significant at conventional levels in either model.

[50] Though using log of the climate variables yield a very simple expression for the WTP number, it is not possible to represent preferences consistent with the idea that households have an optimal temperature using that specification. Thus, I do not present results using such a functional form for the utility function. Using log of the other amenities but quadratic climate variables does not improve results.

Figure 5.1 Plot of Utility Against Average Winter Temperature

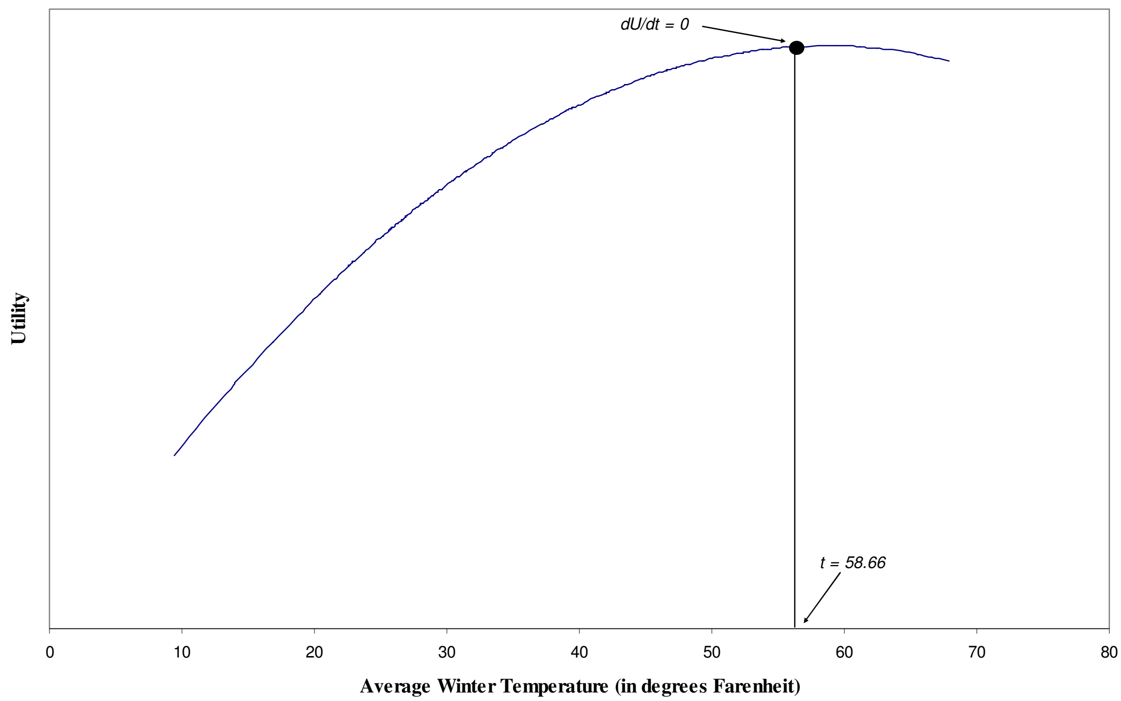


Table 5.6 Second Stage Results

Using Estimates from Specification 1 (Number of Observations =286)	With Census Regions Model 1		With Census Divisions Model 2	
	Coefficient	**t-statistic**	**Coefficient**	**t-statistic**
INTERCEPT	0.4362	0.06	-3.5680	-0.46
MEAN PM10	0.0191	2.21	0.0226	2.60
TOTAL CRIME RATE	-5.1297	-1.79	-5.6049	-1.97
POP DENSITY PER SQ MILE OF LAND	0.0001	3.97	0.0001	4.00
TRANSPORTATION	0.0032	1.75	0.0032	1.79
EDUCATION	0.0062	3.43	0.0064	3.53
ARTS	0.0104	5.58	0.0096	5.18
HEALTHCARE	-0.0005	-0.31	0.0007	0.50
RECREATION	0.0136	7.03	0.0131	6.78
MSA ON THE COAST	-0.1713	-1.97	-0.1772	-2.05
WINTER TEMP AVG	0.1054	3.54	0.0802	2.57
WINTER TEMP AVG SQUARED	-0.0010	-2.78	-0.0007	-1.82
SUMMER TEMP AVG	-0.0692	-0.34	0.0627	0.29
SUMMER TEMP AVG SQUARED	0.0004	0.29	-0.0004	-0.29
WINTERPR	-0.0503	-1.74	-0.0649	-2.06
WINTERPR SQUARED	0.0011	0.97	0.0023	1.91
SUMMERPR	0.0869	1.58	0.0736	1.19
SUMMERPR SQUARED	-0.0013	-0.59	-0.0016	-0.70
NORTH EAST[51]	0.0180	0.13		
MID WEST	-0.1274	-0.90		
WEST	0.8267	2.96		
MIDDLE ATLANTIC[52]			-0.6400	-1.58
EAST NORTH CENTRAL			-0.6330	-1.52
WEST NORTH CENTRAL			-0.9348	-1.97
SOUTH ATLANTIC			-0.4636	-1.15
EAST SOUTH CENTRAL			-0.8483	-2.07
WEST SOUTH CENTRAL			-0.7470	-1.87
MOUNTAIN			0.2459	0.92
NEW ENGLAND			-0.3280	-0.84
Adjusted R-Squared	**0.7525**		**0.7837**	
R-Squared	**0.7698**		**0.7629**	

[51] The left out category is the SOUTH.
[52] The left out category includes the PACIFIC division

Table 5.7 Marginal Effects of Climate Variables for Movers

Using Estimates from Specification 1	Model 1 With Census Regions		Model 2 With Census Divisions	
	Coefficient	**t-statistic**	**Coefficient**	**t-statistic**
WINTER TEMPERATURE	0.0304	3.14	0.0291	2.80
SUMMER TEMPERATURE	-0.0110	-0.63	0.0002	0.01
WINTER PRECIPITATION	-0.0296	-2.54	-0.0222	-1.53
SUMMER PRECIPITATION	0.0584	4.18	0.0375	1.90

The marginal effects of the climate variables, calculated at sample means, are presented, together with their t-statistics, in Table 5.7. This table indicates that winter temperature is an amenity. The results also imply that households may value higher summer precipitation and lower winter precipitation, though this is imprecisely estimated. Summer temperature does not have a significant marginal effect in either specification.[53]

5.3.2 Other Specifications of Climate Variables

Fall and spring variables were also included in one specification. However, it is difficult to distinguish their effects as they are very highly correlated with the winter and summer variables. For example, winter temperature has a correlation of 0.94 with spring temperature and 0.96 with fall temperature.[54] The number of observations also limits the number of degrees of freedom.

[53] These results are robust even using moving costs calculated from birthplace. The marginal effects are presented in Table A5.7 in the Appendix.

[54] Please see Table 4.4.

Degree days have also been used in the literature as measures of climate variables. I use both annual measures of degree days and seasonal measures of degree days (DD). As noted in chapter 4, however, degree days are highly correlated with temperature (see Table 4.4) and thus should not yield different results. These results for the second stage model and the marginal effects of the climate variables are presented in Tables A5.3 and A5.4 of the Appendix. Annual HDD have a negative marginal effect indicating that warmer temperatures are preferred. Annual cooling degrees have a positive marginal effect which indicates that warmer temperatures are preferred. This is true both in models where regions are used as well as when divisions are used. However, the marginal effects of both annual HDD and CDD are not statistically significant at the 5% level. The use of annual DD sacrifices seasonality. I therefore focus on results using mean summer and winter temperature and precipitation.

I also estimate the second stage model using number of days an MSA faces temperatures in 15 degree Fahrenheit intervals or "bins." However, these variables are not statistically significant at the 5% level for most of the "bins." Finer definitions of "bins" do not yield statistically significant results either. These results are presented in Table A5.5 of the Appendix.

Preferences for extreme weather are accounted for using days where the average temperature is below 35 and above 75 degrees Fahrenheit. This is constructed using the "bin data" described in Chapter 4. None of the marginal effects is significant at conventional levels however. These results are presented in Table A5.6 of the Appendix.

The marginal value of precipitation may depend on temperature, e.g., location in extremely wet places which are also very hot may not be desirable, due to humidity. To take this fact into account, I interact temperature and precipitation for both the summer and winter months. However, this results in the marginal effect of the winter temperature being statistical insignificant. The marginal effects are presented in Table A5.6 of the Appendix. This may be due to the inclusion of variables which are highly correlated with each other.

Though the humidity data has problems (please see chapter 4 for details), I estimate the model using this data. The marginal effects (presented in Table A5.6 of the Appendix) for winter temperature and summer precipitation are statistically significant and larger in magnitude in comparison to the models without humidity[55].

5.3.3 Sensitivity of Temperature and Precipitation Results to Equation Specification

Keeping in mind that PM 2.5 is known to have more severe health effects, I use the mean value of PM 2.5 instead of PM 10. The variable is statistically significant at conventional levels but as with PM 10, has the wrong sign. The results (presented in the Table 5.8 below) are very similar to model 2. The optimal winter temperature calculated for this specification is roughly 57 degrees Fahrenheit.

As the variation in means of the pollution variables are very low relative to the mean (Table 4.5), I alternately use the 95th percentile of the PM 10 and the PM 2.5 (presented in the Table 5.8 below) variables. These yield very similar results. The optimal

[55] Since this model is run using a much smaller set of MSAs (156), I also ran the model using just these 156 MSAs without the humidity data to facilitate comparison.

winter temperatures calculated from both these runs are 58 and 56 degrees Fahrenheit, respectively.

To account for bad air quality in terms of other criteria pollutants, in an alternative specification, I use a dummy variable that is one if an MSA is out of attainment with the NAAQS.[56] Once again, the results (presented in the Table 5.8 below) are very similar.

To explicitly capture the effects of city size, I use population levels in an MSA instead of population density. The results (presented in the Table 5.8 below) do change – the marginal effect of the winter temperature is slightly higher and the summer precipitation (significant at the 10%) level is slightly lower. The other variables are still insignificant.

As households may react differently to property as opposed to violent crime, I use these two measures. The violent crime variable has the wrong sign but is statistically insignificant at all conventional levels. The marginal effects (presented in the Table 5.8 below) of the climate variables are once again very similar to the results obtained previously.

Since households may have a different preference for the living on the Pacific coast than the Great Lakes, I split the coastal dummy into three parts – the Pacific coast dummy, the Great Lakes dummy and the Atlantic/Gulf of Mexico dummy. The only significant coastal variable is the Pacific coast dummy and this has a negative sign. Once again, the marginal effects of the climate variables (presented in the Table 5.8 below) are unaffected by this change in specification.

---

[56] Please see Chapter 5 for a more detailed description of this variable.

Table 5.8 Marginal Effects of Climate Variables Using Variants of other Amenities

Using PM2.5	With Census Divisions	
	Coefficient	t-statistic
WINTER TEMPERATURE	0.02655	2.54
SUMMER TEMPERATURE	0.01772	0.93
WINTER PRECIPITATION	-0.02469	-1.53
SUMMER PRECIPITATION	0.03529	1.77
Using MSA Out of Attainment with NAAQS	With Census Divisions	
	Coefficient	t-statistic
WINTER TEMPERATURE	0.03098	2.9
SUMMER TEMPERATURE	0.01371	0.75
WINTER PRECIPITATION	-0.01953	-1.34
SUMMER PRECIPITATION	0.03261	1.64
Using Population	With Census Divisions	
	Coefficient	t-statistic
WINTER TEMPERATURE	0.03341	3.53
SUMMER TEMPERATURE	-0.02610	-1.45
WINTER PRECIPITATION	-0.02064	-1.56
SUMMER PRECIPITATION	0.03471	1.93
Using Two Crime Variables	With Census Divisions	
	Coefficient	t-statistic
WINTER TEMPERATURE	0.02913	2.81
SUMMER TEMPERATURE	-0.00148	-0.08
WINTER PRECIPITATION	-0.02151	-1.48
SUMMER PRECIPITATION	0.03623	1.83
Using 3 Coastal Dummies	With Census Divisions	
	Coefficient	t-statistic
WINTER TEMPERATURE	0.03259	2.98
SUMMER TEMPERATURE	-0.00754	-0.37
WINTER PRECIPITATION	-0.02290	-1.58
SUMMER PRECIPITATION	0.03244	1.61

Sensitivity of Results to α_H

I have used a value of α_H =0.25 to construct the dependent variable for the second stage as this is the median value for the share of income spent on housing in my sample. To check if the results are sensitive to this choice of α_H, I construct dependent variables using values of 0.2 and 0.3 and use these to estimate Models 1 and 2. The resulting coefficients are very similar to those obtained from using a value of 0.25. The marginal effects of the climate variables are presented in Table 5.9 below. These marginal effects are also very similar to those in Table 5.8.

Table 5.9 Marginal Effects of Climate Variables for Movers Using Different Alpha Values

Using α_H =0.2	Model 1		Model 2	
	With Census Regions		With Census Divisions	
	Coefficient	t-statistic	Coefficient	t-statistic
WINTER TEMPERATURE	0.0299	3.11	0.0286	2.77
SUMMER TEMPERATURE	-0.0100	-0.57	0.0011	0.06
WINTER PRECIPITATION	-0.0294	-2.54	-0.0218	-1.51
SUMMER PRECIPITATION	0.0589	4.24	0.0377	1.93
Using α_H =0.3	Model 1		Model 2	
	With Census Regions		With Census Divisions	
	Coefficient	t-statistic	Coefficient	t-statistic
WINTER TEMPERATURE	0.0310	3.18	0.0296	2.83
SUMMER TEMPERATURE	-0.0121	-0.68	-0.0006	-0.03
WINTER PRECIPITATION	-0.0297	-2.54	-0.0225	-1.54
SUMMER PRECIPITATION	0.0579	4.12	0.0373	1.88

5.3.4 Effects of Amenities on Living Costs

The second stage estimates presented in the previous section reflect the impacts of amenities on the cost of living index as well as on the MSA-specific constants. As living costs are used to construct the dependent variable, it is interesting to see how amenities affect these costs: It is of interest to see whether amenities explain variation (i.e., are capitalized in) living costs. These results are included in Table 5.10. The pollution variable has a positive sign here though it is insignificant. Places with lower crime rates are more expensive. MSAs on the coast are more expensive. The marginal effects of the climate variables are shown in Table 5.11. Only the temperature variables are significant – places with higher winter temperatures and those with lower summer temperatures have higher housing costs. These results suggest that it is not lack of variation in summer temperature that accounts for the results in section 5.3.1.

Table 5.10 Effects of Amenities on Living Costs

Dependent Variable: Living Costs (Number of Observations =286)	With Census Regions		With Census Divisions	
	Model 1		Model 2	
	Coefficient	t-statistic	Coefficient	t-statistic
INTERCEPT	5.3343	2.35	4.7130	2.09
MEAN PM10	0.0009	0.35	0.0014	0.55
TOTAL CRIME RATE	-0.8900	-1.02	-1.5127	-1.82
POP DENSITY PER SQ MILE OF LAND	0.0000	3.91	0.0000	4.55
TRANSPORTATION	-0.0021	-3.77	-0.0020	-3.76
EDUCATION	0.0007	1.22	0.0007	1.27
ARTS	0.0029	5.08	0.0025	4.65
HEALTHCARE	0.0013	2.87	0.0015	3.55
RECREATION	0.0006	1.10	0.0009	1.57
MSA ON THE COAST	0.0950	3.58	0.0784	3.10
WINTER TEMP AVG	0.0173	1.90	0.0187	2.05
WINTER TEMP AVG SQUARED	-0.0001	-0.84	-0.0001	-1.06
SUMMER TEMP AVG	0.0279	0.45	0.0396	0.63
SUMMER TEMP AVG SQUARED	-0.0003	-0.78	-0.0004	-0.91
WINTERPR	0.0016	0.18	-0.0052	-0.57
WINTERPR SQUARED	-0.0003	-0.80	-0.0001	-0.20
SUMMERPR	-0.0396	-2.36	-0.0193	-1.06
SUMMERPR SQUARED	0.0013	1.97	0.0007	0.98
NORTH EAST[57]	-0.1123	-3.07		
MID WEST	-0.1746	-4.04		
WEST	-0.1082	-1.31		
MIDDLE ATLANTIC[58]			-0.1141	-0.96
EAST NORTH CENTRAL			-0.1513	-1.24
WEST NORTH CENTRAL			-0.1544	-1.11
SOUTH ATLANTIC			-0.1778	-1.50
EAST SOUTH CENTRAL			-0.2939	-2.45
WEST SOUTH CENTRAL			-0.1999	-1.71
MOUNTAIN			-0.0915	-1.16
NEW ENGLAND			0.1309	1.14
Adjusted R-Squared	0.7484		0.7623	
R-Squared	0.7660		0.7831	

[57] The left out category is the SOUTH.
[58] The left out category includes the PACIFIC division

Table 5.11 Marginal Effects of Climate Variables on Living Costs

Marginal Effects on Living Costs	Model 1		Model 2	
	Coefficient	t-statistic	Coefficient	t-statistic
WINTER TEMP	0.0104	3.53	0.0100	3.30
SUMMER TEMP	-0.0205	-3.82	-0.0175	-3.05
WINTERPR	-0.0036	-1.02	-0.0065	-1.54
SUMMERPR	-0.0104	-2.43	-0.0045	-0.78

5.3.5 Marginal Willingness to Pay

This section presents estimates of willingness to pay for climate variables, i.e. the marginal rate of substitution between wage earnings and climate based on the systematic portion of the household's utility function. Calculating a complete welfare measure would entail looking at the impact of a change in the vector of amenities on the expected utility. However, this is not the focus of this study.

In this study, I have used a quadratic form for the function g(.). i.e., $g(E_j) = \alpha_0 E_j + \alpha_1 (E_j)^2$. The marginal rate of substitution is thus $((\alpha_0 + 2\alpha_1(E_j))/\ \alpha_C + \alpha_H\ ))^* W_{ij}$. This is the MWTP of household i for climate amenity E_j. It should be noted that this depends on the value of E_j as well as the sign of the alphas. Thus, for example, a household which is already experiencing a low winter temperature would be willing to pay more to increase it by a degree, than a household experiencing a higher winter temperature. The ratio of the marginal effect to the coefficient on the wage variable obtained in the first stage represents the percentage of income that a household would be willing to pay for a 1 unit change around the mean.

Using the estimates from the movers models in Table 5.7, the percentage of income that households are willing to pay for a unit change in the variable around the mean is calculated. A household facing an average winter temperature of 37 degrees is willing to pay about 2.99% of their annual income to raise it by a degree and 2.28% to lower winter precipitation by an inch from its mean level of 9 inches for the model with Census divisions. Households are willing to pay 3.86% of their annual income to raise summer precipitation by an inch from a level of about 11 inches. For example, a household with mean earnings of 45,000 dollars would be willing to pay 1,300 dollars to raise average winter temperature from 37 degrees by a degree. However, a household which is in a warmer MSA would be willing to pay much less. Table 5.12 provides examples of the marginal willingness to pay to raise average winter temperature in a few MSAs. This table illustrates that MWTP is lower for warmer MSAs.

Table 5.12 WTP to Increase Average Winter Temperature by a Degree: Examples

Name of MSA	**Average Winter Temperature**	**WTP** (as a % of income)
Washington, DC/MD/VA	34.83	3.35
Memphis, TN/AR/MS	40.02	2.62
Greenville, NC	44.67	1.97
San Jose, CA	50.03	1.21
Houston Brazoria, TX	54.17	0.63
Orlando, FL	60.11	-0.20
Naples, FL	65.53	-0.97

Chapter 6: Conclusions

There is a large literature that has attempted to value climate amenities or to estimate the role that they play in migration decisions in the United States. This dissertation contributes to this literature by modeling the location choices of households who changed MSAs between 1995 and 2000. The results provide estimates of the rate at which movers substitute income for temperature and precipitation—of marginal willingness to pay for changes in these climate variables. The results could also be used to simulate the impact of a counterfactual climate scenario on the migration patterns of households in the U.S.

A Random Utility Model framework is used to characterize the utility that a household derives from living in an MSA. This utility depends on climate amenities along with household earnings, housing costs and other location-specific amenities. Households choose the MSA that maximizes their utility. A two step procedure is used to estimate the model. In the first stage, location-specific constants (one for each MSA) are estimated together with other parameters of the utility function. In the second stage, the constants are regressed on location-specific amenities to estimate the average utility attached to these amenities. The model also allows for migration costs.

Two questions arise when estimating the preferences of movers for climate amenities. Do stayers have the same preferences as movers? Can the preferences of stayers be estimated based on their location choices? For the latter to be possible, it must be the case that stayers are in equilibrium. While it is difficult formally to test the hypothesis that stayers are in equilibrium, some information can be provided by estimating the migration model using both movers and stayers and comparing the results with estimates using movers only. The migration model is therefore estimated for all households in the PUMS sample, including those who did not change MSAs between 1995 and 2000, i.e., “stayers.” However, the low coefficient on wages in the first stage results suggests that stayers are not in equilibrium. Thus, I focus on results from the movers sample.

The results show that households facing an average winter temperature of 37 degrees Fahrenheit are willing to pay approximately about 3% of their income for an increase in average winter temperature by one degree. The value of changes in winter temperature however, declines as temperature increases. Indeed, my results suggest the optimum winter temperature is between 53 and 59 degrees. Households are willing to pay 2% of their income to lower winter precipitation by an inch from its mean level of 9 inches. Raising summer precipitation by one inch from a mean level of 11 inches is worth roughly 3% of income.

The study also provides estimates of quality of life indices for 297 Metropolitan Areas. These indices, which capture the value of locational amenities net of housing costs, are often used in the urban economics literature to rank cities.

Future Research

My future research will involve incorporating a more complete set of amenities. Some of variables to consider may be altitude and visibility, though the latter is highly correlated with pollution and humidity. Explicit data on snowfall might be more appropriate as many cities with high precipitation are in the South. These numbers, therefore reflect rainfall rather than snowfall. Snowfall may be valued differently from rainfall.

Estimating preferences for different population groups is also part of my research agenda. Preferences are likely to vary across different demographic groups. A random parameters model could be used to estimate a distribution of preferences.

Using a utility function that allows the marginal utility from consuming a housing bundle to be a function of household characteristics might also be interesting.[59] These characteristics could include the size of household, number of children and marital status of household head. For example, larger households may consume more housing than smaller households. Household characteristics may also affect moving costs. Married couples might find it more difficult to move due to the problem of supporting dual careers. Households with children might potentially have more difficulty moving.

Finally, the estimates from this model could potentially be linked to General Circulation Models to predict how different climate change scenarios might affect migration. Thus, in addition to providing welfare measures of climate changes, this study can be used to estimate the effects of climate change on the migration patterns of households.

[59] I have estimated initial versions of this model but this needs to be explored further.

Appendices

Appendix A.2[60]

The following model is presented to illustrate the bias that can arise if households adjust slowly in response to shocks to the economy. In this model, I abstract from considering firm behavior and land markets. It is assumed that the labor demand function for any particular region is given. To simplify the model, linear functional forms are used and I assume that there are only two regions in the economy.

i=A,B (2 regions)

t=1,…..,n (time period)

$$V_t^i = \alpha_w w_t^i + \alpha_E E_t^i \quad (A2.1)$$

$$w_t^i = \beta^i - \beta_p^i P_t^i \quad (A2.2)$$

$$P_t^i = \gamma P_{t-1}^i + (1-\gamma) P^{i*} \quad (A2.3)$$

$$\overline{P} = P^A + P^B \quad (A2.4)$$

$$V^{A*} = V^{B*} \quad (A2.5)$$

Equation A2.1 represents the indirect utility function of a representative household living in region i at time t. It is a function of wages and amenities, and I assume α_w ,$\alpha_E > 0$. Equation A2.2 depicts the labor demand for region i at time t. β^i reflects different productivities in different regions. This could be due to certain region specific factors such as the availability of cheap raw materials. It is also assumed that $\beta^i > 0$ and $\beta_p^i > 0$. Given the population, P_t^i, equation A2.2 determines the wage in region i.

[60] Appendix A.2 corresponds to Chapter 2.

Amenity levels (E_t^i) are exogenously given. A partial adjustment model is used to describe the process of adjustment of households; specifically, the population in a region changes according to equation (A2.3). Thus, population in the current period is a linear combination of population in the previous period and the equilibrium population, which is determined by the condition that utility is the same in both regions (Equation A2.5). Here (1-γ) reflects the speed of adjustment of households where γ is a positive fraction. The total population remains constant over time as given by (A2.4).

Suppose the system is in equilibrium at time 0, i.e., utility in region A is the same as utility in region B. Setting $V_0^A = V_0^B$, the equilibrium levels of population in both regions are obtained. They are given by the following equations:

$$P_0^A = \frac{\alpha_w(\beta^A - \beta^B) + \alpha_w \beta_p^B \overline{P} + \alpha_E (E_0^A - E_0^B)}{\alpha_w(\beta_p^A + \beta_p^B)} \tag{A2.6}$$

$$P_0^B = \overline{P} - P_0^A = \frac{\alpha_w \beta_p^A \overline{P} - \alpha_w(\beta^A - \beta^B) - \alpha_E (E_0^A - E_0^B)}{\alpha_w(\beta_p^A + \beta_p^B)} \tag{A2.7}$$

In equation (A2.6), $(\beta^A - \beta^B)$ can be interpreted as representing the location-specific differences in income, $\beta_p^B \overline{P}$ represents the maximum impact of population on wages (in region B) and ($E_0^A - E_0^B$) represents the impact of differences in environment.

Suppose that there is a one time exogenous shock in region A and E_0^A increases to E_1^A but there is no change in the value of the exogenous amenity in region B. Utility in region A goes up relative to that of region B in period 0. This difference in utilities causes households to migrate from region B to Region A. However, since there is a lag in adjustment, not all people migrate instantaneously. The influx of people into region A from region B causes wages to drop in region A and rise in region B. Thus utility in region A (region B) now falls (increases) due to a fall (rise) in wages. To determine the equilibrium populations after the shock, the equilibrium utility in region A after the shock (V^{A*}) must equal that in region B after the shock (V^{B*}). Thus

$$V^{A*} = V^{B*} \Rightarrow \alpha_E (E_0^B - E_1^A) + \alpha_w \{\beta^B - \beta^A - \beta_p^B \overline{P} + (\beta_p^A + \beta_p^B) P^{A*}\} = 0 \qquad \text{(A2.8)}$$

Solving equation (A2.7), the value of P^{A*} can be obtained. The expression for the differences in utility in period t is given by

$$V_t^B - V_t^A = \alpha_E (E_0^B - E_1^A) + \alpha_w [\beta^B - \beta^A - \{\beta_p^B \overline{P} - (\beta_p^A + \beta_p^B) P^{A*}\} - \gamma^t (\beta_p^A + \beta_p^B)(P^{A*} - P_0^A)]$$

(A2.9)

As γ lies between 0 and 1, the difference $V_t^B - V_t^A \rightarrow V^{B*}$ - V^{A*} (= 0) as t → ∞.Thus equilibrium is reached asymptotically.

At the margin, willingness to pay for the amenity E is (α_E / α_w).This is given in equilibrium by the ratio of the wage differences to the amenity differences between the regions. Table A2.1 provides two examples where this ratio is calculated assuming equilibrium after a shock occurs. The value of γ used is 0.75.[61] The true ratio is 0.5. However, as can be seen from the table, if a compensating differential is calculated assuming that the economy is in equilibrium, the results obtained for α_E / α_w are biased.

[61] In these two models, the assumptions made about the underlying parameters of the model are as follows: $\overline{P}$ = 10, $\beta^A = 3$, $\beta^B =$ 3, $\beta_p^A = 0.1$, $\beta_p^B = 0.2$, $\alpha_E = 0.2$ and α_w = 0.4 . The tolerance limit was set at 1%.

Table A2.1 Calculated Values of α_E / α_w in Different Periods Assuming Equilibrium.

t	(small shock)	(large shock)
0	0.5	0.5
1	4.2500	-0.9063
2	3.3125	-0.5548
3	2.6095	-0.2910
4	2.0820	-0.0933
5	1.6865	0.0550
6	1.3900	0.1663
7	1.1675	0.2498
8	1.0005	0.3123
9	0.8755	0.3593
10	0.7815	0.3944
11	0.7110	0.4209
12	0.6585	0.4406
13	0.6190	0.4554
14	0.5890	0.4666
15	0.5000	0.4749
16		0.5

Appendix A.4.1[62]

Table A4.1: Descriptive Statistics of Wage Income by Metropolitan Statistical Area

Metropolitan area	N	Mean	Std Dev
Abilene, TX	2037	24664.54	26251.41
Akron, OH	11264	35448.52	34732.95
Albany, GA	1688	29421.91	27436.88
Albany-Schenectady-Troy, NY	14922	35778.49	33428.78
Albuquerque, NM	10746	31259.20	30452.68
Alexandria, LA	1893	26236.90	28303.91
Allentown-Bethlehem-Easton, PA/NJ	9866	34113.59	29219.48
Altoona, PA	2187	27179.49	21399.15
Amarillo, TX	3578	28302.71	26655.61
Anchorage, AK	3435	36855.41	28366.99
Ann Arbor, MI	7281	43702.61	40033.01
Anniston, AL	1709	25783.02	24174.90
Appleton-Oskosh-Neenah, WI	6401	32907.23	25023.89
Asheville, NC	4000	29268.61	31124.87
Athens, GA	2999	27831.00	29419.25
Atlanta, GA	65455	40094.94	41217.57
Atlantic City, NJ	6059	33492.80	31494.87
Auburn-Opelika, AL	1692	25742.77	22998.40
Augusta-Aiken, GA-SC	6775	31028.02	29467.27
Austin, TX	20344	37782.70	40414.88
Bakersfield, CA	7353	31200.72	27919.37
Baltimore, MD	42120	39949.42	36733.91
Barnstable-Yarmouth, MA	2082	38234.76	40851.32
Baton Rouge, LA	8995	31415.80	30604.51
Beaumont-Port Arthur-Orange,TX	5616	30738.03	28940.35
Bellingham, WA	2707	29362.69	26253.25
Benton Harbor, MI	2900	32494.16	30022.98
Billings, MT	1763	29196.68	28189.49
Biloxi-Gulfport, MS	3937	27968.79	26561.10

[62] Appendix A.4.1 corresponds to Chapter 4, Section 4.1.

Binghamton, NY	4965	30777.98	28703.89
Birmingham, AL	12198	35704.71	37752.01
Bloomington, IN	1945	27044.81	30095.87
Bloomington-Normal, IL	3497	35080.31	31993.50
Boise City, ID	5760	31555.73	29588.61
Boston, MA	58365	47301.82	49559.63
Lawrence-Haverhill, MA/NH	4174	45244.41	49393.38
Lowell, MA/NH	5134	43253.35	38370.22
Bremerton, WA	3532	36067.05	31935.93
Bridgeport, CT	5639	49047.77	58265.31
Brockton, MA	4530	38081.51	31201.15
Brownsville-Harlingen-San Benito, TX	3968	22903.00	25885.79
Bryan-College Station, TX	2219	25708.52	34055.58
Buffalo-Niagara Falls, NY	18284	34409.58	31709.09
Canton, OH	7117	30812.69	28673.97
Cedar Rapids, IA	2961	34248.66	31015.66
Champaign-Urbana-Rantoul, IL	3586	30149.60	28586.59
Charleston-N.Charleston,SC	6601	31482.71	35548.78
Charlotte-Gastonia-Rock Hill, SC	24470	36287.56	37072.68
Charlottesville, VA	2992	31518.51	32946.00
Chattanooga, TN/GA	7126	31696.45	32900.60
Chicago-Gary-Lake, IL	119682	43019.76	43825.36
Gary-Hammond-East Chicago, IN	9576	34741.85	29878.20
Chico, CA	2941	29269.99	31892.12
Cincinnati OH/KY/IN	25105	37224.63	37769.78
Clarksville-Hopkinsville, TN/KY	1644	25754.62	23301.85
Cleveland, OH	38697	36564.08	35565.79
Colorado Springs, CO	8145	33514.48	32721.05
Columbia, MO	2189	28310.59	27945.52
Columbia, SC	8880	33048.05	32953.27
Columbus, GA/AL	2615	28626.49	30514.11
Columbus, OH	25768	36458.72	35082.29
Corpus Christi, TX	4044	29846.86	29154.01
Dallas-Fort Worth, TX	57202	40603.68	43913.86
Fort Worth-Arlington, TX	27229	36433.21	36933.00
Danbury, CT	3234	61087.30	69500.52

Danville, VA	2116	26072.80	24691.67
Davenport, IA Rock Island-Moline, IL	4159	31407.84	29263.91
Dayton-Springfield, OH	15772	33206.25	29277.07
Daytona Beach, FL	6586	27783.41	26983.57
Decatur, AL	2516	31089.83	27428.08
Decatur, IL	2548	32392.11	30698.85
Denver-Boulder-Longmont, CO	33716	39490.79	38881.89
Boulder-Longmont, CO	3692	41461.64	44972.06
Des Moines, IA	4749	35311.55	33961.63
Detroit, MI	60619	41750.70	38738.62
Dothan, AL	2746	26345.54	24502.45
Dover, DE	2416	29691.96	25274.03
Duluth-Superior, MN/WI	3520	31016.62	25902.04
Dutchess Co., NY	4768	41285.76	37286.33
Eau Claire, WI	2637	28011.15	24257.91
El Paso, TX	7869	25175.52	27814.00
Elkhart-Goshen, IN	2981	31772.86	29680.03
Erie, PA	4790	29278.67	25222.81
Eugene-Springfield, OR	4705	29441.47	27949.77
Evansville, IN/KY	4506	31571.17	30176.88
Fargo-Morehead, ND/MN	1653	28191.07	23966.91
Fayetteville, NC	3168	27270.46	26647.62
Fayetteville-Springdale, AR	5517	28314.05	28317.10
Fitchburg-Leominster, MA	2573	34455.20	27188.95
Flagstaff, AZ-UT	2346	27836.24	27314.29
Flint, MI	2995	29614.47	26091.53
Florence, AL	2263	28622.65	27352.36
Fort Collins-Loveland, CO	4098	34842.57	32349.23
Fort Lauderdale-Hollywood-Pompano Beach, FL	24182	36929.42	38380.17
Fort Myers-Cape Coral, FL	6522	30653.85	32848.93
Fort Pierce, FL	4229	32094.80	36682.22
Fort Smith, AR/OK	2391	26994.04	27194.56
Fort Walton Beach, FL	2578	28529.93	27203.55
Fort Wayne, IN	9033	32601.94	28483.55
Fresno, CA	11209	29686.59	29739.25
Gadsden, AL	1549	26618.01	22730.06

Gainesville, FL	4076	29493.15	32175.74
Galveston-Texas City, TX	4420	37357.84	35947.23
Glens Falls, NY	3052	28578.72	25637.63
Goldsboro, NC	1517	26201.24	24009.71
Grand Rapids, MI	13813	35065.13	31930.36
Grand Junction, CO	1565	27883.10	27489.28
Greeley, CO	3846	30614.57	27941.86
Green Bay, WI	3602	34232.25	30349.57
Greensboro-Winston Salem-High Point, NC	22648	32655.59	32529.56
Greenville, NC	2521	28209.91	26192.78
Greenville-Spartanburg-Anderson SC	13122	31298.65	29873.47
Hagerstown, MD	2670	31835.24	26729.75
Hamilton-Middleton, OH	5632	34951.64	32076.30
Harrisburg-Lebanon-Carlisle, PA	11515	32335.31	27997.07
Hartford-Bristol-Middleton-New Britain, CT	12335	43543.95	45457.22
Hickory-Morgantown, NC	6711	28181.91	25522.41
Hattiesburg, MS	1452	26128.68	27660.27
Honolulu, HI	12072	34940.59	28973.04
Houma-Thibodoux, LA	1183	27110.88	26278.76
Houston-Brazoria, TX	60278	38182.09	41064.48
Brazoria, TX	4512	35639.00	30890.87
Huntsville, AL	5665	35046.01	33208.09
Indianapolis, IN	24265	35904.61	33876.69
Iowa City, IA	1779	32228.11	35076.05
Jackson, MI	2560	32986.45	28284.45
Jackson, MS	6369	30549.62	33458.67
Jackson, TN	1704	30472.10	37929.45
Jacksonville, FL	18478	33947.56	36049.13
Jacksonville, NC	1525	22529.95	18068.68
Jamestown-Dunkirk, NY	3275	27676.41	25020.32
Janesville-Beloit, WI	2501	33755.29	28102.04
Johnson City-Kingsport-Bristol, TN/VA	5069	28523.39	29320.02
Johnstown, PA	5249	25284.71	22193.13
Joplin, MO	2937	24883.85	22879.83
Kalamazoo-Portage, MI	7907	32521.88	30067.47
Kankakee, IL	2117	32178.08	24421.59

Kansas City, MO-KS	26925	36177.88	33277.01
Kenosha, WI	2088	34680.46	28413.50
Kileen-Temple, TX	4063	26586.81	27259.42
Knoxville, TN	9736	31141.56	31216.27
Kokomo, IN	1928	36482.38	29537.31
LaCrosse, WI	1640	28869.38	28725.88
Lafayette, LA	3517	29246.05	30176.34
Lafayette-W. Lafayette, IN	3354	28272.62	28874.90
Lake Charles, LA	2859	29708.31	28610.59
Lakeland-Winterhaven, FL	7258	28869.29	29432.07
Lancaster, PA	6925	33031.52	28551.76
Lansing-E. Lansing, MI	8487	34683.55	29765.44
Laredo, TX	2025	22003.61	21710.06
Las Cruces, NM	1984	23813.03	23085.59
Las Vegas, NV	23475	33018.61	33180.09
Lexington-Fayette, KY	3899	34658.84	37606.16
Lima, OH	3208	29658.93	20964.78
Lincoln, NE	3411	31501.62	29112.25
Little Rock-North Little Rock, AR	7910	30501.04	30910.04
Longview-Marshall, TX	2740	28385.34	29908.61
Los Angeles-Long Beach, CA	136506	36819.58	41369.85
Orange County, CA	47625	43170.17	45917.01
Louisville, KY/IN	17261	34898.83	33888.55
Lubbock, TX	4082	27671.28	29601.69
Lynchburg, VA	3661	29314.52	27387.59
Macon-Warner Robins, GA	5291	30682.48	29070.19
Madison, WI	6559	35966.83	32955.38
Manchester, NH	1477	32360.78	27643.93
Mansfield, OH	2185	29151.17	22735.05
McAllen-Edinburg-Pharr-Mission, TX	5494	20900.24	23023.31
Medford, OR	2449	28417.42	26656.42
Melbourne-Titusville-Cocoa-Palm Bay, FL	8006	33097.78	31523.70
Memphis, TN/AR/MS	14407	35974.85	40763.78
Merced, CA	2443	27627.37	24071.77
Miami-Hialeah, FL	29666	31472.55	36850.65
Milwaukee, WI	18870	38099.61	37347.61

Minneapolis-St. Paul, MN	41118	40819.40	38053.97
Mobile, AL	7396	29191.97	29934.79
Modesto, CA	6279	31890.44	30318.91
Monmouth-Ocean, NJ	18332	47600.29	47350.90
Monroe, LA	2172	27338.72	30548.08
Montgomery, AL	4666	30377.37	30507.07
Muncie, IN	2105	27694.39	23924.97
Myrtle Beach, SC	2797	26921.68	25207.49
Naples, FL	2874	36138.66	45139.92
Nashua, NH	1624	41920.38	35701.73
Nashville, TN	20799	35146.89	36703.18
New Bedford, MA	2993	31836.99	30353.49
New Haven-Meriden, CT	5654	39599.01	37519.69
New Orleans, LA	17297	30787.18	31514.02
New York-Northeastern NJ	110327	44194.61	52009.57
Nassau Co, NY	46483	52384.79	56286.56
Bergen-Passaic, NJ	22127	48812.93	51066.09
Jersey City, NJ	9575	35911.38	34665.00
Middlesex-Somerset-Hunterdon, NJ	22676	50187.79	47944.24
Newark, NJ	34701	49299.02	52073.32
Newburgh-Middletown, NY	5542	41019.09	36726.82
Norfolk-VA Beach-Newport News, VA	22315	30650.37	29394.94
Ocala, FL	3380	26421.15	25878.54
Odessa, TX	3499	29242.64	30627.85
Oklahoma City, OK	11405	29983.82	29805.91
Olympia, WA	3437	34175.15	29993.67
Omaha, NE/IA	6837	34567.55	33235.51
Orlando, FL	27534	33563.37	35300.47
Panama City, FL	2402	28011.11	31432.22
Pensacola, FL	6128	28836.09	30006.73
Peoria, IL	7416	33624.87	29356.43
Philadelphia, PA/NJ	73366	41361.40	40071.23
Phoenix, AZ	48928	36341.75	36313.50
Pittsburgh-Beaver Valley, PA	35627	34334.71	34712.67
Portland, ME	2786	35277.33	35838.82
Portland-Vancouver, OR	29339	37249.56	34552.47

Providence-Fall River-Pawtucket, MA/RI	17113	33438.68	29283.56
Provo-Orem, UT	5053	29807.33	30091.60
Pueblo, CO	2010	26924.11	24218.34
Punta Gorda, FL	1630	29352.82	33189.69
Racine, WI	2339	37486.73	34088.02
Raleigh-Durham, NC	23274	38121.93	35629.90
Reading, PA	6760	33372.90	26945.19
Redding, CA	2545	30679.37	29427.62
Reno, NV	6470	34159.57	36634.53
Richland-Kennewick-Pasco, WA	2871	34329.17	29068.54
Richmond-Petersburg, VA	17883	36745.28	36872.44
Riverside-San Bernadino, CA	41992	33453.28	30702.65
Roanoke, VA	4217	31141.15	30671.19
Rochester, MN	2020	37720.73	35230.61
Rochester, NY	19407	34949.15	32173.86
Rockford, IL	5590	34388.54	28689.41
Rocky Mount, NC	2638	28038.27	25506.52
Sacramento, CA	26206	37766.02	34283.24
Saginaw-Bay City-Midland, MI	7272	33269.45	28697.44
St. Cloud, MN	3702	29352.73	23967.19
St. Joseph, MO	1826	27551.80	23982.19
St. Louis, MO-IL	39389	35905.14	35469.22
Salem, OR	4673	29884.43	24581.07
Salinas-Sea Side-Monterey, CA	3792	33270.58	30566.63
Salt Lake City-Ogden, UT	18693	33509.64	31459.94
San Antonio, TX	21098	30441.57	31341.99
San Diego, CA	41080	38105.76	40528.93
San Francisco-Oakland-Vallejo, CA	29888	53018.04	56785.43
Oakland, CA	40156	47770.27	47402.47
Vallejo-Fairfield-Napa, CA	8040	38049.31	32601.41
San Jose, CA	29321	55612.84	55093.91
San Luis Obispo-Atascad-P Robles, CA	3450	33551.78	34562.17
Santa Barbara-Santa Maria-Lompoc, CA	5934	37577.40	41734.92
Santa Cruz, CA	4239	41978.06	44903.91
Santa Fe, NM	2282	34621.25	30654.45
Santa Rosa-Petaluma, CA	7504	40580.80	38356.82

Sarasota, FL	7757	33033.18	36962.67
Savannah, GA	3191	32510.24	35390.84
Scranton-Wilkes-Barre, PA	12129	29749.67	26022.43
Seattle-Everett, WA	39354	42600.50	42128.90
Sharon, PA	2636	27268.56	21969.50
Sheboygan, WI	2097	33647.80	27161.36
Shreveport, LA	6000	28594.92	30421.96
Sioux City, IA/NE	1551	29571.11	29871.46
Sioux Falls, SD	1473	30388.66	29709.18
South Bend-Mishawaka, IN	4709	31338.58	28989.72
Spokane, WA	6673	30998.08	30572.28
Springfield, IL	1895	33373.02	29740.68
Springfield, MO	5503	27341.50	28945.27
Springfield-Holyoke-Chicopee, MA	9726	32939.47	29321.26
Stamford, CT	5234	80274.38	97703.75
State College, PA	2895	27594.58	28491.80
Stockton, CA	8008	33514.17	29812.36
Sumter, SC	1365	24846.37	20313.52
Syracuse, NY	16411	32337.01	29668.38
Tacoma, WA	11818	34669.07	29776.53
Tallahassee, FL	5263	30764.06	31026.86
Tampa-St. Petersburg-Clearwater, FL	36626	33093.10	34760.02
Terre Haute, IN	2941	28063.17	27247.53
Toledo, OH/MI	10689	32832.44	29957.26
Topeka, KS	2413	30768.57	25953.76
Trenton, NJ	5677	48839.08	53098.53
Tucson, AZ	12538	30701.25	31953.85
Tulsa, OK	8921	32282.62	31833.24
Tuscaloosa, AL	2161	29045.21	28136.56
Tyler, TX	2737	29705.26	30539.35
Utica-Rome, NY	6838	28523.87	23305.03
Ventura-Oxnard-Simi Valley, CA	11700	43180.81	43590.10
Vineland-Milville-Bridgetown, NJ	2456	30873.38	26557.97
Visalia-Tulare-Porterville, CA	4802	27761.76	26111.97
Waco, TX	3990	28505.14	32892.99
Washington, DC/MD/VA	87862	47140.09	43764.65

Waterbury, CT	1669	31952.25	26571.05
Waterloo-Cedar Falls, IA	1798	28014.56	25798.25
Wausau, WI	2818	32154.85	28713.63
West Palm Beach-Boca Raton-Delray Beach, FL	15764	39322.84	45036.13
Wichita, KS	8327	32333.30	26257.52
Wichita Falls, TX	1940	25379.58	25187.62
Williamsport, PA	3360	28096.63	24645.07
Wilmington, DE/NJ/MD	8850	40816.49	38600.26
Wilmington, NC	3793	31032.31	33345.84
Worcester, MA	4710	37291.20	32976.63
Yakima, WA	2723	28199.20	27750.90
Yolo, CA	2625	34069.85	31445.14
York, PA	6993	32774.57	25308.91
Youngstown-Warren, OH-PA	9653	30423.14	26645.83
Yuba City, CA	1943	29495.62	29789.44
Yuma, AZ	1716	25058.91	26461.08

Table A4.2: Descriptive Statistics of Variables used in Wage Hedonics

Variable	Mean [63]	Std Dev across MSAs
Wage rate	18.72058	2.89765
Highschool (left out category is no high school)	0.25459	0.06621
Some college	0.32715	0.04428
College graduate	0.20531	0.04575
Higher education	0.10847	0.03678
age	40.38501	0.95586
married	0.60153	0.03852
male	0.53759	0.02290
black	0.10777	0.09083
other	0.11476	0.08739
Speaks english	0.97052	0.02339
hispanic	0.10895	0.13363
businessoperations occupation (left out category is Management occupation)	0.02659	0.00678
financialspecialists occupation	0.02689	0.00698
computerandmath occupation	0.03332	0.01562
engineering occupation	0.02769	0.01116
lifephysicalsocialsc occupation	0.01068	0.00680
socialservices occupation	0.01696	0.00422
legal occupation	0.01165	0.00456
teachers occupation	0.04543	0.01241
othereduc occupation	0.00708	0.00300
artssportsmedia occupation	0.01696	0.00484
healthcarepractitioners occupation	0.04742	0.01233
healthcaresupport occupation	0.01784	0.00653
protectiveservices occupation	0.02302	0.00841
foodandserving occupation	0.03295	0.00985
maintenance occupation	0.02526	0.00615
personalcareservice occupation	0.01535	0.00679
highskillsales occupation	0.02971	0.00490
lowskillsales occupation	0.07146	0.01213
officesupport occupation	0.17473	0.01652
constructiontrades occupation	0.04777	0.01203
extractionworkers occupation	0.00046	0.00173
maintenanceworkers occupation	0.04234	0.00912
production occupation	0.08723	0.04801
transportation occupation	0.05506	0.01560
construction industry (left out category is mining and utilities)	0.05998	0.01465
manufacturing industry	0.16275	0.08509
wholesale industry	0.04211	0.01024
retail industry	0.10414	0.01782

[63] This mean is calculated using all observations in the 40 weeks sample.

Variable	Mean [63]	Std Dev across MSAs
transportation industry	0.04703	0.01417
informationcomm industry	0.03848	0.01229
finance industry	0.08348	0.02830
profscientificmngmntservices industry	0.09502	0.02647
educhealthsocialservices industry	0.19356	0.04456
recreationfoodservices industry	0.05927	0.02952
otherservices industry	0.03905	0.00600
publicad industry	0.06091	0.03485

Table A4.3: Descriptive Statistics of Variables used in Housing Price Hedonics

Variable	Mean	Std Dev
User Costs Includuing Insurance And Utility Costs	1361.39	1175.14
Own (=1 if the house is owned)	0.6497106	0.4770606
Bedroom3 (left out category is less than three bedrooms)	0.2548508	0.4357774
Bedroom4	0.3818743	0.4858461
Bedroom5	0.1581861	0.3649154
Bedroomgt5	0.0353323	0.1846183
Room2 (left out category is less than two rooms)	0.0512963	0.2206014
Room3	0.0993347	0.299111
Room4	0.1371428	0.343998
Room5	0.1904571	0.3926617
Room6	0.1884442	0.3910665
Roomgt6	0.3091816	0.4621563
Completekitchen	0.9937028	0.0791048
Completeplumbing	0.9945736	0.0734639
Acres1to10	0.6409834	0.4797121
Ageofstructure_0to1years (left out category is age of structure over 61 years	0.0187451	0.1356236
Ageofstructure_2to5years	0.0648173	0.2462032
Ageofstructure_6to10years	0.0687146	0.2529682
Ageofstructure_11to20years	0.1491919	0.3562776
Ageofstructure_21to30years	0.1754633	0.380363
Ageofstructure_31to40years	0.1470899	0.3541956
Ageofstructure_41to50years	0.1460088	0.3531151
Ageofstructure_51to60years	0.0790465	0.2698113
Unitsinstructure _Singleattached (left out category is units in structure single family detached	0.067834	0.2514608
Units_In_Structure _2	0.0483438	0.2144917
Units_In_Structure _3to4	0.0529012	0.2238362
Units_In_Structure _5to9	0.0522223	0.222475
Units_In_Structure _10to19	0.0457203	0.2088779
Units_In_Structure _20to49	0.0386258	0.1927014
Units_In_Structure _Over50	0.0640468	0.2448363

Figure A4.1

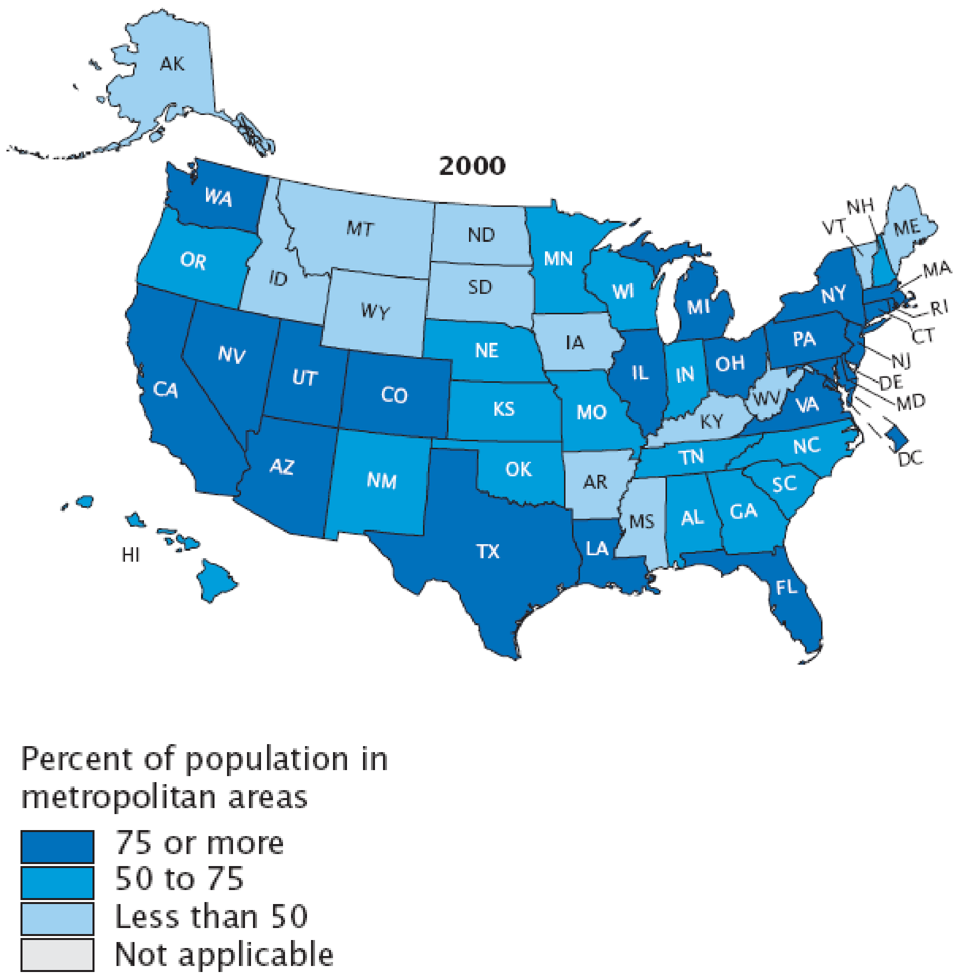

2000
AK
WA
OR
CA
ID
NV
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MS
MI
IN
KY
TN
AL
OH
WV
GA
FL
PA
VA
NC
SC
NY
VT
NH
ME
MA
RI
CT
NJ
DE
MD
DC
HI
Percent of population in metropolitan areas
75 or more
50 to 75
Less than 50
Not applicable

Figure A4.2

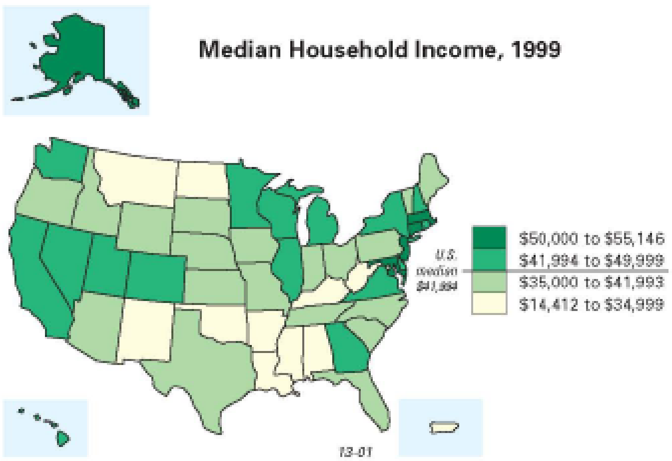


Figure A4.3

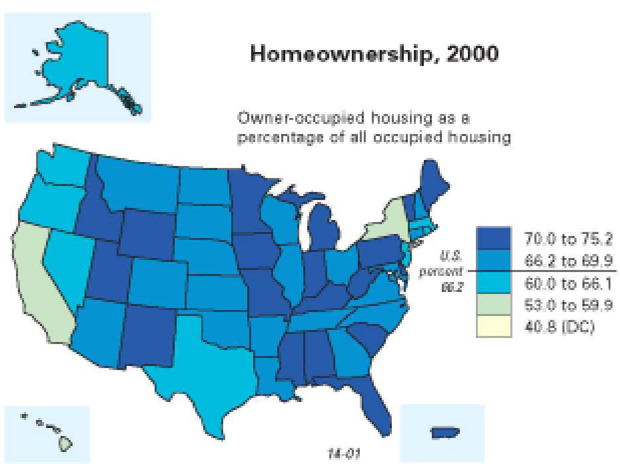


Figure A4.4

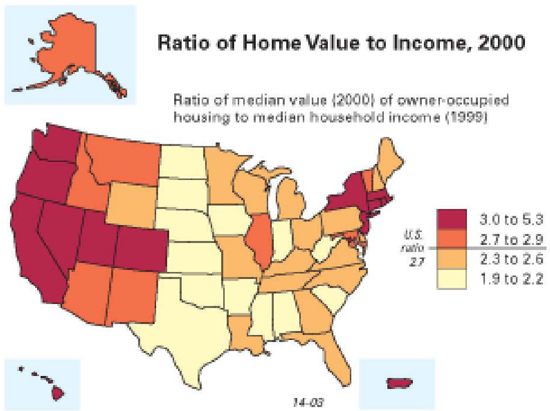


Figure A4.5

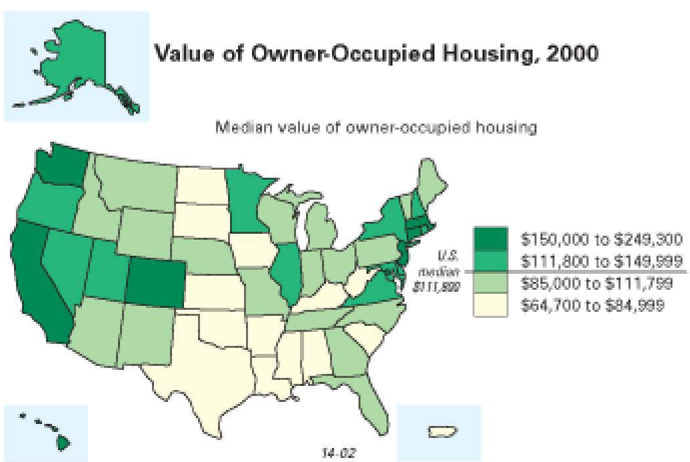


Figure A4.6

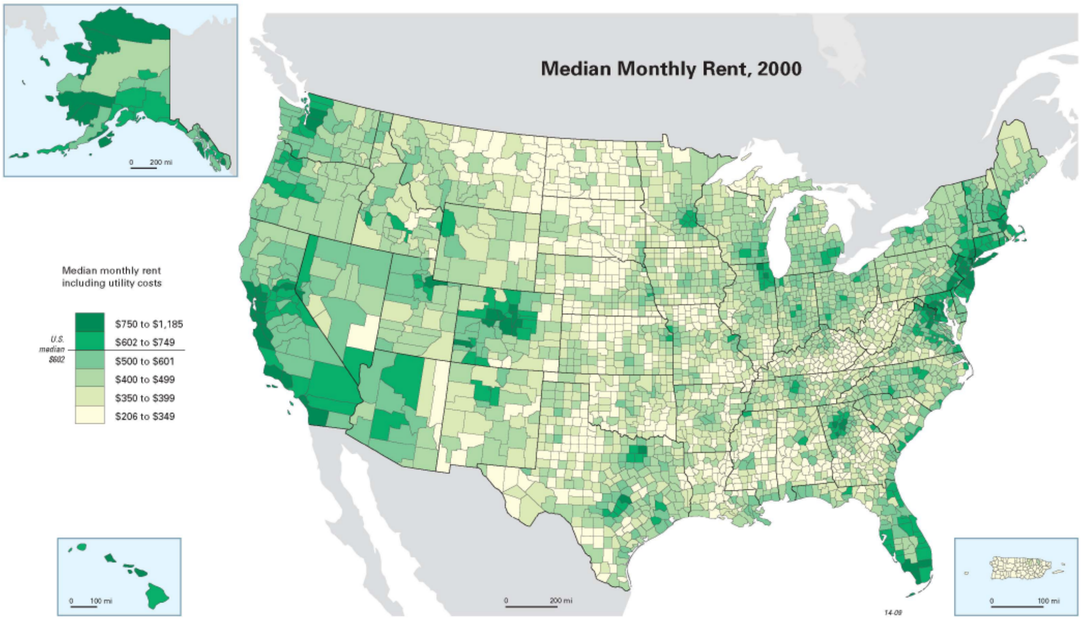


Figure A4.7

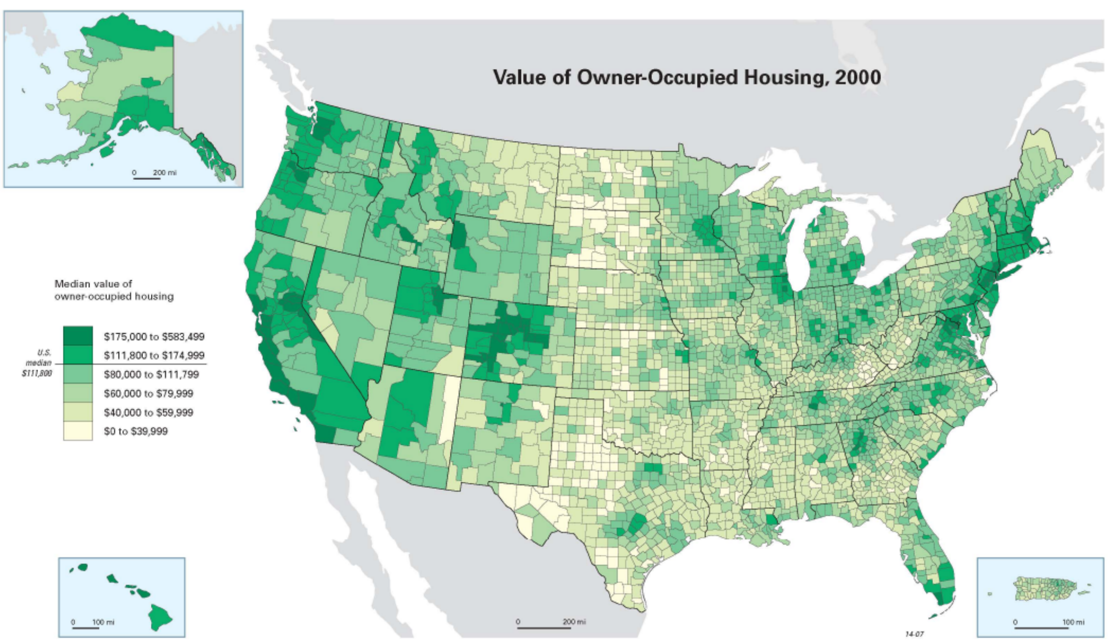
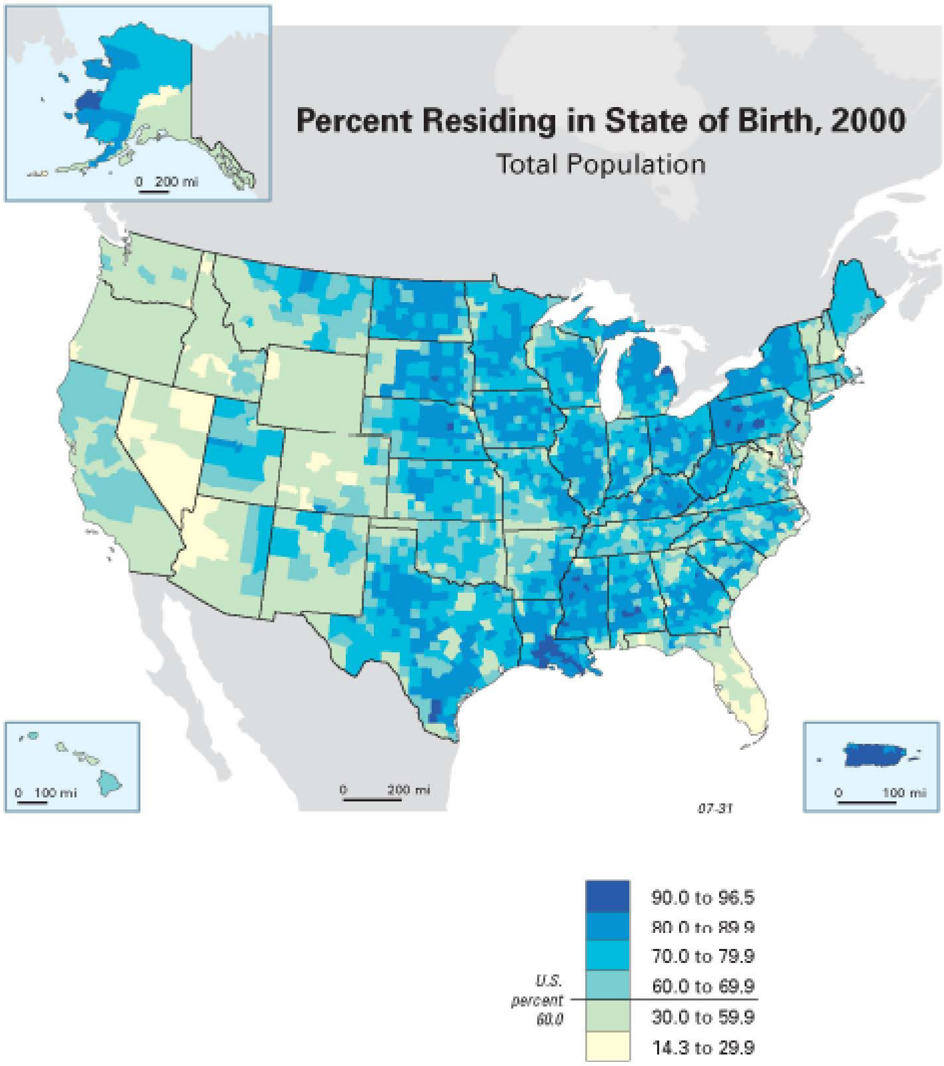


Figure A.4.8


Percent Residing in State of Birth, 2000
Total Population
0 200 mi
0 100 mi
0 200 mi
07-31
0 100 mi
90.0 to 96.5
80.0 to 89.9
70.0 to 79.9
U.S. percent 60.0
60.0 to 69.9
30.0 to 59.9
14.3 to 29.9

Appendix A.4.2[64]

For the climate and crime data, there are missing values for certain counties. I use an appropriate strategy to address these missing values. These strategies are described in sections A.4.2.1 and A.4.2.2. A description of variables used to create the rankings for Places Rated Almanac data is provided in Section A.4.2.3. Sections A.4.2.4 and A.4.2.5 provide the summary statistics of the amenities by Census Regions and Divisions respectively.

A.4.2.1. Climate

The geographical unit for the climate normals data is climate divisions. According to NOAA, "a climate division represents a region within a state that is as climatically homogeneous as possible." There are 344 climate divisions in the US. There can potentially be up to 10 climate divisions per state though some states have fewer. In general, climate divisions "coincide with county borders except in the western US, where they are based largely on drainage basins." A map showing the different climate divisions and counties in the US is shown in the following page.[65]

It is reasonable to assume that the households' perceptions of the climate in a county are not very different from their perceptions of the climate in neighboring counties. To obtain the population weighted average value for the MSA, if the climate division for a county is missing, the numbers of neighboring counties are used and the population shares of the other counties are re-scaled so that they add to 1. For

[64] Appendix A.4.2 corresponds to Chapter 4, Section 4.2.
[65] The map was generated using GIS software by Mike Squires, a meteorologist from NOAA.

independent cities located in the middle of a county (or in between two counties which are a part of the same climate division), the missing divisions are filled in with those of the surrounding county (or counties).

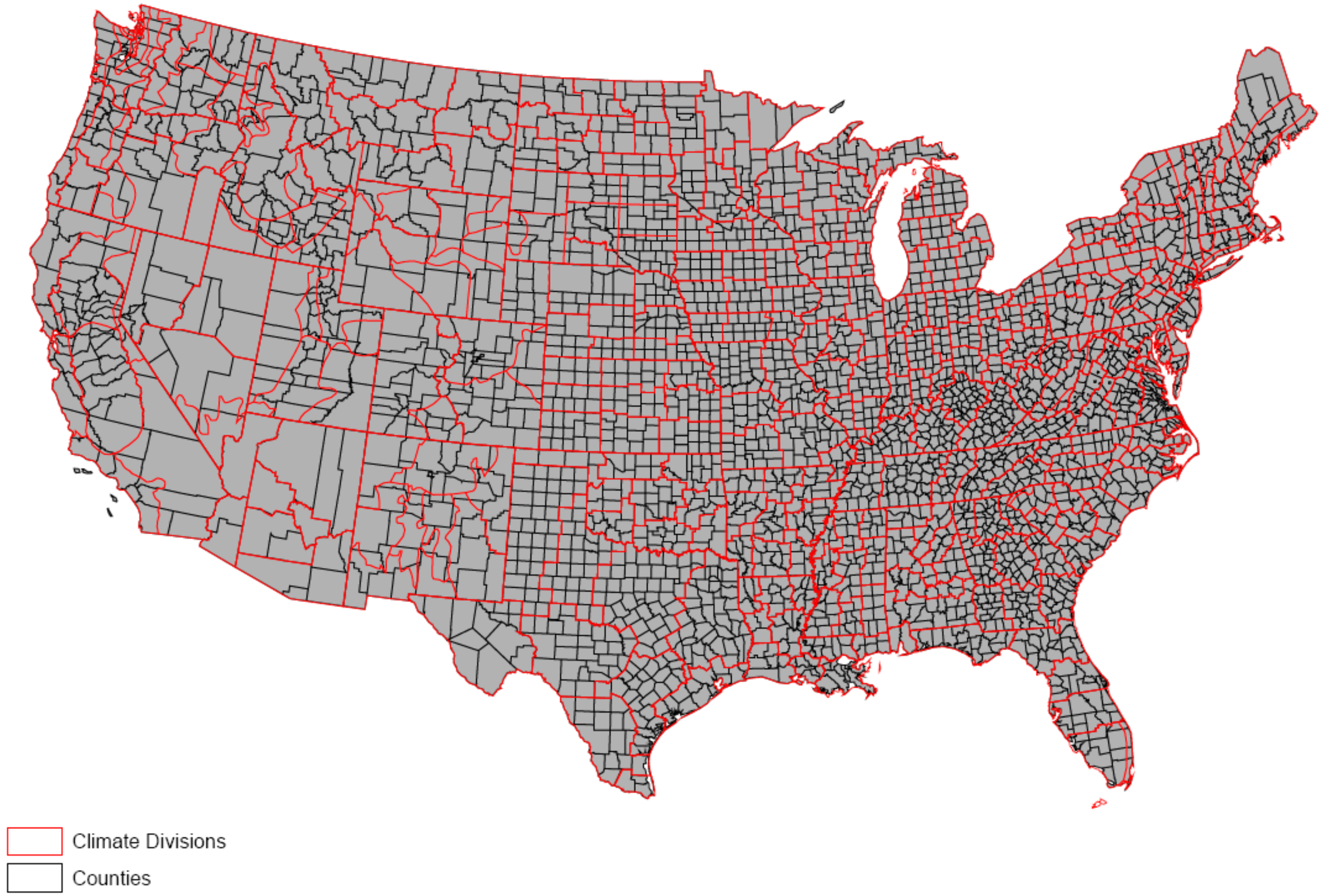

Climate Divisions
Counties

Table A4.4 Correlation between Monthly Temperature Values

Variable	JAN_TEMP	FEB_TEMP	MAR_TEMP	APR_TEMP	MAY_TEMP	JUN_TEMP	JUL_TEMP	AUG_TEMP	SEP_TEMP	OCT_TEMP	NOV_TEMP	DEC_TEMP
MEAN	34.89	38.53	46.14	54.42	63.29	71.02	75.16	73.86	67.55	57.29	46.95	38.11
STD	12.72	12.22	10.54	8.51	7.11	6.20	5.39	5.67	6.97	8.09	9.59	11.37
N	295.00	295.00	295.00	295.00	295.00	295.00	295.00	295.00	295.00	295.00	295.00	295.00
JAN_TEMP	1.00	0.99	0.98	0.92	0.83	0.72	0.69	0.77	0.89	0.94	0.97	0.99
FEB_TEMP	0.99	1.00	0.99	0.94	0.85	0.75	0.72	0.80	0.91	0.95	0.96	0.98
MAR_TEMP	0.98	0.99	1.00	0.98	0.92	0.84	0.81	0.87	0.95	0.98	0.97	0.98
APR_TEMP	0.92	0.94	0.98	1.00	0.98	0.92	0.89	0.94	0.99	0.99	0.96	0.94
MAY_TEMP	0.83	0.85	0.92	0.98	1.00	0.98	0.95	0.97	0.98	0.96	0.92	0.87
JUN_TEMP	0.72	0.75	0.84	0.92	0.98	1.00	0.99	0.98	0.95	0.89	0.83	0.77
JUL_TEMP	0.69	0.72	0.81	0.89	0.95	0.99	1.00	0.99	0.94	0.86	0.79	0.73
AUG_TEMP	0.77	0.80	0.87	0.94	0.97	0.98	0.99	1.00	0.97	0.91	0.85	0.80
SEP_TEMP	0.89	0.91	0.95	0.99	0.98	0.95	0.94	0.97	1.00	0.98	0.95	0.91
OCT_TEMP	0.94	0.95	0.98	0.99	0.96	0.89	0.86	0.91	0.98	1.00	0.98	0.96
NOV_TEMP	0.97	0.96	0.97	0.96	0.92	0.83	0.79	0.85	0.95	0.98	1.00	0.99
DEC_TEMP	0.99	0.98	0.98	0.94	0.87	0.77	0.73	0.80	0.91	0.96	0.99	1.00

A.4.2.2. Crime

The source of the crime data is the crimes reported file (Part 4) of the County-Level Detailed Arrest and Offense Data, 2000 (ICPSR 3451). These include murder, rape, robbery, aggravated assault, burglary, larceny, auto theft, and arson (classified by the FBI as Part I crimes).The data were originally collected by the Federal Bureau of Investigation (FBI) from reports submitted by agencies and states who participated in the Uniform Crime Reporting (UCR) Program.

To adjust for incomplete reporting by individual law enforcement jurisdictions, an imputation algorithm has been adopted by ICPSR. For each active Originating Agency Identifier (ORI) that reports less than 12 months of data, adjustments were made through weighting of partial year data or by substitution of a value based on population group and state. ICPSR has also created a Coverage Indicator (CI) to provide users with a diagnostic measure of aggregated data quality in a particular county. The CI represents the proportion of county data that is not imputed for a given year. This variable ranges from 100 (indicating that all ORIs in the county reported 12 months) to 0. In the crimes reported file, it was not possible for ICPSR to do any estimation for agencies reporting 0 months. All the crime variables will have a value of zero for these counties and thus a zero for the CI indicates missing data.[66]

In my dataset, there were 34 such counties. Other than Lincoln County in St. Louis, MO and St. Bernard Parish in New Orleans, LA, all the other counties are in

[66] For a detailed description of these adjustments and the construction of the CI, please refer to the Codebook for ICPSR 3451.

Kentucky and Illinois. Out of 28 counties in Illinois, 21 are missing crime data. If I delete all 10 MSAs which have missing county data, then I will be deleting the state of Illinois. Similarly, for Kentucky, 11 out of 18 counties are missing data. There are 5 MSAs in Kentucky in my dataset and if I delete MSAs which are missing county data, I will be deleting the state of Kentucky.

I have deleted MSAs which are missing data for any one county if the county has a "high" population share in the MSA.[67] I tried 3 alternative exclusion criteria. First, I deleted MSAs if the county for which data were missing had a population share greater than 0.3. Alternately, I used shares greater than 0.15 and 0.1 to see if the average values for the MSA changed using this stricter condition. Since the averages were not very different, I use a share of 0.3. This enables me to include 3 additional MSAs in my dataset. The MSAs which are I delete using each of the 3 criteria described above are listed in the following table.

	Share of county population with missing data in the MSA	**MSAs in Illinois**	**MSAs in KY**
1.	0.3	1040, 1400, 1960, 2040, 3740 and 6120	1660
2.	0.15	(1.) + 6800 and 7040	(1.) + 2440
3.	0.1	(2.) + 1600	(2.) + 4520

For the MSAs which have counties that have missing crime data but have "small" population shares, I re-weight the shares of the remaining counties such that

[67] Potentially, it could be the case that there are several counties that are missing data and have a "small" population share but the shares add up to a large number. However, this does not occur in Illinois or Kentucky.

they sum to 1. I use the crime data for these counties to calculate the population weighted average for the MSA.

A4.2.3 Places Rated Almanac

There are numerous aspects of a city that households value when choosing a residential location. The Places Rated Almanac provides scores and rankings for certain location specific characteristics for 354 metropolitan areas in North America.[68] The scores for these characteristics are intended to provide a measure of the attractiveness of a metropolitan area. This publication provides scores for nine location-specific characteristics including Transportation, Education, Arts, Healthcare and Recreation.[69] I use these scores as measures of location-specific amenities in my study. The Places Rated Almanac does not provide precise definitions of the factors that are used to rank each city. Nor does it describe how the numerical scores underlying the rankings are computed. It does, however, describe the factors that play a role in determining the city rankings.

<u>TRANSPORTATION</u>

“Three broad factors are used to rate each metro area for transportation:

[68] These include areas in Canada as well as the United States.
[69] It also includes scores for Cost of Living, Climate, Crime and Jobs.

(1) its supply of public transit and the typical time it takes to get to work and back;

(2) its connectivity with other metro areas via national highways, scheduled air service, and passenger rail service;

(3) its relative nearness to all other metro areas"

EDUCATION

There are 4 criteria that are used for ranking areas on the basis of education.

(1) School Support "combines metro area averages for the number of pupils per classroom teacher (the fewer the better) and the percent of funding the schools receive from local as opposed to state and federal- sources (the more the better)."

(2) Library Popularity

"The number of books on library shelves tells half the story of a place's reading habits. How much use those volumes get, or the metro area's circulation rate, is the other half. When the circulation figure is added to the number of volumes, and that sum divided by the population served, the result is Library Popularity."

(3) College Town is "enrollment weighted by number of years of typical attendance to get the highest degree offered" (e.g. "doctoral enrollment is multiplied by 9"). "This large number is then divided by the metro area's population".

(4) College Options is the "variety of higher education institutions that meet the need of the residents: low-cost night and weekend continuing education courses for people who work, full-time graduate courses in the professions, courses leading to occupational certification in 2-year colleges, and the traditional bachelor's degree curriculum offered in a college or university."

HEALTHCARE

The Places Rated Almanac "doesn't assess the quality of healthcare but its supply". There are the 5 criteria are used to rate the supply of healthcare in a metro area.

(1) General/Family Practitioners per 100,000

(2) Medical Specialists per 100,000

(3) Surgical Specialists per 100,000

(4) Accredited General Hospital Beds: This includes the number of hospital beds in short term general hospitals[70] that are accredited by the Joint Commission on Accreditation of Healthcare Organizations.[71]

(5) "Hospitals with physician teaching programs certified by the American Medical Association". This is included as facilities with physician training programs "tend to be larger urban institutions where the interaction between students and faculty encourages the development and use of latest techniques, equipment and therapy".

ARTS

The Almanac uses eight criteria to derive a metro area's score for arts[72]:

(1) Number of art museums

(2) Annual museum attendance

(3) Per capita museum

[70] "In the Short-Term General Hospitals category, Places Rated Almanac counts only hospitals classified by the American Hospital Association as acute care facilities whose patients stay fewer than 30 days."

[71] "In US metro areas, 91% of short-term general hospitals are accredited by the JCAHO."

[72] The first three are reflect collected art displayed in museums and gardens and the rest reflect live performances.

(4) Annual ballet performances

(5) Touring artist bookings

(6) Opera performances

(7) Professional theater performance

(8) Symphony performances

RECREATION

Thirteen criteria are used to rate a metro area's supply of recreation assets:

(1) Amusement and theme parks

(2) Aquariums

(3) Auto racing

(4) College sports

(5) Gambling

(6) Golf courses

(7) Good restaurants

(8) Movie theatre screens

(9) Professional sports

(10) Protected recreation areas

(11) Skiing

(12) Water area

(13) Zoos

A4.2.4. Amenities by Census Regions

North East Region						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	53	28.382	3.624	19.933	34.495	28.581
SUMMER TEMPERATURE	53	69.178	2.237	64.633	73.000	68.767
SUMMER CDD	53	158.505	51.266	74.667	252.000	144.046
WINTER HDD	53	1098.620	108.718	915.038	1351.500	1092.570
WINTER PRECIPITATION	53	10.163	1.263	7.672	12.393	10.494
SUMMER PRECIPITATION	53	11.991	0.826	10.505	13.450	11.985
ANNUAL CDD	53	547.066	193.514	235.500	907.000	486.096
ANNUAL HDD	53	6333.630	819.100	4972.010	8261.500	6354.810
TOTAL ANNUAL PRCP	53	45.544	2.997	38.528	50.820	45.980
DAYS WITH AVG TEMPERATURE <35	53	81.666	18.178	45.329	127.853	83.282
DAYS WITH AVG TEMPERATURE > 75	53	25.295	13.117	4.405	58.457	21.347
TRANSPORTATION	52	45.370	32.188	1.130	99.710	40.220
EDUCATION	52	55.073	31.927	1.690	99.710	58.640
ARTS	52	56.167	30.006	0.570	100.000	58.505
HEALTHCARE	52	48.568	28.729	0.000	94.900	49.430
RECREATION	52	56.560	25.766	0.000	98.300	59.485
TOTAL CRIME RATE	53	0.029	0.009	0.004	0.055	0.028
MSA OUT OF ATTAINMENT WITH NAAQS	53	0.302	0.463	0.000	1.000	0.000
MEAN PM 2.5	53	12.908	1.777	9.593	17.509	13.059
95TH PERCENTILE OF PM 2.5	53	27.122	3.924	20.063	36.638	27.655
MEAN PM 10	53	21.808	4.651	11.582	32.256	22.502
PM 10(95TH PERCENTILE)	53	42.687	9.511	24.959	65.517	43.119
POPULATION DENSITY PER SQUARE MILE	53	1100.280	2052.520	72.900	13043.600	579.200
POPULATION	53	885626.3	1494398.6	120044.0	9314235.0	381751.0
MSA ON COAST	53	0.434	0.500	0	1	0
GREATLAKES	53	0.075	0.267	0	1	0
PACIFIC	53	0	0	0	0	0
ATLANTICGULF	53	0.358	0.484	0	1	0
NEW ENGLAND	53	0.358	0.484	0	1	0
MIDDLE ATLANTIC	53	0.642	0.484	0	1	1
EAST NORTH CENTRAL	53	0	0	0	0	0
WEST NORTH CENTRAL	53	0	0	0	0	0
SOUTH ATLANTIC	53	0	0	0	0	0
EAST SOUTH CENTRAL	53	0	0	0	0	0
WEST SOUTH CENTRAL	53	0	0	0	0	0
MOUNTAIN	53	0	0	0	0	0
PACIFICDIV	53	0	0	0	0	0

Mid West Region						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	73	25.252	5.482	9.442	35.163	26.100
SUMMER TEMPERATURE	73	70.997	2.804	62.739	78.090	71.100
SUMMER CDD	73	214.215	67.610	59.759	408.636	210.333
WINTER HDD	73	1194.140	164.614	897.263	1670.550	1168.670
WINTER PRECIPITATION	73	5.616	2.128	1.540	11.876	5.960
SUMMER PRECIPITATION	73	11.822	0.998	9.430	13.645	12.050
ANNUAL CDD	73	794.899	281.172	192.714	1592.890	781.000
ANNUAL HDD	73	6598.730	1119.720	4655.820	9863.630	6440.000
TOTAL ANNUAL PRCP	73	36.359	4.966	21.253	49.657	37.034
DAYS WITH AVG TEMPERATURE <35	72	91.490	23.113	44.760	146.629	90.394
DAYS WITH AVG TEMPERATURE > 75	72	40.481	19.029	4.014	87.620	36.251
TRANSPORTATION	73	54.662	27.952	5.940	100.000	57.790
EDUCATION	73	60.246	26.015	5.380	99.150	60.900
ARTS	73	58.086	24.564	3.970	99.160	56.380
HEALTHCARE	73	50.672	30.031	1.980	97.730	47.870
RECREATION	73	52.776	26.820	3.680	99.710	49.570
TOTAL CRIME RATE	67	0.039	0.012	0.018	0.085	0.038
MSA OUT OF ATTAINMENT WITH NAAQS	73	0.137	0.346	0.000	1.000	0.000
MEAN PM 2.5	73	13.490	2.479	7.601	17.669	13.488
95TH PERCENTILE OF PM 2.5	73	27.794	3.682	18.141	34.932	27.792
MEAN PM 10	73	23.782	3.768	17.591	35.748	23.285
PM 10(95TH PERCENTILE)	73	45.849	8.978	32.817	84.209	44.605
POPULATION DENSITY PER SQUARE MILE	73	341.063	267.217	32.400	1634.200	256.200
POPULATION	73	657291.8	1176524.2	101541.0	8272768.0	250291.0
MSA ON COAST	73	0.205	0.407	0	1	0
GREATLAKES	73	0.205	0.407	0	1	0
PACIFIC	73	0	0	0	0	0
ATLANTICGULF	73	0	0	0	0	0
NORTHEAST	73	0	0	0	0	0
MIDDLE ATLANTIC	73	0	0	0	0	0
EAST NORTH CENTRAL	73	0.712	0.456	0	1	1
WEST NORTH CENTRAL	73	0.288	0.456	0	1	0
SOUTH ATLANTIC	73	0	0	0	0	0
EAST SOUTH CENTRAL	73	0	0	0	0	0
WEST SOUTH CENTRAL	73	0	0	0	0	0
MOUNTAIN	73	0	0	0	0	0
PACIFICDIV	73	0	0	0	0	0

South Region						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	112	47.711	8.602	31.533	67.922	46.585
SUMMER TEMPERATURE	112	78.888	3.074	70.833	84.800	79.306
SUMMER CDD	112	427.965	92.051	194.333	608.333	439.753
WINTER HDD	112	541.759	228.584	68.333	1004.000	558.667
WINTER PRECIPITATION	112	10.990	4.036	1.520	17.212	11.272
SUMMER PRECIPITATION	112	14.113	4.438	5.520	23.300	12.856
ANNUAL CDD	112	2145.880	831.247	706.000	4171.000	2021.140
ANNUAL HDD	112	2595.700	1290.490	240.667	5588.000	2580.920
TOTAL ANNUAL PRCP	112	49.035	10.775	13.190	66.590	51.155
DAYS WITH AVG TEMPERATURE <35	111	14.593	14.514	0.000	60.789	11.109
DAYS WITH AVG TEMPERATURE > 75	111	121.061	43.252	30.245	237.273	120.605
TRANSPORTATION	111	49.213	27.852	0.000	98.300	49.290
EDUCATION	111	46.270	28.183	0.280	100.000	45.890
ARTS	111	42.366	28.879	0.000	99.720	39.380
HEALTHCARE	111	49.481	28.975	0.840	98.300	52.120
RECREATION	111	53.687	29.539	0.280	100.000	55.800
TOTAL CRIME RATE	111	0.051	0.014	0.002	0.089	0.052
MSA OUT OF ATTAINMENT WITH NAAQS	112	0.018	0.133	0.000	1.000	0.000
MEAN PM 2.5	112	13.679	2.741	7.325	19.535	13.724
95TH PERCENTILE OF PM 2.5	112	26.287	5.865	13.757	42.941	26.047
MEAN PM 10	112	23.972	3.311	16.897	36.560	24.047
PM 10(95TH PERCENTILE)	112	43.727	7.559	27.874	73.718	43.655
POPULATION DENSITY PER SQUARE MILE	112	298.929	223.445	57.500	1346.500	213.350
POPULATION	112	647890.5	860337.4	103459.0	4923153.0	320987.5
MSA ON COAST	112	0.313	0.466	0	1	0
GREATLAKES	112	0	0	0	0	0
PACIFIC	112	0	0	0	0	0
ATLANTICGULF	112	0.313	0.466	0	1	0
NEW ENGLAND	112	0	0	0	0	0
MIDDLE ATLANTIC	112	0	0	0	0	0
EAST NORTH CENTRAL	112	0	0	0	0	0
WEST NORTH CENTRAL	112	0	0	0	0	0
SOUTH ATLANTIC	112	0.491	0.502	0	1	0
EAST SOUTH CENTRAL	112	0.179	0.385	0	1	0
WEST SOUTH CENTRAL	112	0.330	0.472	0	1	0
MOUNTAIN	112	0	0	0	0	0
PACIFICDIV	112	0	0	0	0	0

West Region						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	57	39.927	8.866	22.967	55.033	40.367
SUMMER TEMPERATURE	57	69.338	6.116	60.848	89.733	69.367
SUMMER CDD	57	186.479	152.245	32.005	760.000	159.475
WINTER HDD	57	755.478	265.154	315.000	1264.330	742.667
WINTER PRECIPITATION	57	10.422	8.049	1.500	28.084	9.780
SUMMER PRECIPITATION	57	3.062	1.988	0.440	7.242	2.070
ANNUAL CDD	57	784.310	754.133	111.783	3892.000	557.618
ANNUAL HDD	57	4677.090	1877.220	1343.000	8089.000	4170.340
TOTAL ANNUAL PRCP	57	25.498	16.697	5.080	66.747	19.140
DAYS WITH AVG TEMPERATURE <35	57	27.909	33.448	0.000	100.800	11.629
DAYS WITH AVG TEMPERATURE > 75	57	37.638	39.570	0.600	164.765	23.029
TRANSPORTATION	59	51.563	30.330	1.410	98.580	51.550
EDUCATION	59	44.946	29.519	0.000	98.010	41.350
ARTS	59	54.026	29.158	3.690	99.440	54.400
HEALTHCARE	59	43.495	26.521	0.280	95.460	43.900
RECREATION	59	46.780	31.316	4.810	96.880	43.340
TOTAL CRIME RATE	59	0.044	0.011	0.023	0.071	0.042
MSA OUT OF ATTAINMENT WITH NAAQS	59	0.525	0.504	0.000	1.000	1.000
MEAN PM 2.5	57	10.675	3.330	5.382	18.948	10.263
95TH PERCENTILE OF PM 2.5	57	27.595	11.088	9.389	58.177	26.951
MEAN PM 10	57	23.800	7.069	10.930	44.384	21.681
PM 10(95TH PERCENTILE)	57	47.655	14.556	19.124	96.148	42.574
POPULATION DENSITY PER SQUARE MILE	59	394.466	630.263	5.400	3605.600	153.400
POPULATION	59	921997.3	1425447.2	116255.0	9519338.0	399347.0
MSA ON COAST	59	0.339	0.477	0	1	0
GREATLAKES	59	0	0	0	0	0
PACIFIC	59	0.339	0.477	0	1	0
ATLANTICGULF	59	0	0	0	0	0
NEW ENGLAND	59	0	0	0	0	0
MIDDLE ATLANTIC	59	0	0	0	0	0
EAST NORTH CENTRAL	59	0	0	0	0	0
WEST NORTH CENTRAL	59	0	0	0	0	0
SOUTH ATLANTIC	59	0	0	0	0	0
EAST SOUTH CENTRAL	59	0	0	0	0	0
WEST SOUTH CENTRAL	59	0	0	0	0	0
MOUNTAIN	59	0.339	0.477	0	1	0
PACIFICDIV	59	0.661	0.477	0	1	1

A4.2.5. Amenities by Census Divisions

NEW ENGLAND DIVISION						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	19	27.721	3.053	21.533	31.433	27.623
SUMMER TEMPERATURE	19	68.508	1.457	65.183	70.933	68.561
SUMMER CDD	19	139.487	32.381	75.167	197.667	139.065
WINTER HDD	19	1118.510	92.135	1006.330	1304.000	1121.590
WINTER PRECIPITATION	19	11.387	0.620	9.970	12.393	11.390
SUMMER PRECIPITATION	19	11.560	0.686	10.505	12.858	11.592
ANNUAL CDD	19	462.363	117.378	235.500	678.000	456.858
ANNUAL HDD	19	6522.430	647.888	5653.000	7907.500	6493.590
TOTAL ANNUAL PRCP	19	47.598	1.836	43.980	50.820	47.340
DAYS WITH AVG TEMPERATURE <35	19	84.963	14.681	61.364	104.800	85.074
DAYS WITH AVG TEMPERATURE > 75	19	19.624	7.030	9.000	34.027	19.546
TRANSPORTATION	19	36.807	34.221	1.130	94.610	34.840
EDUCATION	19	52.434	34.699	6.510	99.710	47.590
ARTS	19	55.932	28.933	0.850	98.590	58.930
HEALTHCARE	19	43.278	28.101	0.000	94.900	42.770
RECREATION	19	55.133	23.272	7.360	94.610	58.350
TOTAL CRIME RATE	19	0.029	0.007	0.013	0.044	0.029
MSA OUT OF ATTAINMENT WITH NAAQS	19	0.526	0.513	0.000	1.000	1.000
MEAN PM 2.5	19	11.736	1.178	9.911	13.266	11.658
95TH PERCENTILE OF PM 2.5	19	24.510	3.253	20.063	30.209	23.324
MEAN PM 10	19	20.816	3.067	15.166	24.546	22.212
PM 10(95TH PERCENTILE)	19	39.551	5.190	29.623	46.016	41.277
POPULATION DENSITY PER SQUARE MILE	19	926.700	416.218	388.900	1754.900	817.900
POPULATION	19	565806.5	752923.9	142284.0	3406829.0	301686.0
MSA ON COAST	19	0.526	0.513	0	1	1
GREATLAKES	19	0	0	0	0	0
PACIFIC	19	0	0	0	0	0
ATLANTICGULF	19	0.526	0.513	0	1	1
NORTHEAST	19	1	0	1	1	1
MIDWEST	19	0	0	0	0	0
WEST	19	0	0	0	0	0
SOUTH	19	0	0	0	0	0

MIDDLE ATLANTIC DIVISION						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	34	28.751	3.901	19.933	34.495	29.267
SUMMER TEMPERATURE	34	69.552	2.515	64.633	73.000	69.400
SUMMER CDD	34	169.132	56.964	74.667	252.000	162.833
WINTER HDD	34	1087.510	116.789	915.038	1351.500	1072.330
WINTER PRECIPITATION	34	9.480	0.981	7.672	11.290	9.312
SUMMER PRECIPITATION	34	12.231	0.807	10.650	13.450	12.299
ANNUAL CDD	34	594.401	212.259	249.000	907.000	583.000
ANNUAL HDD	34	6228.130	892.336	4972.010	8261.500	6127.500
TOTAL ANNUAL PRCP	34	44.397	2.921	38.528	49.820	44.710
DAYS WITH AVG TEMPERATURE <35	34	79.823	19.834	45.329	127.853	81.732
DAYS WITH AVG TEMPERATURE > 75	34	28.465	14.673	4.405	58.457	24.437
TRANSPORTATION	33	50.300	30.396	6.230	99.710	48.440
EDUCATION	33	56.592	30.672	1.690	99.430	60.050
ARTS	33	56.302	31.049	0.570	100.000	58.080
HEALTHCARE	33	51.613	29.070	3.110	94.050	52.690
RECREATION	33	57.382	27.413	0.000	98.300	60.900
TOTAL CRIME RATE	34	0.028	0.009	0.004	0.055	0.027
MSA OUT OF ATTAINMENT WITH NAAQS	34	0.176	0.387	0.000	1.000	0.000
MEAN PM 2.5	34	13.563	1.728	9.593	17.509	13.724
95TH PERCENTILE OF PM 2.5	34	28.582	3.516	21.055	36.638	28.221
MEAN PM 10	34	22.363	5.298	11.582	32.256	23.336
PM 10(95TH PERCENTILE)	34	44.440	10.910	24.959	65.517	44.794
POPULATION DENSITY PER SQUARE MILE	34	1197.280	2552.820	72.900	13043.600	349.850
POPULATION	34	1064349.0	1765788.3	120044.0	9314235.0	429163.5
MSA ON COAST	34	0.382	0.493	0	1	0
GREATLAKES	34	0.118	0.327	0	1	0
PACIFIC	34	0	0	0	0	0
ATLANTICGULF	34	0.265	0.448	0	1	0
NORTHEAST	34	1	0	1	1	1
MIDWEST	34	0	0	0	0	0
WEST	34	0	0	0	0	0
SOUTH	34	0	0	0	0	0

EAST NORTH CENTRAL DIVISION						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	52	26.151	4.062	14.133	35.163	26.550
SUMMER TEMPERATURE	52	70.618	2.253	64.133	75.299	70.895
SUMMER CDD	52	201.646	52.972	79.667	322.124	203.599
WINTER HDD	52	1166.680	121.739	897.263	1527.670	1155.000
WINTER PRECIPITATION	52	6.437	1.703	3.059	11.876	6.365
SUMMER PRECIPITATION	52	11.703	0.874	9.490	13.115	12.028
ANNUAL CDD	52	746.247	225.088	266.000	1279.820	752.098
ANNUAL HDD	52	6483.640	874.064	4655.820	9209.000	6329.500
TOTAL ANNUAL PRCP	52	37.481	4.073	31.010	49.657	37.390
DAYS WITH AVG TEMPERATURE <35	51	89.660	19.060	45.357	131.045	90.172
DAYS WITH AVG TEMPERATURE > 75	51	37.294	15.870	16.322	78.409	34.250
TRANSPORTATION	52	54.152	30.312	5.940	100.000	59.625
EDUCATION	52	58.946	26.795	5.380	99.150	61.890
ARTS	52	57.953	25.753	3.970	99.160	56.805
HEALTHCARE	52	44.525	29.038	1.980	94.330	38.810
RECREATION	52	55.618	26.763	3.680	99.710	55.805
TOTAL CRIME RATE	46	0.038	0.010	0.018	0.059	0.036
MSA OUT OF ATTAINMENT WITH NAAQS	52	0.135	0.345	0.000	1.000	0.000
MEAN PM 2.5	52	14.575	1.942	10.007	17.669	14.890
95TH PERCENTILE OF PM 2.5	52	29.388	2.683	23.389	34.932	29.013
MEAN PM 10	52	23.241	3.131	17.591	30.469	22.923
PM 10(95TH PERCENTILE)	52	43.766	7.005	32.817	61.043	43.279
POPULATION DENSITY PER SQUARE MILE	52	398.096	291.148	81.400	1634.200	293.650
POPULATION	52	746424.9	1319023.2	101541.0	8272768.0	352876.0
MSA ON COAST	52	0.269	0.448	0	1	0
GREATLAKES	52	0.269	0.448	0	1	0
PACIFIC	52	0	0	0	0	0
ATLANTICGULF	52	0	0	0	0	0
NORTHEAST	52	0	0	0	0	0
MIDWEST	52	1	0	1	1	1
WEST	52	0	0	0	0	0
SOUTH	52	0	0	0	0	0

WEST NORTH CENTRAL DIVISION						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	21	23.025	7.668	9.442	34.667	22.700
SUMMER TEMPERATURE	21	71.934	3.750	62.739	78.090	72.000
SUMMER CDD	21	245.341	88.729	59.759	408.636	243.454
WINTER HDD	21	1262.120	230.102	911.667	1670.550	1270.670
WINTER PRECIPITATION	21	3.583	1.680	1.540	7.500	3.062
SUMMER PRECIPITATION	21	12.117	1.230	9.430	13.645	12.320
ANNUAL CDD	21	915.371	366.088	192.714	1592.890	883.000
ANNUAL HDD	21	6883.700	1563.890	4693.000	9863.630	6811.000
TOTAL ANNUAL PRCP	21	33.581	5.925	21.253	44.980	33.350
DAYS WITH AVG TEMPERATURE <35	21	95.932	30.967	44.760	146.629	99.452
DAYS WITH AVG TEMPERATURE > 75	21	48.219	23.811	4.014	87.620	45.143
TRANSPORTATION	21	55.924	21.619	16.430	97.450	52.690
EDUCATION	21	63.464	24.297	23.220	96.310	59.200
ARTS	21	58.416	21.929	4.820	98.020	56.380
HEALTHCARE	21	65.893	27.478	9.340	97.730	75.920
RECREATION	21	45.739	26.269	7.080	96.030	41.640
TOTAL CRIME RATE	21	0.042	0.015	0.018	0.085	0.041
MSA OUT OF ATTAINMENT WITH NAAQS	21	0.143	0.359	0.000	1.000	0.000
MEAN PM 2.5	21	10.803	1.365	7.601	13.165	10.939
95TH PERCENTILE OF PM 2.5	21	23.846	2.734	18.141	29.604	23.618
MEAN PM 10	21	25.120	4.846	17.608	35.748	25.133
PM 10(95TH PERCENTILE)	21	51.007	11.217	36.383	84.209	49.940
POPULATION DENSITY PER SQUARE MILE	21	199.838	107.253	32.400	489.700	183.800
POPULATION	21	436581.3	689058.7	102490.0	2968806.0	172412.0
MSA ON COAST	21	0.048	0.218	0	1	0
GREATLAKES	21	0.048	0.218	0	1	0
PACIFIC	21	0	0	0	0	0
ATLANTICGULF	21	0	0	0	0	0
NORTHEAST	21	0	0	0	0	0
MIDWEST	21	1	0	1	1	1
WEST	21	0	0	0	0	0
SOUTH	21	0	0	0	0	0

SOUTH ATLANTIC DIVISION						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	55	48.703	10.413	31.533	67.922	46.533
SUMMER TEMPERATURE	55	77.959	3.216	70.833	82.456	78.700
SUMMER CDD	55	399.843	95.924	194.333	536.111	421.000
WINTER HDD	55	521.769	271.766	68.333	1004.000	561.667
WINTER PRECIPITATION	55	10.760	2.664	5.920	15.409	10.894
SUMMER PRECIPITATION	55	16.600	4.119	10.610	23.300	15.870
ANNUAL CDD	55	2102.360	989.111	706.000	4171.000	1876.000
ANNUAL HDD	55	2557.760	1532.890	240.667	5588.000	2628.000
TOTAL ANNUAL PRCP	55	50.907	5.015	40.060	63.130	51.260
DAYS WITH AVG TEMPERATURE <35	54	15.359	16.543	0.000	60.789	10.129
DAYS WITH AVG TEMPERATURE > 75	54	118.816	52.103	30.245	237.273	107.869
TRANSPORTATION	55	50.935	29.030	0.840	98.300	53.250
EDUCATION	55	47.237	28.945	0.280	100.000	48.150
ARTS	55	46.950	29.989	2.270	99.720	47.030
HEALTHCARE	55	47.459	31.224	0.840	98.300	50.140
RECREATION	55	61.437	30.013	0.280	98.580	69.400
TOTAL CRIME RATE	55	0.050	0.013	0.023	0.083	0.050
MSA OUT OF ATTAINMENT WITH NAAQS	55	0.000	0.000	0.000	0.000	0.000
MEAN PM 2.5	55	14.056	2.546	9.418	19.178	14.848
95TH PERCENTILE OF PM 2.5	55	27.045	5.188	17.621	37.749	27.598
MEAN PM 10	55	22.727	3.083	16.897	30.491	23.279
PM 10(95TH PERCENTILE)	55	40.517	6.882	27.874	58.973	40.576
POPULATION DENSITY PER SQUARE MILE	55	359.776	267.417	108.700	1346.500	255.100
POPULATION	55	738132.1	963164.3	104646.0	4923153.0	322549.0
MSA ON COAST	55	0.455	0.503	0	1	0
GREATLAKES	55	0.000	0.000	0	0	0
PACIFIC	55	0	0	0	0	0
ATLANTICGULF	55	0.455	0.503	0	1	0
NORTHEAST	55	0	0	0	0	0
MIDWEST	55	0	0	0	0	0
WEST	55	0	0	0	0	0
SOUTH	55	1	0	1	1	1

EAST SOUTH CENTRAL DIVISION						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	20	43.600	4.916	34.300	52.433	43.658
SUMMER TEMPERATURE	20	77.505	2.088	73.281	80.700	77.600
SUMMER CDD	20	386.291	61.718	262.470	482.000	388.167
WINTER HDD	20	650.226	141.529	401.000	921.667	646.552
WINTER PRECIPITATION	20	15.283	1.666	10.530	17.212	15.897
SUMMER PRECIPITATION	20	13.371	1.915	12.060	19.470	12.658
ANNUAL CDD	20	1722.380	403.109	1039.920	2486.000	1678.000
ANNUAL HDD	20	3194.640	854.827	1732.000	4875.000	3166.510
TOTAL ANNUAL PRCP	20	56.728	4.763	46.150	66.590	56.711
DAYS WITH AVG TEMPERATURE <35	20	19.723	13.851	2.512	49.578	17.592
DAYS WITH AVG TEMPERATURE > 75	20	100.499	25.966	45.403	140.173	96.914
TRANSPORTATION	19	40.013	27.317	2.260	88.950	36.820
EDUCATION	19	53.447	28.811	1.130	93.480	60.620
ARTS	19	40.918	26.164	2.840	85.270	42.210
HEALTHCARE	19	57.473	29.197	2.540	98.010	59.490
RECREATION	19	45.559	29.265	2.260	92.060	40.220
TOTAL CRIME RATE	19	0.046	0.019	0.002	0.089	0.047
MSA OUT OF ATTAINMENT WITH NAAQS	20	0.000	0.000	0.000	0.000	0.000
MEAN PM 2.5	20	16.321	1.613	13.421	19.535	16.342
95TH PERCENTILE OF PM 2.5	20	32.240	4.343	25.958	42.941	31.451
MEAN PM 10	20	25.374	3.593	19.748	36.560	24.870
PM 10(95TH PERCENTILE)	20	47.338	8.550	34.861	73.718	46.052
POPULATION DENSITY PER SQUARE MILE	20	186.645	60.557	113.100	302.300	185.600
POPULATION	20	358396.3	304888.5	103459.0	1231311.0	270044.0
MSA ON COAST	20	0.100	0.308	0	1	0
GREATLAKES	20	0	0	0	0	0
PACIFIC	20	0	0	0	0	0
ATLANTICGULF	20	0.100	0.308	0	1	0
NORTHEAST	20	0	0	0	0	0
MIDWEST	20	0	0	0	0	0
WEST	20	0	0	0	0	0
SOUTH	20	1.000	0	1	1	1

WEST SOUTH CENTRAL DIVISION						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	37	48.459	6.425	37.200	61.100	48.300
SUMMER TEMPERATURE	37	81.017	2.069	75.900	84.800	81.267
SUMMER CDD	37	492.293	62.350	342.333	608.333	499.667
WINTER HDD	37	512.841	178.168	194.000	837.333	510.000
WINTER PRECIPITATION	37	9.011	4.875	1.520	17.040	8.840
SUMMER PRECIPITATION	37	10.816	3.525	5.520	19.130	10.340
ANNUAL CDD	37	2439.480	621.364	1365.000	3943.000	2408.000
ANNUAL HDD	37	2328.350	974.487	714.000	4262.000	2259.000
TOTAL ANNUAL PRCP	37	42.092	14.849	13.190	64.270	47.270
DAYS WITH AVG TEMPERATURE <35	37	10.702	10.414	0.343	38.160	8.728
DAYS WITH AVG TEMPERATURE > 75	37	135.453	30.138	78.575	210.086	133.829
TRANSPORTATION	37	51.377	26.058	0.000	95.460	49.290
EDUCATION	37	41.148	26.450	3.680	98.300	39.370
ARTS	37	36.296	28.034	0.000	93.210	26.070
HEALTHCARE	37	48.383	25.208	2.260	92.350	49.290
RECREATION	37	46.339	26.546	0.560	100.000	45.600
TOTAL CRIME RATE	37	0.055	0.011	0.028	0.072	0.055
MSA OUT OF ATTAINMENT WITH NAAQS	37	0.054	0.229	0.000	1.000	0.000
MEAN PM 2.5	37	11.691	1.980	7.325	16.011	11.976
95TH PERCENTILE OF PM 2.5	37	21.942	3.971	13.757	29.886	22.458
MEAN PM 10	37	25.065	2.835	17.773	33.188	24.874
PM 10(95TH PERCENTILE)	37	46.546	5.950	32.974	60.163	46.963
POPULATION DENSITY PER SQUARE MILE	37	269.176	176.548	57.500	705.700	200.800
POPULATION	37	670230.9	884375.6	126337.0	4177646.0	312952.0
MSA ON COAST	37	0.216	0.417	0	1	0
GREATLAKES	37	0	0.000	0	0	0
PACIFIC	37	0	0.000	0	0	0
ATLANTICGULF	37	0.216	0.417	0	1	0
NORTHEAST	37	0	0	0	0	0
MIDWEST	37	0	0	0	0	0
WEST	37	0	0	0	0	0
SOUTH	37	1.000	0	1	1	1

MOUNTAIN DIVISION						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	20	33.132	9.414	22.967	55.033	30.033
SUMMER TEMPERATURE	20	70.323	7.631	63.233	89.733	67.267
SUMMER CDD	20	214.682	204.795	44.667	760.000	129.833
WINTER HDD	20	959.852	279.790	315.000	1264.330	1052.500
WINTER PRECIPITATION	20	2.750	1.215	1.500	5.070	2.233
SUMMER PRECIPITATION	20	4.472	1.921	1.330	7.242	4.960
ANNUAL CDD	20	887.381	1058.500	140.000	3892.000	439.000
ANNUAL HDD	20	5879.180	2075.520	1343.000	8089.000	6495.260
TOTAL ANNUAL PRCP	20	14.424	3.752	5.080	19.140	15.695
DAYS WITH AVG TEMPERATURE <35	20	58.330	33.079	0.000	100.800	64.943
DAYS WITH AVG TEMPERATURE > 75	20	49.580	50.168	0.686	164.765	28.819
TRANSPORTATION	20	54.486	31.797	11.040	98.580	59.485
EDUCATION	20	47.432	27.685	2.260	92.060	46.880
ARTS	20	53.249	26.002	3.690	97.740	52.840
HEALTHCARE	20	45.533	28.485	0.280	85.260	53.960
RECREATION	20	47.077	31.586	8.490	96.310	46.170
TOTAL CRIME RATE	20	0.047	0.011	0.032	0.071	0.043
MSA OUT OF ATTAINMENT WITH NAAQS	20	0.650	0.489	0.000	1.000	1.000
MEAN PM 2.5	20	8.164	1.651	5.382	11.516	8.013
95TH PERCENTILE OF PM 2.5	20	18.789	6.434	9.389	35.986	17.419
MEAN PM 10	20	22.651	7.931	10.930	44.384	21.084
PM 10(95TH PERCENTILE)	20	45.251	15.059	19.124	81.090	41.172
POPULATION DENSITY PER SQUARE MILE	20	171.745	208.352	5.400	824.700	82.450
POPULATION	20	659382.0	820794.0	116255.0	3251876.0	315387.0
MSA ON COAST	20	0	0	0	0	0
GREATLAKES	20	0	0	0	0	0
PACIFIC	20	0	0	0	0	0
ATLANTICGULF	20	0	0	0	0	0
NORTHEAST	20	0	0	0	0	0
MIDWEST	20	0	0	0	0	0
WEST	20	1	0	1	1	1
SOUTH	20	0	0	0	0	0

PACIFIC DIVISION						
Variable	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**Median**
WINTER TEMPERATURE	37	43.600	5.998	29.300	53.100	44.733
SUMMER TEMPERATURE	37	68.805	5.159	60.848	78.583	69.783
SUMMER CDD	37	171.234	115.062	32.005	422.833	176.000
WINTER HDD	37	645.006	179.728	361.333	1074.670	610.667
WINTER PRECIPITATION	37	14.569	7.044	3.820	28.084	11.660
SUMMER PRECIPITATION	37	2.299	1.581	0.440	5.853	1.780
ANNUAL CDD	37	728.596	533.145	111.783	1909.500	705.000
ANNUAL HDD	37	4027.310	1404.070	2133.000	6775.670	3892.000
TOTAL ANNUAL PRCP	37	31.484	17.925	10.150	66.747	21.400
DAYS WITH AVG TEMPERATURE <35	37	11.465	19.257	0.000	84.829	1.514
DAYS WITH AVG TEMPERATURE > 75	37	31.183	31.390	0.600	121.329	19.629
TRANSPORTATION	39	50.064	29.862	1.410	93.480	50.700
EDUCATION	39	43.671	30.689	0.000	98.010	38.520
ARTS	39	54.425	30.970	5.670	99.440	54.680
HEALTHCARE	39	42.451	25.779	5.090	95.460	40.220
RECREATION	39	46.628	31.590	4.810	96.880	41.350
TOTAL CRIME RATE	39	0.043	0.011	0.023	0.067	0.042
MSA OUT OF ATTAINMENT WITH NAAQS	39	0.462	0.505	0.000	1.000	0.000
MEAN PM 2.5	37	12.033	3.228	7.311	18.948	10.847
95TH PERCENTILE OF PM 2.5	37	32.355	10.150	15.593	58.177	28.798
MEAN PM 10	37	24.420	6.588	15.598	38.336	21.755
PM 10(95TH PERCENTILE)	37	48.954	14.315	30.445	96.148	42.799
POPULATION DENSITY PER SQUARE MILE	39	508.682	738.308	43.100	3605.600	237.000
POPULATION	39	1056671.790	1646073.060	139149.000	9519338.000	417939.000
MSA ON COAST	39	0.513	0.506	0	1	1
GREATLAKES	39	0	0	0	0	0
PACIFIC	39	0.513	0.506	0	1	1
ATLANTICGULF	39	0	0	0	0	0
NORTHEAST	39	0	0	0	0	0
MIDWEST	39	0	0	0	0	0
WEST	39	1	0	1	1	1
SOUTH	39	0	0	0	0	0

Appendix A5[73]

Table A5.1 Summary of Hedonic Wage Coefficients

Variables (Dependent Variable: log(wagerate))	Mean of Estimates from 297 MSAs	Std Dev of Estimates from 297 MSAs
Highschool (left out category is no high school)	0.101	0.040
Somecollege	0.181	0.047
Collgrad	0.387	0.070
Highereduc	0.553	0.076
Age	0.051	0.008
Age squared (divided by 100)	-0.049	0.009
Married	0.095	0.022
Male	0.213	0.040
Black (left out category is white)	-0.067	0.075
Other Race	-0.054	0.058
Speaks English Well	0.111	0.117
Hispanic	-0.043	0.080
Businessoperations_Occ (left out category is Management_Occ)	-0.125	0.067
Financialspecialists_Occ	-0.114	0.078
Computerandmath_Occ	-0.002	0.090
Engineering_Occ	-0.074	0.084
Lifephysicalsocialsc_Occ	-0.183	0.112
Socialservices_Occ	-0.345	0.085
Legal_Occ	-0.040	0.137
Teachers_Occ	-0.200	0.091
Othereduc_Occ	-0.486	0.134
Artssportsmedia_Occ	-0.253	0.098
Healthcarepractitioners_Occ	0.074	0.077
Healthcaresupport_Occ	-0.323	0.081
Protectiveservices_Occ	-0.237	0.106
Foodandserving_Occ	-0.419	0.076

[73] Appendix A.5 corresponds to Chapter 5.

Maintenance_Occ	-0.466	0.079
Personalcareservice_Occ	-0.413	0.112
Highskillsales_Occ	-0.135	0.068
Lowskillsales_Occ	-0.228	0.064
Officesupport_Occ	-0.298	0.052
Constructiontrades_Occ	-0.239	0.094
Extractionworkers_Occ	-0.261	0.292
Maintenanceworkers_Occ	-0.185	0.067
Production_Occ	-0.310	0.085
Transportation_Occ	-0.356	0.074
Construction_Ind (left out category is Mining And Utilities)[74]	-0.178	0.098
Manufacturing_Ind	-0.118	0.108
Wholesale_Ind	-0.185	0.099
Retail_Ind	-0.342	0.098
Transportation_Ind	-0.093	0.110
Informationcomm_Ind	-0.139	0.114
Finance_Ind	-0.173	0.107
Profscientificmngmntservices_Ind	-0.223	0.106
Educhealthsocialservices_Ind	-0.274	0.096
Recreationfoodservices_Ind	-0.378	0.114
Otherservices_Ind	-0.361	0.101
Publicad_Ind	-0.131	0.100

[74] Since these two industries have a very low number of observations, we bundled them together as the left out category.

Table A5.2: Coefficients of the Hedonic Housing Equation

Dependent Variable : Log(user costs including insurance and utility costs) Number of Observations Used: 3346588 Adjusted R-Sq: 0.5737		
Variables	**Coefficient**	**t-statistic**
Intercept	5.625	499.41
Own (=1 if the house is owned)	0.505	633.95
Bedroom3 (left out category is less than three bedrooms)	0.129	100.49
Bedroom4	0.154	99.42
Bedroom5	0.284	162.02
Bedroomgt5	0.486	225.45
Room2 (left out category is less than two rooms)	0.139	69.27
Room3	0.140	73.67
Room4	0.169	79.76
Room5	0.233	103.99
Room6	0.329	141.05
Roomgt6	0.533	224.08
Completekitchen	-0.035	-9.65
Completeplumbing	0.218	55.93
Acres1to10	-0.214	-97.46
Ageofstructure_0to1years (left out category is age of structure over 61 years)	0.390	192.86
Ageofstructure_2to5years	0.369	292.34
Ageofstructure_6to10years	0.314	255.44
Ageofstructure_11to20years	0.216	216.27
Ageofstructure_21to30years	0.108	113.21
Ageofstructure_31to40years	0.058	59.24
Ageofstructure_41to50years	0.020	20.83
Ageofstructure_51to60years	-0.025	-22.00
Unitsinstructure _Singleattached (left out category is units in	-0.157	-139.84

structure single family detached)		
Units_In_Structure _2	-0.270	-106.24
Units_In_Structure _3to4	-0.326	-127.97
Units_In_Structure _5to9	-0.353	-137.60
Units_In_Structure _10to19	-0.330	-125.94
Units_In_Structure _20to49	-0.382	-142.59
Units_In_Structure _Over50	-0.367	-143.29
Akron, OH (left out MSA is Abilene, TX	0.309	27.86
Albany, GA	0.049	3.2
Albany-Schenectady-Troy, NY	0.377	34.53
Albuquerque, NM	0.347	31.07
Alexandria, LA	-0.064	-4.33
Allentown-Bethlehem-Easton, PA/NJ	0.364	32.24
Altoona, PA	0.001	0.09
Amarillo, TX	0.154	11.87
Anchorage, AK	0.614	46.01
Ann Arbor, MI	0.629	53.61
Anniston, AL	-0.112	-7.28
Appleton-Oskosh-Neenah, WI	0.278	23.08
Asheville, NC	0.321	24.94
Athens, GA	0.226	16.37
Atlanta, GA	0.400	38.52
Atlantic City, NJ	0.535	45.67
Auburn-Opelika, AL	0.111	6.81
Augusta-Aiken, GA-SC	0.066	5.58
Austin, TX	0.527	48.85
Bakersfield, CA	0.235	20.63
Baltimore, MD	0.436	41.75
Barnstable-Yarmouth, MA	0.656	49.09
Baton Rouge, LA	0.145	12.65
Beaumont-Port Arthur-Orange,TX	-0.011	-0.91
Bellingham, WA	0.491	35.65
Benton Harbor, MI	0.205	15.39
Billings, MT	0.167	10.49
Biloxi-Gulfport, MS	0.134	10.57
Binghamton, NY	0.142	11.67
Birmingham, AL	0.177	16.07
Bloomington, IN	0.315	20.55
Bloomington-Normal, IL	0.251	18.76
Boise City, ID	0.260	21.35

Boston, MA	0.883	85.14
Lawrence-Haverhill, MA/NH	0.643	52
Lowell, MA/NH	0.650	53.18
Bremerton, WA	0.509	39.64
Bridgeport, CT	0.793	66.95
Brockton, MA	0.593	47.48
Brownsville-Harlingen-San Benito, TX	-0.111	-8.79
Bryan-College Station, TX	0.282	18.9
Buffalo-Niagara Falls, NY	0.263	24.51
Canton, OH	0.165	14.24
Cedar Rapids, IA	0.251	18.19
Champaign-Urbana-Rantoul, IL	0.246	18.75
Charleston-N.Charleston,SC	0.390	33.15
Charlotte-Gastonia-Rock Hill, SC	0.324	30.34
Charlottesville, VA	0.382	28.12
Chattanooga, TN/GA	0.146	12.54
Chicago-Gary-Lake, IL	0.679	66.05
Gary-Hammond-East Chicago, IN	0.287	25.53
Chico, CA	0.392	30.05
Cincinnati OH/KY/IN	0.320	30.11
Clarksville-Hopkinsville, TN/KY	0.092	6.18
Cleveland, OH	0.369	35.31
Colorado Springs, CO	0.394	34.35
Columbia, MO	0.167	11.14
Columbia, SC	0.211	18.37
Columbus, GA/AL	0.100	7.55
Columbus, OH	0.339	31.89
Corpus Christi, TX	0.153	12.12
Dallas-Fort Worth, TX	0.425	40.92
Fort Worth-Arlington, TX	0.295	27.81
Danbury, CT	0.927	70.08
Danville, VA	-0.055	-3.79
Davenport, IA Rock Island-Moline, IL	0.185	14.8
Dayton-Springfield, OH	0.251	23.13
Daytona Beach, FL	0.206	17.94
Decatur, AL	-0.003	-0.24
Decatur, IL	0.049	3.6
Denver-Boulder-Longmont, CO	0.582	55.36
Boulder-Longmont, CO	0.777	60.02
Des Moines, IA	0.314	25.3
Detroit, MI	0.448	43.26

Dothan, AL	-0.102	-7.34
Dover, DE	0.198	13.53
Duluth-Superior, MN/WI	0.018	1.43
Dutchess Co., NY	0.600	48.66
Eau Claire, WI	0.132	9.36
El Paso, TX	0.024	2.07
Elkhart-Goshen, IN	0.188	13.6
Erie, PA	0.118	9.53
Eugene-Springfield, OR	0.467	38.18
Evansville, IN/KY	0.135	10.84
Fargo-Morehead, ND/MN	0.151	9.52
Fayetteville, NC	0.160	12.39
Fayetteville-Springdale, AR	0.108	8.91
Fitchburg-Leominster, MA	0.413	30.18
Flagstaff, AZ-UT	0.307	20.28
Flint, MI	0.057	4.38
Florence, AL	-0.036	-2.6
Fort Collins-Loveland, CO	0.493	38.11
Fort Lauderdale-Hollywood-Pompano Beach, FL	0.477	45.2
Fort Myers-Cape Coral, FL	0.356	31.12
Fort Pierce, FL	0.294	24.67
Fort Smith, AR/OK	-0.018	-1.29
Fort Walton Beach, FL	0.229	17.22
Fort Wayne, IN	0.088	7.68
Fresno, CA	0.316	28.65
Gadsden, AL	-0.117	-7.59
Gainesville, FL	0.220	17.34
Galveston-Texas City, TX	0.286	22.91
Glens Falls, NY	0.241	17.85
Goldsboro, NC	0.051	3.08
Grand Rapids, MI	0.279	25.26
Grand Junction, CO	0.291	18.67
Greeley, CO	0.343	25.94
Green Bay, WI	0.352	26.7
Greensboro-Winston Salem-High Point, NC	0.234	21.83
Greenville, NC	0.130	9.22
Greenville-Spartanburg-Anderson SC	0.142	12.78
Hagerstown, MD	0.232	16.94
Hamilton-Middleton, OH	0.294	24.08
Harrisburg-Lebanon-Carlisle, PA	0.270	24.12

Hartford-Bristol-Middleton-New Britain, CT	0.586	53.49
Hickory-Morgantown, NC	0.115	9.49
Hattiesburg, MS	-0.022	-1.38
Honolulu, HI	0.920	83.14
Houma-Thibodoux, LA	-0.002	-0.11
Houston-Brazoria, TX	0.319	30.72
Brazoria, TX	0.156	12.08
Huntsville, AL	0.079	6.53
Indianapolis, IN	0.274	25.71
Iowa City, IA	0.381	23.87
Jackson, MI	0.172	12.11
Jackson, MS	0.079	6.61
Jackson, TN	0.000	-0.01
Jacksonville, FL	0.244	22.63
Jacksonville, NC	0.134	8.48
Jamestown-Dunkirk, NY	0.029	2.17
Janesville-Beloit, WI	0.285	19.98
Johnson City-Kingsport-Bristol, TN/VA	0.031	2.56
Johnstown, PA	-0.126	-10.49
Joplin, MO	-0.090	-6.58
Kalamazoo-Portage, MI	0.184	15.93
Kankakee, IL	0.287	19.49
Kansas City, MO-KS	0.243	22.9
Kenosha, WI	0.455	30.49
Kileen-Temple, TX	0.120	9.53
Knoxville, TN	0.131	11.62
Kokomo, IN	0.126	8.38
LaCrosse, WI	0.228	13.73
Lafayette, LA	0.043	3.31
Lafayette-W. Lafayette, IN	0.256	19.14
Lake Charles, LA	0.012	0.9
Lakeland-Winterhaven, FL	0.098	8.37
Lancaster, PA	0.361	30.64
Lansing-E. Lansing, MI	0.281	24.38
Laredo, TX	-0.019	-1.28
Las Cruces, NM	0.073	4.79
Las Vegas, NV	0.427	40.07
Lexington-Fayette, KY	0.250	19.77

Lima, OH	0.041	3.09
Lincoln, NE	0.284	21.25
Little Rock-North Little Rock, AR	0.140	12.06
Longview-Marshall, TX	0.016	1.17
Los Angeles-Long Beach, CA	0.870	84.76
Orange County, CA	0.969	92.9
Louisville, KY/IN	0.258	23.93
Lubbock, TX	0.110	8.71
Lynchburg, VA	0.061	4.68
Macon-Warner Robins, GA	0.045	3.76
Madison, WI	0.544	45.44
Manchester, NH	0.543	34.65
Mansfield, OH	0.094	6.53
McAllen-Edinburg-Pharr-Mission, TX	-0.219	-18.16
Medford, OR	0.439	32.04
Melbourne-Titusville-Cocoa-Palm Bay, FL	0.210	18.55
Memphis, TN/AR/MS	0.210	19.25
Merced, CA	0.302	22.11
Miami-Hialeah, FL	0.567	54.03
Milwaukee, WI	0.483	44.95
Minneapolis-St. Paul, MN	0.468	44.67
Mobile, AL	0.077	6.65
Modesto, CA	0.423	36.05
Monmouth-Ocean, NJ	0.702	65.48
Monroe, LA	-0.017	-1.15
Montgomery, AL	0.099	7.99
Muncie, IN	0.049	3.35
Myrtle Beach, SC	0.226	16.34
Naples, FL	0.683	54.79
Nashua, NH	0.585	36.6
Nashville, TN	0.321	29.9
New Bedford, MA	0.417	32.1
New Haven-Meriden, CT	0.618	52.24
New Orleans, LA	0.259	24.09
New York-Northeastern NJ	0.883	85.95
Nassau Co, NY	1.010	96.84
Bergen-Passaic, NJ	0.990	92.95
Jersey City, NJ	0.787	70.43
Middlesex-Somerset-Hunterdon, NJ	0.823	76.93
Newark, NJ	0.830	79.04

Newburgh-Middletown, NY	0.593	49.33
Norfolk-VA Beach-Newport News, VA	0.301	28.36
Ocala, FL	0.046	3.59
Odessa, TX	0.008	0.64
Oklahoma City, OK	0.113	10.17
Olympia, WA	0.465	35.08
Omaha, NE/IA	0.257	21.87
Orlando, FL	0.336	31.7
Panama City, FL	0.179	12.92
Pensacola, FL	0.107	9.13
Peoria, IL	0.176	15.15
Philadelphia, PA/NJ	0.501	48.46
Phoenix, AZ	0.417	40.04
Pittsburgh-Beaver Valley, PA	0.201	19.22
Portland, ME	0.477	35.06
Portland-Vancouver, OR	0.606	57.38
Providence-Fall River-Pawtucket, MA/RI	0.448	41.65
Provo-Orem, UT	0.335	26.63
Pueblo, CO	0.130	9.2
Punta Gorda, FL	0.245	17.9
Racine, WI	0.398	27.69
Raleigh-Durham, NC	0.407	38.01
Reading, PA	0.305	25.84
Redding, CA	0.373	27.84
Reno, NV	0.566	47.55
Richland-Kennewick-Pasco, WA	0.271	19.63
Richmond-Petersburg, VA	0.276	25.59
Riverside-San Bernadino, CA	0.455	43.58
Roanoke, VA	0.148	11.9
Rochester, MN	0.237	15.3
Rochester, NY	0.331	30.77
Rockford, IL	0.242	20.12
Rocky Mount, NC	0.089	6.16
Sacramento, CA	0.558	52.75
Saginaw-Bay City-Midland, MI	0.137	11.75
St. Cloud, MN	0.104	7.73
St. Joseph, MO	0.004	0.29
St. Louis, MO-IL	0.251	24
Salem, OR	0.436	34.97
Salinas-Sea Side-Monterey, CA	0.986	78.9
Salt Lake City-Ogden, UT	0.401	36.97

San Antonio, TX	0.166	15.56
San Diego, CA	0.820	78.55
San Francisco-Oakland-Vallejo, CA	1.358	128.94
Oakland, CA	1.003	96.01
Vallejo-Fairfield-Napa, CA	0.719	62.5
San Jose, CA	1.339	126.42
San Luis Obispo-Atascad-P Robles, CA	0.791	61.7
Santa Barbara-Santa Maria-Lompoc, CA	0.991	84.73
Santa Cruz, CA	1.161	92.23
Santa Fe, NM	0.638	44.36
Santa Rosa-Petaluma, CA	0.955	82.83
Sarasota, FL	0.466	41.72
Savannah, GA	0.292	22.86
Scranton-Wilkes-Barre, PA	0.173	15.72
Seattle-Everett, WA	0.790	75.49
Sharon, PA	0.052	3.71
Sheboygan, WI	0.320	21.29
Shreveport, LA	-0.008	-0.69
Sioux City, IA/NE	0.086	5.35
Sioux Falls, SD	0.238	14.17
South Bend-Mishawaka, IN	0.131	10.6
Spokane, WA	0.274	23.47
Springfield, IL	0.197	13.3
Springfield, MO	0.053	4.34
Springfield-Holyoke-Chicopee, MA	0.409	36.55
Stamford, CT	1.286	108.61
State College, PA	0.275	19.54
Stockton, CA	0.485	42.46
Sumter, SC	-0.062	-3.64
Syracuse, NY	0.211	19.42
Tacoma, WA	0.505	45.64
Tallahassee, FL	0.250	20.38
Tampa-St. Petersburg-Clearwater, FL	0.291	27.86
Terre Haute, IN	-0.018	-1.33
Toledo, OH/MI	0.200	17.88
Topeka, KS	0.104	7.41
Trenton, NJ	0.679	56.58
Tucson, AZ	0.334	30.36
Tulsa, OK	0.146	12.84
Tuscaloosa, AL	0.132	8.94
Tyler, TX	0.126	9.28

Utica-Rome, NY	0.112	9.56
Ventura-Oxnard-Simi Valley, CA	0.890	80.03
Vineland-Milville-Bridgetown, NJ	0.307	21.75
Visalia-Tulare-Porterville, CA	0.255	21.38
Waco, TX	0.041	3.18
Washington, DC/MD/VA	0.681	65.97
Waterbury, CT	0.373	25.37
Waterloo-Cedar Falls, IA	0.112	7.26
Wausau, WI	0.163	11.5
West Palm Beach-Boca Raton-Delray Beach, FL	0.491	46.04
Wichita, KS	0.120	10.43
Wichita Falls, TX	0.038	2.58
Williamsport, PA	0.132	9.9
Wilmington, DE/NJ/MD	0.426	37.31
Wilmington, NC	0.366	29.18
Worcester, MA	0.493	40.8
Yakima, WA	0.295	21.66
Yolo, CA	0.637	46.48
York, PA	0.268	22.7
Youngstown-Warren, OH-PA	0.076	6.75
Yuba City, CA	0.289	20.22
Yuma, AZ	0.136	8.72

Table A5.3 Movers with Annual Degree Days

	With Census Regions		With Census Divisions	
	Model 1		Model 2	
Variable	Coefficient	t-statistic	Coefficient	t-statistic
INTERCEPT	-0.91540	-1.25	0.03297	0.05
MEAN PM10	0.01878	2.22	0.02091	2.44
TOTAL CRIME RATE	-5.33684	-1.88	-6.14108	-2.16
POP DENSITY PER SQ MILE OF LAND	0.00014	4.01	0.00015	4.1
TRANSPORTATION	0.00318	1.8	0.00299	1.7
EDUCATION	0.00632	3.53	0.00650	3.64
ARTS	0.00984	5.37	0.00925	5.08
HEALTHCARE	-0.00038	-0.27	0.00083	0.57
RECREATION	0.01405	7.35	0.01369	7.1
MSA ON THE COAST	-0.14632	-1.71	-0.15979	-1.87
ANNUAL HDD	0.00044	2.73	0.00037	2.22
ANNUAL HDD SQUARED	0.00000	-4.24	0.00000	-3.47
ANNUAL CDD	-0.00026	-0.97	-0.00001	-0.02
ANNUAL CDD SQUARED	0.00000	2.42	0.00000	1.29
WINTERPR	-0.01920	-0.65	-0.04131	-1.28
WINTERPR SQUARED	0.00023	0.21	0.00139	1.2
SUMMERPR	0.09128	1.84	0.08526	1.43
SUMMERPR SQUARED	-0.00144	-0.76	-0.00173	-0.79
NORTH EAST[75]	-0.18981	-1.26		
MID WEST	-0.33059	-2.38		
WEST	0.97100	3.77		
MIDDLE ATLANTIC[76]			-1.08314	-2.74
EAST NORTH CENTRAL			-1.09593	-2.76
WEST NORTH CENTRAL			-1.40008	-3.23
SOUTH ATLANTIC			-0.75551	-1.98
EAST SOUTH CENTRAL			-1.13022	-2.93
WEST SOUTH CENTRAL			-1.00093	-2.62
MOUNTAIN			0.08836	0.34
NEW ENGLAND			-0.77422	-2.02
Number of Observations	**286**		**286**	
Adjusted R-Squared	**0.7599**		**0.7679**	
R-Squared	**0.7768**		**0.7883**	

[75] The left out category is the SOUTH.
[76] The left out category includes the PACIFIC division

Table A5.4 Marginal Effects of Climate Variables in Annual Degree Days Specification

Annual Degree Days	Model 1 With Census Regions		Model 2 With Census Divisions	
	Coefficient	t-statistic	Coefficient	t-statistic
ANNUAL HDD	-0.00004	-0.48	-0.00003	-0.39
ANNUAL CDD	0.00005	0.25	0.00017	0.83
WINTERPR	-0.01478	-1.27	-0.01512	-1.08
SUMMERPR	0.05965	4.26	0.04716	2.46

Table A5.5 Second Stage Estimates Using "Bin" Data

RUNS USING THE "BIN" DATA	With Census Regions Model 1		With Census Divisions Model 2	
INTERCEPT	-6.17367	-1.48	-1.70053	-0.38
MEAN PM10	0.01499	1.88	0.01854	2.28
TOTAL CRIME RATE	-6.71816	-2.33	-6.78776	-2.36
POP DENSITY PER SQ MILE OF LAND	0.00013	3.74	0.00013	3.74
TRANSPORTATION	0.00268	1.49	0.00268	1.49
EDUCATION	0.00695	3.82	0.00705	3.86
ARTS	0.01010	5.48	0.00921	5.02
HEALTHCARE	0.00026	0.17	0.00115	0.76
RECREATION	0.01371	7.04	0.01386	7.1
MSA ON THE COAST	-0.12386	-1.42	-0.16004	-1.83
NUMBER OF DAYS WHEN MEAN TEMPERATURE IS BETWEEN 10 AND 25 DEGREES FAHRENHEIT[77]	0.00177	0.09	-0.01343	-0.64
NUMBER OF DAYS WHEN MEAN TEMPERATURE IS BETWEEN 25 AND 40 DEGREES FAHRENHEIT	0.02027	2.09	0.01084	1.01
NUMBER OF DAYS WHEN MEAN TEMPERATURE IS BETWEEN 40 AND 55 DEGREES FAHRENHEIT	0.01875	1.49	0.00725	0.55
NUMBER OF DAYS WHEN MEAN TEMPERATURE IS BETWEEN 55 AND 70DEGREES FAHRENHEIT	0.01490	1.32	0.00584	0.48
NUMBER OF DAYS WHEN MEAN TEMPERATURE IS BETWEEN 70 AND 85 DEGREES FAHRENHEIT	0.01774	1.56	0.00901	0.74
NUMBER OF DAYS WHEN MEAN TEMPERATURE IS BETWEEN 85 AND 95 DEGREES FAHRENHEIT	0.02373	2.03	0.01278	1.01
WINTERPR	-0.01157	-0.39	-0.04059	-1.29
WINTERPR SQUARED	-0.00001	-0.01	0.00118	1.03
SUMMERPR	0.05174	1.02	0.08300	1.35
SUMMERPR SQUARED	0.00005	0.03	-0.00187	-0.8
NORTH EAST[78]	-0.00655	-0.04		
MID WEST	-0.05463	-0.36		
WEST	0.85716	3.29		
MIDDLE ATLANTIC[79]			-0.90890	-2.21
EAST NORTH CENTRAL			-0.88261	-2.17
WEST NORTH CENTRAL			-1.20941	-2.82
SOUTH ATLANTIC			-0.80009	-2.05

[77] The left out category is the number of days below 10 degrees Fahrenheit.
[78] The left out category is the SOUTH.

[79] The left out category includes the PACIFIC division

EAST SOUTH CENTRAL			-1.13288	-2.89
WEST SOUTH CENTRAL			-1.06837	-2.74
MOUNTAIN			-0.08206	-0.3
NEW ENGLAND			-0.58133	-1.43
Adjusted R-Squared	0.7598		0.7673	
R-Squared	0.7785		0.7895	

Table A5.6 Marginal Effects Using Other Climate Specifications

Extreme Temperature	**With Census Divisions**	
	Coefficient	**t-statistic**
DAYS BELOW 35 DEGREES FAHRENHEIT	-0.00359	-1.13
DAYS ABOVE 75 DEGREES FAHRENHEIT	0.00242	1.01
WINTER PRECIPITATION	-0.015	-1.03
SUMMER PRECIPITATION	0.03312	1.71
Including Humidity (NUMBER OF OBSERVATIONS: 156)	**With Census Divisions**	
	Coefficient	**t-statistic**
WINTER TEMPERATURE	0.03541	2.55
SUMMER TEMPERATURE	-0.0113	-0.45
WINTER PRECIPITATION	0.00159	0.09
SUMMER PRECIPITATION	0.0646	2.54
Interacting Precipitation with Temperature	**With Census Divisions**	
	Coefficient	**t-statistic**
WINTER TEMPERATURE	0.01003	0.74
SUMMER TEMPERATURE	0.00908	0.34
WINTER PRECIPITATION	0.002	0.12
SUMMER PRECIPITATION	0.05206	2.31

Table A5.7 Marginal Effects of Climate Variables for Movers (with Birthplace Moving Costs)

Using Estimates from Specification 3	Model 1 With Census Regions		Model 2 With Census Divisions	
	Coefficient	t-statistic	Coefficient	t-statistic
WINTER TEMPERATURE	0.0411	4.08	0.0354	3.31
SUMMER TEMPERATURE	-0.0213	-1.16	0.0022	0.11
WINTER PRECIPITATION	-0.0261	-2.16	-0.0257	-1.72
SUMMER PRECIPITATION	0.0352	2.42	0.0253	1.25

Bibliography

Abt Associates Inc., BenMAP, *The Environmental Benefits Mapping and Analysis Program: User's Manual* (June, 2005) *Prepared for* Office of Air Quality Planning and Standards U.S. Environmental Protection Agency Research Triangle Park, NC

Bayer, Patrick; Keohane, Nathaniel and Timmins, Christopher. "Migration and Hedonic Valuation: The Case of Air Quality", *NBER Working Paper*, No. 12106 (March, 2006)

Berry, S., O. B. Linton and A. Pakes.. "Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems." *Review of Economic Studies* 71(3):613-54. (2004)

Blomquist, Glenn C.; Berger, Mark C.; Hoehn, John P. "New Estimates of Quality of Life in Urban Areas", *The American Economic Review,* Vol. 78, Issue 1, March (1988): 89-107.

Census 2000 PHC-T-3, Ranking Tables for Metropolitan Areas: 1990 and 2000; Table1: Metropolitan Areas and their Geographic Components in Alphabetic Sort, 1990 and 2000 Population, and Numeric and Percent Population Change: 1990 to 2000

Cragg, M. and Kahn, Matthew. "New Estimates of Climate Demand: Evidence from Location Choice." *Journal of Urban Economics ,* Vol. 42, (1997): 261-284.

Cropper, M.L. The value of urban amenities. *Journal of Regional Science,* 1981: **21(3),** 359-374.

Cropper, M.L. and Arriaga-Salinas A.S. "Inter-city Wage Differentials and the Value of Air Quality" *Journal of Urban Economics*, Vol. 8 (1980): 236-254.

Cushing, Brian J.,"A Note on the Specification of Climate Variables in Models of Population Migration"

Follain, J.R. and S. Malpezzi. "Are Occupants Accurate Appraisers?" Review of Public Data Use (1981) 9: 47-55.

Freddie Mac, Primary Mortgage Market Survey, Available Online at http://www.freddiemac.com/dlink/html/PMMS/display/PMMSOutputYr.jsp?year=2008

Freeman, Myrick A. "The Measurement of Environmental and Resource Values: Theory and Methods" (1993) Resources For the Future

Goodman, J.L. Jr. and J.B. Ittner. "The Accuracy of Home Owners' Estimates of House Value." Journal of Housing Economics (1992) 2: 339-357.

Graves and Mueser, "The Role of Equilibrium and Disequilibrium in Modeling Regional Growth and Decline: a Critical Reassessment", *Journal of Regional Science*, Vol. 33, No. 1 (1993): 69-84

Greene, William H. 2003. *Econometric Analysis, Fourth edition*. Prentice-Hall International

Greenwood, Michael J.; Hunt, Gary L.; Rickman, Dan S.; Treyz George I. "Migration, Regional Equilibrium and the Estimation of Compensating Differentials." *The American Economic Review*, Vol. 81, Issue 5, Dec (1991):1382-1390.

Gyourko, J. and Tracy, J. "The Structure Of Local Public Finance And The Quality Of Life", *Journal of Political Economy*, **99**, (1991):774:806

Hoch, Irving and Judith Drake, 1974: Wages, Climate, and the Quality of life. *Journal of Environmental Economics and Management,* **1,** 268-296.

Malpezzi, Stephen; Chun, Gregory H. and Green, Richard K.,"New Place-to-Place Housing Price Indexes for U.S. Metropolitan Areas, and their Determinants", *Real Estate Economics*, Vol. 26, No.2 (1998): 235-274

McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behavior," in: P. Zarembka, ed., *Frontiers of Econometrics*. (1973) New York: Academic Press.

McFadden, D. Modeling the choice of residential location, in: A. Karlqvist, L. Lundquist, F. Snickars, J. Weibull (Eds.), Spatial Interaction Theory and Planning Models (1978) Amsterdam, North Holland.

Moore, Thomas Gale. "Health and Amenity Effects of Global Warming." *Economic Inquiry*, Vol. 26, July (1998):23-41.

Moshe, Ben-Akiva ; and Lerman, S. R.. 1985. "Discrete Choice Analysis: Theory and Application to Travel Demand." MIT Press, Cambridge.

Mueser, Peter R. and Graves, Philip E. "Examining the Role of Economic Opportunity and Amenities In Explaining Population Redistribution." *Journal of Urban Economics* Vol. 37 (1995): 176-200.

Roback, Jennifer. "Wages, Rents, and the Quality of Life" *The Journal of Political Economy* Vol. 90, Issue 6, Dec. (1982): 1257- 1278.

Roback, Jennifer. "Wages, Rents, and Amenities: Differences among Workers and Regions" *Economic Inquiry*, Vol. 26, Issue 1, Jan (1988):23-41

Rosen, Sherwin. "Hedonic prices and implicit markets: Product differentiation in pure competition." *Journal of Political Economy,* Vol. 82, (1974): 34-55.

Ruggles, Steven, Sobek, Matthew, Alexander, Trent, Fitch, Catherine A., Ronald, Goeken , Hall, Patricia Kelly, King , Miriam, and Ronnander, Chad. Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2004.

Savageau, David and Ralph D'Agostino. Places Rated Almanac: Millennium Edition, (2000) Hungry Minds Inc., New York, NY.

Smith, V.K., 1983: The role of site and job characteristics in hedonic wage models. *Journal of Urban Economics,* **13(3),** 296-321.

Suchan, Trudy A., Perry, Marc J., Fitzsimmons, James D. Juhn, Anika E., Tait, Alexander M. Brewer, Cynthia A. Census Atlas of the United States, Series CENSR-29. U.S. Census Bureau, Washington, DC, 2007

US Census Bureau, "Demographic Trends in the 20th Century", Nov. 2002

US Census Bureau, Census Regions and Divisions of the United States, Available online at http://www.census.gov/geo/www/us_regdiv.pdf

U.S. Climate Normals, 1971-2000, **Monthly Divisional Normals & Standard Deviations - CLIM85,**http://cdo.ncdc.noaa.gov/cgibin/climatenormals/climatenormals.pl September 8, 2003

U.S. Dept. of Justice, Federal Bureau of Investigation. Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest And Offense Data, 2000 [Computer File]. 3rd ICPSR Ed. Ann Arbor, MI: Inter-University Consortium. For Political and Social Research [Producer And Distributor], 2002.

U.S. Environmental Protection Agency. Environmental Benefits Mapping and Analysis Program (BenMAP), Available online at http://www.epa.gov/air/benmap/