7-5-2012

Responding to Natural Hazards: The Effects of Disaster on Residential Location Decisions and Health Outcomes

James Ian Price

James Ian Price
Candidate

Department of Economics
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Alok K. Bohara, Chairperson

Robert P. Berrens

Janie M. Chermak

Wendy L. Hansen

RESPONDING TO NATURAL HAZARDS:

THE EFFECTS OF DISASTER ON RESIDENTIAL LOCATION DECISIONS AND HEALTH OUTCOMES

by

JAMES IAN PRICE

B.A., Sociology, Calvin College, 2001
M.A., Economic Theory, University of New Mexico, 2010

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Economics

The University of New Mexico
Albuquerque, New Mexico

May, 2012

DEDICATION

To Emily,

with whom life is a joy

ACKNOWLEDGEMENTS

This work would not have been possible without the support of my committee members, family and friends. I thank the chair of my committee, Professor Alok K. Bohara, for his assistance during the writing of this dissertation—and in particular for challenging and inspiring me to engage new research questions and methodologies. My skills as an applied economist are immensely better because of his guidance and encouragement. I thank Professor Robert Berrens, Professor Janie M. Chermak, and Professor Wendy L. Hansen for their interest in my research and their valuable insights, which have improved this work. In sum, I am grateful to my entire committee for their extraordinary mentorship throughout my time at the University of New Mexico. I am also indebted to Mrs. Judy Parker for her generosity through the Alfred L. Parker Award. The financial assistance provided through this award allowed me to pursue my dissertation research with greater focus and determination.

Numerous other people have contributed to my career development. I would especially like to acknowledge my parents, Tim and Stephanie Price, who have provided unwavering support and encouragement to me in every endeavor. They instilled in me the confidence to attempt things that initially seem beyond my capabilities. Finally, I thank my wife Emily for her counsel and remarkable patience during my graduate career, as well as her willingness to edit my research papers whenever needed. Her warmth and joyfulness helped me through the challenges and were instrumental to my success as a graduate student.

Responding to Natural Hazards:
The Effects of Disaster on Residential Location Decisions and Health Outcomes

by

James Ian Price

B.A., Sociology, Calvin College, 2001

M.A., Economic Theory, University of New Mexico, 2010

Ph.D., Economics, University of New Mexico, 2012

ABSTRACT

Exposure to natural hazards is rapidly increasing due to growing populations within floodplains and along hazard-prone coastlines. This trend, coupled with potential increases in the frequency and intensity of extreme weather events from climate change, underscores the importance of disaster research and continued advancements in hazard risk mitigation. This dissertation conducts analyses regarding the effects of natural hazards on residential location choice, county migration rates, mental health status, and displacement. The results have practical implications for disaster risk management.

Chapter 2 estimates household willingness-to-pay to live in lower hazard-risk areas. A model of residential location choice is developed in which households select the location that maximizes expected utility. Empirical estimates are obtained using a two-

stage estimation process that exploits spatial variation in labor markets, housing markets, and environmental amenities across U.S. metropolitan statistical areas. Results indicate an annual willingness-to-pay of $275 per household for a marginal reduction in the expected number of earthquake, hurricane, and flood events per 1000 years.

Chapter 3 estimates the relationship between county-level net in-migration rates and the expected frequency of hazard events. Empirical estimation is complicated by the presence of spatial dependency and heterogeneity. These issues are addressed using spatial simultaneous autoregressive estimation and geographically weighted regression. Results show that net in-migration rates are negatively correlated with expected frequency. Moreover, the effects of hazard risk are strongest in the Southern U.S.; a region susceptible to increased hazard intensity from climate change.

Chapter 4 contains two separate analyses regarding the wellbeing of individuals affected by Hurricanes Katrina and Rita. The first analysis evaluates the effect of post-disaster stress and vulnerability on long-term mental health. Results show that the likelihood of being diagnosed with an adverse mental health condition increases with stress and vulnerability levels. The second analysis evaluates the determinants of displacement and the duration of displacement. Results show that housing damage is the most important predictor of displacement and displacement duration. Social support has a positive impact on displacement but a negative impact on the displacement duration, implying that social networks provide accommodations during hazard events as well as assistance in returning home.

Table of Contents

List of Figures

List of Tables

List of Acronyms

2SLS	Two-Stage Least Squares
ACCRA	American Chamber of Commerce Researchers Association
ACS	American Community Survey
ASC	Alternative Specific Constant
CBP	County Business Patterns
CCD	Core of Common Data
CCDB	City and County Database
DRM	Disaster Risk Management
DSI	Disaster Severity Index
EM-DAT	Emergency Events Database
EPA	Environmental Protection Agency
ESRI	Environmental Systems Research Institute
FEMA	Federal Emergency Management Agency
GAD	Generalized Anxiety Disorder Survey
GIS	Geographic Information Systems
GMM	Generalized Methods-of-Moments
GRDP	Global Risk Data Platform
GS2SLS	Generalized Spatial Two-Stage Least Squares
GWR	Geographically Weighted Regression
IPPC	Intergovernmental Panel on Climate Change

IV	Instrumental Variable
ML	Maximum Likelihood
MSA	Metropolitan Statistical Areas
NAICS	North American Industry Classification System
NCDC	National Climatic Data Center
NOAA	National Oceanic and Atmospheric Administration
NPL	National Priorities List
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
PDVI	Post Disaster Vulnerability Index
PHQ	The Patient Heath Questionnaire
PSID	Panel Survey of Income Dynamics
PTSD	Posttraumatic Stress Disorder
PUMA	Public Use Microdata Area
PUMS	Public Use Microdata Sample
PRA	Places Rated Almanac
RUM	Random Utility Model
SAC	Spatial Simultaneous Autoregressive Estimation
SSI	Social Support Index
UNISDR	United Nations International Strategy for Disaster Reduction
WTP	Willingness-to-Pay

Chapter 1: Introduction

1.1 Natural Hazard Exposure

Natural hazards (e.g. earthquakes, cyclones, floods, wildfire, and volcanic activity) have profound consequences for affected populations. The salient effects are adverse health outcomes, property damage, reduced income, and disruption of service flows from community and environmental amenities. According to the Emergency Events Database (EM-DAT), 3740 major natural hazards occurred globally between 2000 and 2011.[1] These events resulted in an estimated 943,000 deaths and $1.2 trillion in damages. They directly or indirectly affected 1.7 billion people. During the same period, the U.S. experienced 246 major hazard events—as well as several thousand localized events. Major events resulted in 4293 deaths and $350 billion in damage. They affected an estimated 2.1 million Americans.

Hazard effects have been, and continue to be, exacerbated by urban growth within floodplains, wildfire zones, and along hazard-prone coastlines. For instance, the number of people living in areas susceptible to flooding increased from 32.5 to 69.4 million (114%) between 1970 and 2010 (UNISDR, 2011). Although experienced in every region,

---

[1] Authors calculation based on EM-DAT data. This figure, and subsequent impact estimates, pertain to earthquakes, floods, storms, volcanic activity, and wildfire. Additional hazards, such as drought, epidemics and extreme temperatures are not included.

these increases were most pronounced in South Asia and East Asia where exposed population rose by 37 million. During the same period, the number of people living in areas susceptible to tropical cyclones increased from 65.9 to 122.5 million (192%). These increases were largest in East Asia and Organisation for Economic Co-operation and Development (OECD) countries. In the U.S., more than one-third of the population currently resides in a hazard-prone area (Dilley et al., 2005).

Mitigating the cost of natural hazards will require advances in scientific knowledge, public policy and technology. Despite a growing body of literature, the behavioral and health responses of economic agents to natural hazards are still not fully understood. This dissertation addresses these gaps through an examination of individuals and households in the U.S. Specifically, the dissertation contains two analyses regarding the effects of hazard risk, as opposed to hazard events, on residential location choice. It also contains analyses of long-term mental health outcomes and displacement duration for households affected by Hurricanes Katrina and Rita.

1.2 Theoretical Framework and Terminology

A common theoretical framework underlies the majority of recent research on natural hazards. The framework's central concept is hazard risk (or disaster risk), which is defined as the likelihood that societal functions will be severely disrupted due to an extreme weather, climate, or geologic event over a specified time period (IPCC, 2011; UNISDR, 2009). Hazard risk is typically subdivided into extensive and intensive risk. Extensive risk refers to the risk associated with low-severity high-frequency disasters, such as periodic flooding, storms, and drought. While destructive, extensive-risk disasters are highly localized and typically confined to rural areas or urban margins (UNISDR, 2009). Intensive risk refers to the risk associated with high-severity low-frequency disasters. Major earthquakes, volcanic eruptions, flooding, tsunamis, and tropical cyclones are examples of intensive-risk disasters. These risks have wide-ranging impacts that often correspond to densely populated and urban areas (UNISDR, 2009).

In the U.S., unlike in low- and middle-income countries, disaster-related mortality is primarily related to extensive risk. Since 1960, approximately 89% of disaster-related mortality resulted from extensive-risk hazards (UNISDR, 2011). Only in 2005, with the devastation caused by Hurricane Katrina, did intensive-risk mortality exceed that of extensive risk. High hazard mortality rates are concentrated in the Midwest, West, and South census regions. The highest rates occur in low-income rural areas of Montana, North Dakota, South Dakota, Nebraska, Kansas, Utah, Oklahoma, Texas, and Arkansas. In contrast to mortality, hazard-related economic loss in the U.S. is primarily related to intensive risk. Between 1960 and 2008, greater than two-thirds of hazard-related damage is attributed to high-severity low-frequency disasters (UNISDR, 2011).

Hazard risk, either extensive or intensive, for a given community is determined by three factors: hazard events, exposure, and vulnerability. Hazard events simply refer to the occurrence of extreme weather, climate, or geologic evenst within a specific time period. Exposure is used to describe the presence of people and assets (i.e. economic, social, environmental, or cultural assets) within hazard-prone areas. Finally, vulnerability is used to describe a community's susceptibility, independent of exposure, to hazard-related damage and loss. Vulnerability depends on a number of community attributes, including socioeconomic characteristics, infrastructure design and quality, land use patterns, and disaster preparedness.

Although they comprise three distinct concepts, hazard risk, exposure, and vulnerability are highly interrelated. For instance, exposure is partially determined by hazard events. An increase in hazard events, or the probability of hazard events, encourages economic agents to locate away from hazard-prone areas. These location decisions reduce population, assets, and ultimately the level of exposure within these areas. Similarly, exposure is partially determined by vulnerability. Higher levels of vulnerability place economic agents at greater risk and encourage migration to safer areas. Reversing this relationship, a community's vulnerability is also affected by its level of exposure. In many urban areas, particularly in low- and middle-income countries, rapid population growth has outpaced infrastructure development. In these situations, those without access to adequate infrastructure are highly vulnerable to natural hazards, which in turn increases vulnerability within the entire urban area.

A community's level of hazard risk will vary over time in response to changes in hazard events, exposure, and vulnerability. Policy interventions mostly influence hazard risk through changes in vulnerability—although exposure levels and, to a lesser extent, hazard events can also be affected. Construction of disaster-resistant infrastructure, implementation of building codes and land use restrictions, and disaster preparedness measures (e.g. contingency planning, emergency supply stockpiles, and public information systems) are examples of interventions designed to reduce hazard risk. These policies have proven to substantially reduce hazard-related losses. The most striking improvements concern mortality rates. The annual mortality rate for extensive-risk disasters in the U.S. declined by 35% since 1989 (UNISDR, 2011). While public policies also mitigate economic loss, continued population growth and economic development (i.e. higher levels of exposure) have caused an overall rise in damage. Since 1960, annual economic losses from extensive-risk hazards have increased by approximately 60% in the U.S. (UNISDR, 2011). At the global level, this upward trend in hazard-related economic losses is even more pronounced—with large spatial and inter-annual variation.

1.3 Hazard Risk and Climate Change

In recent years, considerable attention has been given to identifying the effects of climate change on extreme weather events. A recently published special report from the Intergovernmental Panel on Climate Change (IPCC) is the most comprehensive treatment of the subject. Due to high levels of uncertainty in climate prediction, the IPCC report attaches confidence rankings (e.g. virtually certain, very likely, likely, unlikely, and very unlikely) to its key findings. Rankings are based on the assessed validity of scientific evidence and agreement within the scientific community.

Increasing exposure has been, and will remain, the major cause of long-term trends in hazard risk. There is, however, evidence to suggest that changing climate conditions will augment hazard risk in coming decades. It is virtually certain (99-100% probability) that climate change will increase the frequency and magnitude of temperature extremes over most land areas (IPCC, 2011). It is likely (66-100% probability) that the frequency of heavy precipitation will increase over many land areas (IPCC, 2011). Even regions that experience an overall decline in precipitation may notice an increase in heavy precipitation. Heavier precipitation is expected to impact local and regional flood patterns; however, limited evidence and complex hydrologic systems have prevented more explicit flood predictions. It is also very likely (90-100% probability) that rising sea levels will contribute to extreme high water events, causing erosion and inundation along coastal areas (IPCC, 2011). These events may contribute to considerable economic loss for small island states and other low-lying coastal zones. While the precise impact will vary by location, the combined effect of these changes (i.e. temperature extremes, heavy precipitation, and sea level rise) will likely be an increase in hazard risk and hazard-related losses.

With respect to cyclone activity, it is likely (66-100% probability) that the frequency of hazard events will either decrease or remain unchanged. This prediction holds at both the global level and regionally in the Atlantic Ocean (IPCC, 2011). It is also likely (66-100% probability) that changing climate conditions will increase the average wind speed of tropical cyclones (IPCC, 2011). This suggests that future hurricane activity in the U.S. will be less frequent but more destructive. Because of these opposing effects, implications for hazard risk are ambiguous.

According to IPCC findings, climate change will impose considerable, if uncertain, costs on society through more frequent and intense extreme weather events. This reemphasizes the importance of continued research on behavioral and health responses to natural hazards. It also highlights the need for researchers and policymakers to consider how economic agents adapt to changing hazard risk.

1.4 Hazard Risk Mitigation

Hazard risk can be reduced at the individual, community, and governmental level. The policies and actions undertaken to reduce the adverse impacts of natural hazard risk are commonly referred to as disaster risk management (DRM). Individuals have several options for reducing hazard risk, including purchasing disaster insurance, improving their home's ability to withstand extreme weather, and migrating to lower-risk areas. As previously mentioned, government options for reduction of hazard risk include disaster-resistant infrastructure, building codes, land use restrictions, and disaster preparedness measures. Additional examples include flood control measures (i.e. levees, channel diversion, and pumping stations), restoration and protection of critical natural resources (i.e. coral reefs, coastal wetlands, and mangrove forests), and disaster insurance programs. Government-operated disaster insurance is available in several countries. In the U.S. both state and federal agencies have established entities to provide disaster insurance. The most prominent example is the National Flood Insurance Program, which was established in 1968 to protect property owners, within participating communities, against flood losses. As of 2010, approximately 5.5 million properties purchased insurance through this program (Holladay and Schwartz, 2010).

Despite these options, the tendency to heavily discount future risks suggests that both individuals and governments under invest in DRM (UNISDR, 2011). From a policy perspective, short political time horizons discourage government investment even when interventions are cost-effective. Results from an analysis by the United Nations International Strategy for Disaster Reduction (UNISDR) found that few countries have a dedicated budget for DRM (UNISDR, 2011). Where dedicated budgets do exist, they comprise only a small proportion of government revenue. Thorough analyses of DRM in Mexico and Colombia indicate annual investment of 0.01 and 0.08% of government revenue, respectively (Moreno and Cardona, 2011). Moreover, DRM investments usually focus on extensive-risk disasters, often ignoring intensive risk (UNISDR, 2011). Spending related to intensive-risk disasters is usually in the form of hazard relief rather than prevention or mitigation.

In recent years there has been a considerable, and growing, international effort to improve investments in DRM. Governments have been encouraged to systematically identify and quantify the benefits and costs of DRM interventions. To this end, sophisticated modeling techniques have been developed to estimate the potential economic loss from hazard events. A detailed discussion of these models is available in Cardona et al. (2008). Despite advances in these models they remain only partial estimates of hazard-related cost. In particular, they fail to account for non-market values, medium- to long-term economic impacts, and health impacts brought about by hazard events.

1.5 Contributions of this Dissertation

This dissertation conducts four analyses regarding the behavior and health effects of natural hazards in the U.S. It makes extensive use of publicly available data and geographic information systems (GIS). Household survey data is taken from Public Use Microdata Samples (PUMS) and the Panel Survey of Income Dynamics (PSID). Socioeconomic data for various geographic entities is primarily derived from the U.S. Census Bureau's City and County Database (CCDB) and County Business Patterns (CBP). Environmental characteristics, discussed in detail in subsequent chapters, are obtained from numerous sources, including the National Climate Data Center (NCDC), the Environmental Protection Agency (EPA), and the Global Risk Data Platform (GRDP). This data, which is often constructed using GIS, represents a unique collection of sub-national environmental information.

The first analysis, presented in Chapter 2, estimates household willingness-to-pay (WTP) to live in areas that are less likely to experience a major hazard event. A model of residential location choice is developed in which households select the location that maximizes their expected indirect utility. Empirical estimates are obtained using a two-stage estimation process that exploits spatial variation in labor markets, housing markets, and environmental amenities across U.S. metropolitan statistical areas (MSA). The first stage employs a conditional logit model to estimate a set of alternative specific constants (ASC)—following the method developed by Berry et al. (1995). The vector of ASC reflects household preferences for MSAs and is loosely interpreted as a quality-of-life index. In the second stage, ASC values are regressed against location-specific attributes (e.g. economic, demographic and environmental characteristics). Coefficients from this regression are used to estimate implicit prices and to determine the relative impact of each attribute on quality-of-life.

Chapter 3 models domestic migration patterns among U.S. counties. Migration decisions are motivated by a number of economic, social, and environmental considerations. Among these considerations is the potential for high-risk low-probability hazard events. Chapter 3 estimates the relationship between county-level net in-migration rates and the expected frequency of earthquakes, hurricanes, and floods. Empirical estimation is complicated by the presence of spatial dependency and heterogeneity. These issues are addressed using two separate regression techniques: spatial simultaneous autoregressive estimation (SAC) and geographically weighted regression (GWR). Coefficients from the SAC regression quantify the effects of hazard risk on migration rates, after controlling for socioeconomic, environmental characteristics, and spatial autocorrelation. GWR results examine spatial heterogeneity, identifying regions where hazard risk has the greatest impact on migration.

Chapter 4 consists of two separate analyses regarding the wellbeing of individuals affected by Hurricanes Katrina and Rita. The first part evaluates the effects of post-disaster stress on long-term mental health. This analysis is conducted using a simultaneous equations model. The second part evaluates the determinants of household displacement, which has been linked to adverse health and economic outcomes. A hurdle model is developed in order to estimate the relationship between household characteristics, displacement, and displacement duration. Particular attention is given to the role of social support in reducing adverse outcomes. The econometric techniques employed in both analyses contribute to the literature by advancing previous work.

Results from the dissertation offer important insight into the behavioral and health effects of natural hazards, which is essential to developing successful and cost effective policy interventions. Foremost among these findings are WTP estimates for hazard-risk reduction. These estimates have practical implications as a point of departure for conducting benefit-cost analyses of DRM interventions—particularly with regards to interventions that lower the frequency of hazard occurrence (e.g. flood control measures). Results also reveal the role of hazard risk in migration decisions, stressing the importance of continued research in the area of climate change adaptation. Finally, the dissertation provides useful information regarding the determinants of mental health and displacement following intensive disasters.

Chapter 2: Residential Sorting and the Value of Hazard Risk Reduction

2.1 Introduction

Growing populations and continued urban development in hazard-prone areas highlight the need for DRM policies. DRM encompasses a variety of interventions designed to reduce vulnerability to natural disasters. Common interventions include disaster-resistant infrastructure (e.g. reinforced electrical and transportation systems), flood control measures (e.g. levees, channel diversion, and pumping stations), building codes, land use restrictions, and reforestation in critical coast and mountain areas (Freeman et al., 2003). DRM policy decisions are often evaluated within the framework of benefit-cost analysis. In most cases, the costs of risk-mitigation (e.g. construction and monitoring) are readily available and easy to quantify. Benefits, on the other hand, must be estimated. Quantifying these benefits is vital to the development of effective and efficient interventions.

Hedonic property models are suitable for valuing the benefits of risk mitigation within localized housing markets. These models establish a statistical relationship between prices and housing characteristics, and subsequently recover marginal WTP values. Several hedonic studies have estimated WTP for risk mitigation (Bin et al., 2008). While WTP values vary, findings from these studies consistently indicate that homeowners pay a premium to reside in safer neighborhoods (i.e. outside a floodplain or away from a fault line). Because these studies are conducted within localized markets they primarily capture the value placed on avoiding property damage. They do not, however, capture the value placed on avoiding broader consequences of natural disaster, such as temporary reductions in income and the disruption of service flows from community and environmental amenities. Estimating these values, which occur regardless of damage to the homeowner's property, requires an inter-market analysis.

This paper employs a residential sorting model, developed by Bayer et al. (2009), that exploits spatial variation in labor markets, housing markets, and local attributes across U.S. MSAs. The analysis estimates parameters of the indirect-utility function and recovers marginal WTP values for several socioeconomic and environmental amenities, including reductions in the expected number of hazard events. This sorting model differs from the hedonic method in two ways. First, the model is conducted across several housing and labor markets. As a result, WTP estimates reflect the value placed on avoiding all disaster-related consequences—not just property damage. Second, the model incorporates migration costs as a control variable. The inability to account for migration is often seen as a shortcoming of the hedonic method that leads to omitted-variable bias (Bayer et al., 2009). Inclusion of migration cost in the residential sorting model eliminates this source of bias and results in more accurate WTP values (Bayer et al., 2009).

2.2 Literature Review

This analysis builds upon two distinct but related lines of research. The first pertains to the non-market valuation of disaster risk, which is primarily conducted using hedonic models of local housing markets.[2] The second pertains to the residential sorting

---

[2] In the hazard literature risk refers to the interaction of events, exposure, and vulnerability. This dissertation, however, often uses risk to mean the probability of

model. These models are increasingly used for non-market valuation of environmental amenities—although they have not been applied to the risk of natural disaster. This section begins with a review of Ehrlich and Becker (1972), a seminal paper that develops a theoretical model of risk-mitigating behavior. This is followed by literature reviews of disaster-risk valuation and residential sorting models.

2.2.1 Theory of Hazard Risk Valuation

Ehrlich and Becker (1972) provide a theoretical basis for valuing risk mitigation. They distinguish between three types of mitigation: market insurance, self-insurance and self-protection. Self-insurance refers to investments that reduce disaster-related losses but do not reduce the probability of disaster occurrence. Examples of self-insurance include use of earthquake-resistant construction techniques, installation of storm shutters and maintenance of fire sprinkler systems. In contrast, self-protection refers to investments that reduce the probability of disaster occurrence. The foremost example of self-protection is migration to lower-risk areas.[3] The distinction between self-insurance and self-protection, although subtle, has important implications within the Ehrlich and Becker (1972) model. In particular, market insurance and self-insurance are regarded as

---

hazard occurrence. This definition is consistent with the economic interpretation of risk within expected utility theory.

[3] Ehrlich and Becker (1972) recognize the distinction between self-insurance and self-protection to be somewhat artificial since many behaviors fall into both categories. Migration from high- to low-risk areas is a prime example (Brookshire et al., 1985; Simmons et al., 2002). In the present context, migration is considered a mechanism for reducing the probability of risk and is classified as self-protection.

substitutes while market insurance and self-protection are regarded as complements. Thus, a decrease in the real price of market insurance would lead to increased demand for self-protection and decreased demand for self-insurance.

Ehrlich and Becker (1972) demonstrate that when market insurance is offered at actuarially fair rates, households will fully insure themselves against disaster-related losses. Moreover, fully insured households are indifferent between states of the world (i.e. a disaster-event state and a no-disaster-event state) and have no financial incentive to adopt methods of self-insurance or self-protection (Ehrlich and Becker, 1972; Simmons et al., 2002). Without investments in self-insurance or self-protection the full value of disaster risk-mitigation is reflected in the equilibrium price of market insurance. In reality, however, the ability of market prices to signal mitigation values is hampered by the presence of imperfect information, moral hazard, and non-financial losses. Evidence to this effect is reflected in the large share of households that remain uninsured (Beron et al., 1997; Bin and Polasky, 2004; Kunreuther, 2006) and the considerable use of self-insurance and self-protection (Fronstin and Holtmann, 1994; Simmons et al., 2002; Bin and Polasky, 2004). On the whole, these observations suggest that non-market valuation techniques are required to accurately value disaster risk mitigation.

There is some concern regarding the validity of WTP values based on self-protection investments. Berger et al. (1987) demonstrate that the value of risk-mitigation is equivalent to the marginal rate of technical substitution between risk mitigation and self-protection, providing that the risk is exogenous. Under a few reasonable assumptions, these values represent the lower bound of WTP. Shogren and Crocker (1999) develop a model of endogenous risk for a hazard event (e.g. exposure to hazardous waste). Under certain assumptions they find that an increase in the level of risk has an ambiguous effect on self-protection, suggesting WTP values are unreliable. Quiggin (1992) extends this analysis; assuming that preferences display decreasing absolute risk aversion and that self-protection activities are separable from unobserved risk factors. Findings indicate that WTP estimates are valid lower-bound measures of risk mitigation. This analysis circumvents these issues by assuming the risk of natural disaster is exogenous to self-protection activities. In the case of earthquakes, hurricanes, and floods this seems a reasonable assumption.

2.2.2 Non-market Valuation of Hazard Risk

Numerous studies have estimated non-market values for disaster risk-mitigation. These studies typically employ hedonic property models within a localized housing market.[4] Studies focus on a variety of hazards types, including earthquake hazards (Brookshire and Schulze, 1980; Brookshire et al., 1985; Bernknopf et al., 1990; Beron et al. 1997; Önder et al., 2004), hurricane hazards (Simmons et al., 2002; Hallstrom and Smith 2005), flood hazards (MacDonald et al. 1987; Speyrer and Ragas, 1991; Harrison et al., 2001; Bin and Polasky, 2004; McKenzie and Levendis, 2008) and wildfire hazards (Loomis, 2004; Donovan et al., 2007).

In addition to hazard type, hedonic studies vary in their measure of disaster risk. Smith (1986) argues that individual behavior, and the value placed on safety, is governed

[4] A notable exception is Brookshire and Schulze (1980), which uses the survey-based method, contingent valuation, to estimate the effects of earthquake risk on the housing market in Los Angeles, California.

by perceived risk rather than statistical estimates. This view is supported by a number of recent studies that find divergence between perceived and actual risk (Baker et al., 2009; Horney et al., 2010). In many cases, however, detailed information regarding perceived risks is unavailable. In its place, researchers have relied on various proxy measures, including scientific risk estimates (MacDonald et al. 1987; Harrison et al., 2001; Önder et al., 2004), government notices (Brookshire et al., 1985; Bernknopf et al., 1990; Donovan et al., 2007) and natural experiments (Beron et al. 1997; Bin and Polasky, 2004; Loomis, 2004; Hallstrom and Smith 2005; McKenzie and Levendis, 2008).

Results from these studies indicate a negative correlation between property values and proxy variables for perceived risk. Brookshire et al. (1985) estimate the premium on single-family residences located outside Special Studies Zones in Los Angeles and San Francisco. Special Studies Zones are designated areas characterized by proximity to faults lines and elevated risk of earthquake activity. Findings indicate a 5.6% reduction in the value of properties within the Special Studies Zones. Bin and Polasky (2004) evaluate the effects of flood hazard on housing prices in Pitt County, North Carolina. On average, houses located within a floodplain sell for 5.7% less than houses outside the floodplain. Hallstrom and Smith (2005) incorporate a difference-in-difference component into their hedonic analysis of hurricane risk in Lee County, Florida. They find that average property values declined by 19% in response to a nearby hurricane event. Often studies compare their estimated premiums to the capitalized value of market insurance rates. In some cases, estimated premiums are similar to market rates (MacDonald et al. 1987), while in other instances premiums are larger (Speyrer and Ragas, 1991; Bin and Polasky, 2004) or smaller (Harrison et al., 2001).

Findings also highlight the dynamic nature of risk perception. Studies that estimate pre- and post-disaster hedonic models often find larger premiums on low-risk housing after the disaster (Bin and Polasky, 2004; Loomis, 2004; McKenzie and Levendis, 2008), suggesting that consumers update their risk perception following hazard events.[5] Whether updated risk perceptions are permanent, or erode over time, has not been thoroughly addressed in the hedonic literature.

2.2.3 Residential Sorting Models

Residential sorting models are increasingly used to estimate non-market goods and services. These models, which are derived from the Tiebout sorting model, use the residential location decision of individual households to recover parameters of the indirect utility function. Empirical estimates are obtained using familiar discrete-choice techniques. Sorting models have been used to recover household preference for school quality (Bayer et al. 2004), public safety (Bayer et al., 2005), and racial segregation (Bayer and McMillan, 2006). With respect to environmental amenities, these models have been used to value green space (Klaiber and Phaneuf, 2010) and air pollution (Tra, 2010; Finney et al., 2011).

Timmins (2007) and Bayer et al. (2009) extend the model to evaluate residential sorting at a national, rather than localized, level. Timmins (2007) analyzes residential sorting across Brazilian states using household data from the 1991 Brazilian Demographic Census. The model values changes in climate amenities (i.e. temperature and precipitation), while controlling for migration costs, population density, and distance

---

[5] In contrast, Beron et al. (1997) find that the premium for homes located outside Special Studies Zones in San Francisco fell after the 1989 Loma Prieta earthquake.

to economic centers. Results indicate a total annual cost of \$1.6-\$8.1 billion for moderate changes in climate. Bayer et al. (2009) analyze residential sorting across U.S. MSAs using household data from the 1990 and 2000 PUMS. Their model estimates WTP for air quality improvements, while controlling for migration costs and socioeconomic characteristics. Results indicate a marginal WTP between \$149-\$185 per household for a one-unit reduction in average concentrations of particulate matter. A distinguishing feature of these models is the inclusion of migration costs as a control variable. Bayer et al. (2009) demonstrate that without controlling for migration costs WTP estimates will be biased.

2.3 Theoretical Framework

This section develops a theoretical model of residential location choice within the context of disaster risk. It begins with a brief discussion of key similarities and differences between hedonic and residential sorting models. Particular emphasis is placed on the underlying assumptions of market equilibrium and mobility costs.

2.3.1 Hedonic and Residential Sorting Models

The hedonic method operates under a number of assumptions. One key assumption is that housing and labor markets are in equilibrium, which itself follows from assumptions of utility maximization and competitive markets. Market equilibrium implies that households must make tradeoffs between prices and amenities (Taylor, 2003). These tradeoffs reveal the marginal implicit prices inherent in the marketplace. On occasion, exogenous shocks will create disequilibrium in housing and labor markets. Disequilibrium is a temporary condition, as markets will adjust, primarily through migration, until a new equilibrium is achieved.

Another key assumption of the hedonic model is costless mobility. This assumption implies that utility-maximizing agents will migrate instantaneously in response to an exogenous shock—and that markets are always in equilibrium. Evidence suggests, however, that mobility costs are substantial. Migration costs include transportation costs, search costs (i.e. for housing and employment), and psychological costs associated with leaving social networks and familiar cultural settings. Table 2.1 shows migration rates between a household head's birth region and their region of residence. The diagonal elements indicate the percent of individuals residing in their birth region, while off-diagonal elements indicate the percent that have migrated. As evidenced in Table 2.1, individuals are much more likely to remain in their birth region than to migrate. This pattern, which also holds for smaller geographical areas (i.e. states and MSAs), suggests movement is inhibited by migration costs.

Bayer et al. (2009) demonstrate that WTP values recovered from hedonic analyses are biased in the presence of migration costs. To see this, consider a two-city scenario where cities have identical characteristics and markets are in equilibrium. An improvement in one city's amenities, for instance a decrease in pollution levels, creates market disequilibrium. Under the assumption of costless mobility, some households will move to the location with improved pollution levels. Their movement shifts supply and demand schedules within the housing and labor markets, resulting in relatively lower wages and higher housing costs at the improved location. Migration will continue until a new equilibrium is reached. In this case, WTP for improved air quality is determined by the difference in wages and housing prices between the two locations. On the other hand, if mobility is costly, price differences understate the value of air quality improvements.

Actual WTP equals the price difference plus the cost of migration, including the psychological cost.

This paper builds on the residential sorting model developed by Bayer et al. (2009), which explicitly controls for the cost of migration. This model exploits patterns of residential location choice to recover preferences of various location-specific attributes. Like the hedonic model, the residential sorting model assumes equilibrium in the labor and housing markets. It is this equilibrium that ensures tradeoffs are being made between prices and amenities, and that WTP values can be recovered. The two approaches differ in their underlying estimation function. The hedonic model estimates parameters for the price function, while the residential sorting model estimates parameters for the indirect utility function.

2.3.2 Theoretical Model

The theoretical model, adapting Day and Winer (2001), combines elements of expected utility theory and discrete choice modeling. Households solve a two-part optimization problem. First, they maximize their expected utility, subject to a budget constraint, in order to determine the optimal allocation of income between consumption goods. Second, households select the residential location that maximizes their utility, taking into account wages, prices, socioeconomic characteristics, environmental amenities and disaster risk. The implicit assumption is that households have knowledge of each location's labor market, housing market, amenities, and risk levels.

For simplicity, households face only two states of the world: the occurrence and non-occurrence of a hazard event. The expected utility at any location is a function of consumption, migration costs, location-specific attributes, and state of the world. Equation 2.1 specifies the utility function for household i at location j.

$$E(U_{ij}) = (1-\pi_j)U_{ND}(C_i, H_i; X_j, M_{ij}) + \pi_j U_D(C_i, H_i; X_j, M_{ij}) \qquad (2.1)$$

Here C represents household consumption of a composite numeraire good, H consumption of housing services, X a vector of location-specific attributes (i.e. socioeconomic characteristics and environmental amenities), M migration costs, and π the probability of hazard occurrence. Utility received during the disaster and non-disaster state of the world are designated U_D and U_{ND}, respectively. Due to reductions in income, housing services, and amenity levels that occur during hazard events, utility in the non-disaster state is assumed to be greater than utility in the disaster state ($U_{ND} > U_D$). Utility in both states of the world is increasing in C and H and twice differentiable.

For each state of the world, households maximize their utility subject to a budget constraint. The budget constraint is presented in Equation 2.2:

$$I_{ij} = C + \rho_j H \qquad (2.2)$$

where I represents household income and ρ_j the price of housing services at location j. Substituting the optimal demand functions into the Equation 2.1 yields the expected indirect utility function, presented in Equation 2.3. Superscripts ND and D are included to emphasize that income, prices and location-specific attributes are state dependent.

$$E(V_{ij}) = (1-\pi_j)V_{ND}(I_{ij}^{ND}, \rho_j^{ND}; X_j^{ND}, M_{ij}) + \pi_j V_D(I_{ij}^{D}, \rho_j^{D}; X_j^{D}, M_{ij}) \qquad (2.3)$$

Having determined optimal utility in both states of the world, households turn to the second part of the optimization problem. During this stage, households evaluate the income, housing price, and location-specific attributes offered at each MSA and select the location that maximizes expected indirect utility. Formally, household i will choose to reside at location j if:

$$E(V_{ij}) > E(V_{ik}) \quad \forall \quad j \neq k, \quad k = 1,2,\ldots,j$$

The household will remain at its current location if utility is greatest at that location. Likewise, the household will migrate if utility, after accounting for migration costs, is greater elsewhere. As households select their optimal location they necessitate changes in labor and housing markets, as well as in location-specific attributes. When markets are in equilibrium, as this analysis assumes, expected indirect utility is constant: $E(V_{ij})=E(\overline{V_{ij}})$.[6] Constant utility implies that households are indifferent among location and must therefore make tradeoffs between expected income, housing prices, and local attributes. Empirical estimates of these tradeoffs, and corresponding WTP values, can be made using discrete-choice modeling techniques (e.g. conditional logit, mixed logit and nested logit models). These techniques require the researcher to specify a utility function and make distributional assumptions about the error term.

2.4 Econometric Model

Empirical estimation of the indirect utility function, which is rooted in the random utility model (RUM) framework, proceeds in stages. The central component of the analysis is a conditional logit model. Due to endogeneity issues the model is estimated in two stages, following the method developed by Berry et al. (1995). Prior to the

[6] Disequilibrium, which arises as consequence of exogenous shocks, creates non-constant utility that ultimately results in household migration. This migration continues until a new equilibrium is reached.

conditional logit it is necessary to generate estimates of income and housing price at each location. A discussion of these estimates is provided towards the end of the section.

2.4.1 Conditional Logit Model

Equation 2.4 specifies the utility function for individual *i* at location *j*. Following Bayer et al. (2009) and Timmins (2007) a Cobb-Douglas utility function is employed.

$$U_{ij} = C_i^{\beta_C} H_i^{\beta_H} e^{X_j^{\beta X} + M_{ij} + \xi_j + \eta_{ij}} \tag{2.4}$$

As before, C denotes consumption of a composite good, H consumption of housing services, X a vector of location-specific attributes, and M migration costs. The term ξ_j captures the average utility individuals receive from unobserved local attributes. The term η_{ij} is an idiosyncratic random error that varies across individual and location. For convenience the empirical model extracts away from the expected utility framework. Instead of expected utility it incorporates disaster-risk directly into utility function as a location-specific attribute.

In the first stage of the optimization problem, households determine their optimal allocation of income between housing and the composite good. The budget constraint is presented in Equation 2.5:

$$I_{ij} = C + \rho_j H \tag{2.5}$$

where I represents the individual's income and ρ_j the price of housing services at location j. Solving the utility maximization problem yields the following demand functions:

$$C_i = \left(\frac{\beta_C}{\beta_C + \beta_H} \right) I_{ij} \tag{2.6}$$

$$\mathrm{H_i} = \left(\frac{\beta_H}{\beta_C + \beta_H} \right) \frac{I_{ij}}{\rho_j} \tag{2.7}$$

The functions presented in Equations 2.6 and 2.7 indicate that optimal consumption is a constant proportion of income, regardless of geographic location. This property will be exploited later in the estimation process to circumvent problems with endogeneity. Substituting the demand functions into Equation 2.4 and taking the natural logarithm produces the indirect utility function, presented in Equation 2.8. As evident from the function, utility increases in income and decreases in housing price.

$$\begin{aligned} &\ln(V_{ij}) = \alpha + \beta_I \ln(I_{ij}) - \beta_H \ln(\rho_j) + \beta_X X_j + M_{ij} + \xi_j + \eta_{ij} \\ &\text{where} \quad \alpha = \beta_c \ln\left(\frac{\beta_c}{\beta_I}\right) + \beta_H \ln\left(\frac{\beta_H}{\beta_I}\right) \\ &\text{and} \quad \beta_I = \beta_C + \beta_H \end{aligned} \tag{2.8}$$

In the second stage of the optimization problem individuals select their utility maximizing location. In equilibrium, individuals make tradeoffs between income, housing price, and location-specific attributes. The exact nature of these tradeoffs is easily recovered from the indirect utility function. In particular, β_I, β_H, and β_X, which denote marginal utilities, can be used to calculate marginal rates of substitution and marginal WTP. Equation 2.9 depicts an individual's WTP for local attribute X_j.

$$WTP_i = \frac{MU_X}{MU_I} = \frac{\beta_X}{\beta_I} I_{ij} \tag{2.9}$$

Here MU_I is the marginal utility of income and MU_X the marginal utility of local attribute X_j. Note that while parameter values are constant the WTP varies with income.

Discrete-choice techniques can be used to recover parameter values. Due to its computational tractability this analysis employs a conditional logit model, where the error term (η_{ij}) is assumed independently and identically distributed (i.i.d.) type-I extreme value.[7] Equation 2.10 presents a generalized version of the model. It indicates the probability that individual i chooses to live at location j.

$$P[\ln(V_{ij}) \geq \ln(V_{ik}) \ \ \forall\, j \neq k] = \frac{e^{V_{ij}(I_{ij},\rho_j;X_j,M_{ij})}}{\sum_{k=1}^{j} e^{V_{ik}(I_{ij},\rho_j;X_j,M_{ij})}} \tag{2.10}$$

When estimated across a large number of individuals these probabilities represent the share of population living at each location (Timmins, 2007; Bayer et al., 2009). Before running the conditional logit model it is necessary to address several estimation issues. First, the model must be adjusted to reflect the inclusion of estimated incomes and housing prices. Second, it is necessary to parameterize migration costs. Third, the model must be modified to remove the simultaneity problem arising from correlation between housing prices and unobserved local attributes (Bruch and Mare, 2011).

By assumption, individuals know the income and housing price associated with each MSA. In practice, these values must be estimated. Following Bayer et al. (2009) hedonic models are used to estimate an individual's potential income at each location and to create an index of housing prices. Detailed descriptions of these models are provided in sections 2.4.2 and 2.4.3. Estimated values, depicted in Equations 2.11 and 2.12, are incorporated into the indirect utility function.

$$\ln\left(I_{ij}\right) = \ln\left(\hat{I}_{ij}\right) + \psi_{ij} \tag{2.11}$$

$$\rho_j = \rho_j^* \tag{2.12}$$

[7] The conditional logit model imposes the assumption of independence of irrelevant alternatives. Although restrictive, this assumption allows the model to be estimated using a sub-sample of alternatives (see below), greatly improving the model's tractability.

Here $\hat{I}$ is the predicted mean of income, ψ is an idiosyncratic error term, and ρ_j^* the estimated housing price index.

Migration costs are included in the indirect utility function, rather than the budget constraint, to capture the long-term psychological costs of moving. Following Timmins (2007) and Bayer et al. (2009), migration costs are parameterized using a series of dummy variables. These variables indicate the location of MSAs relative to an individual's place of birth. Equation 2.13 presents the migration cost specification:

$$M_{ij} = \beta_{MS} M_{ij}^{S} + \beta_{MD} M_{ij}^{D} + \beta_{MR} M_{ij}^{R} \qquad (2.13)$$

The variable M_{ij}^S=1 if MSA j is located outside the state where individual i was born (0 otherwise), M_{ij}^D=1 if MSA j is located outside the census division where individual i was born (0 otherwise), M_{ij}^R=1 if MSA j is located outside the census region where individual i was born (0 otherwise). Parameter values β_{MS}, β_{MD}, and β_{MR} will be negative, reflecting the propensity to reside close to one's place of birth.

Inclusion of housing prices in the indirect utility function introduces an endogenous relationship into the model. Specifically, there is correlation between housing prices and unobserved local attributes. Due to the nonlinearity inherent in discrete choice models, standard methods for dealing with endogeneity (i.e. instrumental variable techniques) cannot be employed. Berry et al. (1995) develop a method to circumvent the problem wherein all location-specific components of utility (i.e. components that are constant across individuals) are collapsed into ASC. Equation 2.14 defines the ASC.

$$\theta_j = \alpha - \beta_H \ln\left(\rho_j\right) + \beta_X X_j + \xi_j \qquad (2.14)$$

The ASC enter the conditional logit model as a vector of dummy variables, denoted by θ_j. Estimated coefficients on these variables are interpreted as a quality-of-life index (Bayer et al., 2009). These coefficients are then used as dependent variables in a separate linear regression, extracting parameter values for location-specific attributes. Since the regression is linear, endogeneity is resolved using standard methods, or as discussed below, by incorporating housing price into the dependent variable. Substituting Equations 2.11 through 2.14 into Equation 2.8 yields the final indirect utility function, presented in Equation 2.15.

$$\begin{aligned}
&\ln(V_{ij}) = \beta_I \ln(\hat{I}_{ij}) + \beta_{MS} M_{ij}^S + \beta_{MD} M_{ij}^D + \beta_{MR} M_{ij}^R + \theta_j + \nu_{ij} \\
&\text{where :} \qquad \theta_j = \alpha - \beta_H \ln(\rho_j^*) + \beta_X X_j + \xi_j \\
&\qquad\qquad\quad \alpha = \beta_c \ln\left(\frac{\beta_c}{\beta_I}\right) + \beta_H \ln\left(\frac{\beta_H}{\beta_I}\right) \\
&\qquad\qquad\quad \nu_{ij} = \beta_I \psi_{ij} + \eta_{ij} \\
&\qquad\qquad\quad \beta_I = \beta_C + \beta_H
\end{aligned} \tag{2.15}$$

The conditional logit model and corresponding log-likelihood function are presented in Equation 2.16 and 2.17, respectively.[8]

$$P[\ln(V_{ij}) \geq \ln(V_{ik}) \ \ \forall\, j \neq k] = \frac{e^{\beta_I \ln(\hat{I}_{ij}) + \beta_{MS} M_{ij}^S + \beta_{MD} M_{ij}^D + \beta_{MR} M_{ij}^R + \theta_j}}{\sum_{k=1}^{j} e^{\beta_I \ln(\hat{I}_{ij}) + \beta_{MS} M_{ij}^S + \beta_{MD} M_{ij}^D + \beta_{MR} M_{ij}^R + \theta_j}} \tag{2.16}$$

$$\ln(L) = \sum_i \sum_j \Omega_{ij} \ln(P_{ij}) \tag{2.17}$$

Here Ω is an indicator variable that equals 1 if individual i chooses to reside in location j, and 0 otherwise. With a large choice set, estimation of the conditional logit model is

[8] When estimating the conditional logit model it is necessary to designate one MSA as a base, omitting it from the model. Madera-Chowchilla, California is omitted.

computationally burdensome. To make the analysis tractable a random sample of 50,000 individuals is selected. The choice set of each individual is limited to their place of residence and a subset of 19 randomly selected alternatives. McFadden (1978) demonstrates that random sampling of alternatives produces consistent estimates when the uniform conditioning property is satisfied.[9]

During the final stage of estimation, the vector of ASC-coefficients is regressed against location-specific attributes.[10] Following Bayer et al. (2009), endogeneity is resolved by incorporating housing prices into the dependent variable. This variable, depicted in Equation 2.18, is interpreted as the housing-price-adjusted quality-of-life index (Bayer et al., 2009).

$$\theta_j + \beta_H \ln(\rho_j) = \alpha + \beta_X X_j + \varepsilon_j \tag{2.18}$$

Use of this method requires an estimate for β_H. Recall from Equation 2.7 that optimal consumption of housing is a constant proportion of income. Rearranging the demand function for housing yields an expression for β_H:

$$\beta_H = \beta_I \frac{\rho_j H_i}{I_{ij}}$$

---

[9] The uniform conditioning property holds that the probability of selecting a subset is independent of the chosen alternative (McFadden, 1978).

[10] In order to obtain consistent estimates of location-specific attributes the vector of ASC-coefficients must be true values (i.e. without estimation error). Berry et al. (2004) demonstrate that estimates approach their true value as the number of individuals increase relative to the number of alternatives.

Thus β_H is a function of β_I, which is obtained from the conditional logit model, and the share of income spent on housing ($\rho_j H_i / I_{ij}$). This share is set to 0.21, the median value of the sample based on survey data. Having constructed the dependent variable, linear regression techniques are used to recover parameters β_X and to calculate marginal WTP values.

2.4.2 Housing Price Estimates

The index for the price of housing services, used in the conditional logit model, is constructed using a hedonic housing model with location fixed-effects. To accommodate a variety of housing structure and tenure types the dependent variable is defined in terms of user cost. Equation 2.19 presents the regression equation:

$$\ln(UC_{ij}) = \delta_0 + \ln(\rho_j) + \delta_D D + \omega_{ij} \tag{2.19}$$

where UC_{ij} is user cost, ρ_j is a vector of MSA fixed-effects, D is a vector of dwelling characteristics, and ω is an idiosyncratic error term. Coefficients obtained from the MSA fixed-effects reflect the value of housing services relative to other locations, after controlling for dwelling characteristics.[11] User costs are defined as the sum of monthly mortgage or rent payments, utilities fees, property taxes, and insurance. Following Sinha (2008), monthly mortgage payments are calculated using the amortization formula given in Equation 2.20.[12]

---

[11] Coefficients obtained from the MSA fixed-effects are estimated values and should be adjusted to reflect variation in standard errors. This is accomplished using the errors-in-variables correction employed by Gawande and Bandyopadhyay (2000).

[12] All prices are standardized to $2009 using inflation-factors reported in the American Community Survey.

$$Monthly\ Payment_{ij} = \frac{K\left(\frac{r}{12}\right)}{1-\left(1+\left(\frac{r}{12}\right)\right)^{-n}} \quad (2.20)$$

In this formula, K is the self-reported value of the home, r the annual interest rate, and n the number of periods. The annual interest rate is obtained from Freddie Mac's Primary Mortgage Market Survey. In particular, r is set to 6.29%, the average value for 30-year fixed-rate mortgages between 2000 and 2009.

2.4.3 Wage Rate Estimates

Separate hedonic wage regressions are calculated for each MSA. Equation 2.21 depicts the regression model:

$$\ln(W_{ij}) = \sigma_0 + \sigma_S S_{ij} + \sigma_P P(R_B, R_D \mid SC) + \sigma_P P(R_B, R_D \mid SC)^2 + \zeta_{ij} \quad (2.21)$$

where W denotes the hourly wage rate, S a vector of socioeconomic characteristics, and ζ an idiosyncratic error term. Use of MSA-specific regressions allows the effects of socioeconomic characteristics to vary across location. Following Bayer et al. (2009) the term $P(R_B, R_D|SC)$ is used to correct for non-random sorting across locations. This term is defined as the probability that an individual born in census region R_B will reside in region R_D, given the individual's social classification (SC). Probabilities are calculated using observed location choices for ten social classifications, which are based on educational attainment (i.e. no high school degree, high school degree, some college education, college degree, or graduate degree) and marital status (i.e. married or not married). Calculated probabilities are presented in Appendix A. A detailed description of this methodology is available in Dahl (2002).

Regression models are used to predict the wage rate individuals would receive at each location. Predicted wages are, in turn, used to calculate annual incomes for the conditional logit model. More specifically, incomes are calculated from predicted wage rates and self-reported work schedules.[13] Individuals are assumed to work the same number hours at each location, thus ignoring potential tradeoffs between labor and leisure.

2.5 Data

Household data for the user cost regression (Equation 2.19), MSA-specific wage regressions (Equation 2.21), and conditional logit model (Equation 2.16) are obtained from the 2005-2009 American Community Survey (ACS) PUMS. The ACS is an ongoing national survey conducted by the U.S. Census Bureau. The survey is sent to nearly 3 million households per year and contains questions regarding demographic, socioeconomic and housing characteristics (U.S. Census Bureau, 2008). Information obtained from the ACS, including Public Use Microdata Area (PUMA) geographic identifiers, is released in PUMS datasets. PUMA are geographic regions defined by the U.S. Census Bureau, each of which has a minimum population of 100,000 (U.S. Census Bureau, 2008).

With respect to user costs, data is restricted to owner-occupied residences (e.g. non-institutional, non-rented single family homes and apartments) located within a

---

[13] $P(R_B, R_D|SC)$ and $P(R_B, R_D|SC)^2$ are included in the regression model in order to produce unbiased parameter estimates. These variables are omitted during the calculation of predicted wage rates.

MSA.[14] User costs for the sample range from $280 to $7,252, with an average of $2,386 per month. Dwelling characteristics include a series of dummy variables indicating housing structure: detached single-family home (UNITS1), attached single-family home (UNITS2), apartment with 2-4 units (UNITS3), apartment with 5-19 units (UNITS4), or apartment with more than 19 units (UNITS5). Mobile homes are the base category. Characteristics also include the number of rooms (ROOMS), the number of bedrooms (BEDROOMS), lot size (ACRE>1), and dummy variables indicating whether the home contains incomplete kitchen facilities (NOKITCH), incomplete plumbing facilities (NOPLUMB), or an incomplete heating system (NOHEAT). Summary statistics for these characteristics are presented in Table 2.2. Single-family homes, both attached and detached, comprise 75% of the sample; apartments account for the remaining 25%. On average, homes have 6.0 rooms and 2.9 bedrooms. Approximately 13% of homes sit on lots larger than one acre, while a small percentage do not possess a heating system (0.5%), complete kitchen facilities (0.3%), or complete plumbing facilities (0.3%).

Data for wage regressions are restricted to U.S.-born household heads, between 20 and 70 years of age, who are living within a MSA. Non-civilians, self-employed workers and those employed in the agricultural, farming, fishing, or forestry sectors are excluded from the analysis. In addition, to ensure the accuracy of predicted wage rates the analysis is limited to full-time workers. Full-time employment is defined as working 30-60 hours per week for 40-52 weeks per year. Hourly wage rates for the sample range

[14] The 2005-2009 ACS indicates household location by PUMA. Using GIS, PUMAs are matched to MSAs. A PUMA is assigned to a MSA if at least 75% of its total land area falls within the MSA's boundaries.

from \$4.71 to \$143.57, with an average of \$27.17. Socioeconomic characteristics include the individual's age (AGE) and dummy variables indicating gender (FEMALE), marital status (MARRIED), and race (WHITE). Characteristics also include dummy variables indicating educational attainment and employment sector. Educational attainment variables are: no high school degree (NOHS_DEG), some college education (SOME_COLL), college degree (COLL_DEG), and graduate degree (GRAD_DEG). Individuals with a high school degree form the base category. Employment sector variables are: service (OCC1), sales and office administration (OCC2), construction and repair (OCC3), and manufacturing (OCC4). Those employed as managers or professionals form the base category. Summary statistics for these characteristics are presented in Table 2.3. Of those included in the sample, approximately 40% are female, 59% are married and 80% are white. The average household head is 43.9 years of age. In terms of education, 0.1% do not have a high school degree, 26% have a high school degree, 32% have some college education, 26% posses a college degree, and 16% posses a graduate degree. Likewise, 10% are employed in the service sector, 24% in sales or office administration, 8% in construction, 11% in manufacturing, and 47% in management or as professionals.

Data used to estimate the conditional logit model are a subset of those used for wage regressions. First, the dataset is limited to household heads that moved to their current house or apartment within the last nine years. These individuals are more likely to be in equilibrium (i.e. $E[V(I_{ij}, \rho_j, X_j, M_{ij})]=E[\overline{V}]$) and will have made their location decisions based on recent market conditions and amenity levels. Second, to improve computational tractability, conditional logit estimates are made using a random sample of

50,000 household heads. Projected annual income ranges from $6,717 to $207,937, with an average of $52,169. With regards to migration, approximately 43% of the sample resides outside of their birth state, 33% outside their birth division, and 26% outside their birth region. Summary statistics for variables included in the conditional logit model are provided in Table 2.4.

Location-specific attributes, which are regressed on the vector of ASC-coefficients, are obtained from numerous sources. Descriptions, summary statistics, and data sources for these variables are provided in Table 2.5. Data on population (lnPOP), unemployment (UNEMP), local per capita taxes (PCTAX), violent crime (VCRIME) and voter participation (VOTERS) are obtained from the U.S. Census Bureau's CCDB. County-level data is aggregated into MSAs using population-weighted averages. Information health services (PHYSICIANS) and entertainment establishments (ARTINDEX) are obtained from the U.S. Census Bureau's CBP. The CBP contains county-level information on the total number of employees and business establishments by NAICS classification. The variable ARTINDEX is constructed using a principle component analysis and information on the per capita number of restaurants, sports venues, performing arts groups, museums, and historical sites.[15] Data regarding public transportation and education are acquired from the Places Rated Almanac (PRA) and the Core of Common Data (CCD), respectively. The PRA reports scores and rankings for MSAs in nine separate categories (e.g. transportation, education, recreation, arts,

[15] See Filmer and Pritchett (2001) for a detailed discussion of the principal component analysis.

healthcare, crime, cost of living, climate and employment).[16] For the purposes of this analysis the PRA is used to identify MSAs with subway or light-rail systems (SUBWAY). The CCD is maintained by the U.S. Department of Education and contains information regarding the demographic (e.g. total enrollment and racial composition), fiscal (e.g. annual revenues and expenditures) and employment (e.g. total number of teachers, administrators and support staff) characteristics of public school districts. CCD data is used to calculate average dropout rates (DROPOUT) and student-teacher ratios (TEACHERS).

Climate variables are acquired from the NCDC, which is a division of the National Oceanic and Atmospheric Administration (NOAA). It contains information on temperature and precipitation climate normals (1981-2010) from 7500 U.S. weather stations.[17] Using GIS, weather stations are matched with MSAs. Once matched, summer temperature (TEMP) and precipitation (PRECIP) climate normals are calculated for each

---

[16] The PRA does not describe how scores and rankings are calculated. Bayer et al. (2009) use PRA rankings to control for variation in the quality of healthcare, entertainment, and transportation. In contrast, this analysis captures similar variation using proxy variables. Although less comprehensive than the PRA rankings, which are based on multiple factors, these proxy variables have clear definitions and interpretations.

[17] A climate normal is defined as the arithmetic average of a climate variable over a 30-year interval (NOAA, 2011).

location.[18] Among MSAs included in this analysis, summer temperatures range from 61 to 89 ºF, with an average of 74 ºF. Likewise, the range of annual precipitation is 4.5 to 64.8 inches. Average annual precipitation is 38.9 inches.

Emissions data are retrieved from the EPA's Air Quality System database. The database contains county-level information on total emissions of hazardous air pollutants in 2002.[19] County-values are aggregated into MSAs and scaled by total population to obtain per capita emissions (EMISSIONS). Per capita, as opposed to total, emissions are used in order to avoid collinearity problems between pollution and population. Average per capita emissions for the study sample is 31 lbs per year. The EPA also provides locations of National Priority List (NPL) sites. The NPL catalogs hazardous waste sites eligible for federally funded cleanup. As of 2008 there were 1247 sites on the NPL. For this analysis, the number of NPL sites (NPLSITES) located within each MSA is determined using GIS. In addition, GIS is used to calculate local park area (PARKS) and determine whether an MSA is adjacent to an ocean or Great Lake (OCEAN). Park area calculations are based on the StreetMap Pro atlas produced by the Environmental Systems Research Institute (ESRI).

Hazard data is obtained from the GRDP. The GRDP is developed and maintained by the United Nations Environment Programme and the UNISDR, in conjunction with

---

[18] Residential location decisions are driven by a variety of climate amenities. An important extension of this analysis would be to include additional, or alternate, climate measures (e.g. humidity, heating degree days, cooling degree days, or drought indices).

[19] Total emissions values are the sum of the 188 substances designated as hazardous air pollutants in the 1990 Clean Air Act.

numerous partner organizations. It provides spatial information on past natural hazard events and estimates of hazard frequency, hazard exposure and hazard risk. Information is available for various hazard types: earthquakes, cyclones, floods, landslides, tsunamis and volcanic eruptions. Dataset resolution, while inadequate for local-area planning, is sufficient to capture hazard variation across MSAs. The hazard variable for this analysis is constructed using hazard frequency estimates for earthquakes, hurricanes and floods.[20] Frequency is defined as the expected number of hazard events per 1000 years (HRISK). Specific to earthquakes, frequency is the expected number of earthquakes classified as 5 (strong) or greater on the Modified Mercalli Intensity scale per 1000 years. Specific to hurricanes, frequency is the expected number of hurricanes categorized as 3 (major) or greater on the Saffir-Simpson scale per 1000 years. Both measures pertain to major hazard events that would likely damage housing structures. There are no intensity scales associated with flooding data. Consequently, flooding frequency is the expected number of events per 1000 years. Hazard variables are calculated, using GIS, as the spatially-weighted average of the expected number of events within MSA boundaries. Expected hazard frequency ranges from 0 to 193.75 events, with an average of 10.99 events. Figure 2.1 displays estimated hazard frequency by MSA. Hazard frequency is highest along the West and Gulf Coasts, where earthquake and hurricane risk predominate.

[20] The GRDP does not contain frequency data for tornadoes, wildfires, droughts, and blizzards. Exclusion of these hazards may bias parameter estimates of the decomposition analysis (i.e. omitted variable bias). An important extension of this analysis would be to include risk measures for the omitted hazards.

2.6 Results

Results for the user cost regression, MSA-specific wage regressions, conditional logit model and ASC decomposition are discussed separately.[21] In general, results are consistent with theoretical expectations and empirical estimates from previous studies.

2.6.1 User Cost Regression

Results from the user cost regression, not including MSA fixed-effects, are presented in Table 2.6. User cost is positively related to property size (ACRE>1), dwelling size (ROOMS and BEDROOMS), and construction year (YBL1-YBL8). Compared to mobile homes, user costs are higher for single-family dwellings and apartments (UINITS1-UNITS5). User cost is lower for dwellings without heating systems (NOHEAT) or complete kitchen facilities (NOKITCH). Coefficients on MSA fixed-effects form a price index for housing services. The 25 most and least expensive MSAs, as specified by this index, are listed in Table 2.7. These ranking are comparable to the cost of living index compiled by the American Chamber of Commerce Researchers Association. A complete list of the fixed-effect coefficients, and corresponding MSA rankings, are reported in Appendix B.

2.6.2 MSA-Specific Wage Regression

Table 2.8 displays average parameter values for the MSA-specific wage regressions. The sample size, which varies across MSAs, ranges from 547 to 99,324. The average sample size is 5,405. Coefficients are significant at standard reference levels 88% of the time and the average R^2 value is 0.37. Regression results exhibit earnings patterns consistent with human capital theory. In particular, there is a positive correlation between

[21] The corresponding Stata and R codes are provided in Appendix I.

hourly wage rates and educational attainment. Individuals without a high school degree (NOHS_DEG) earn less than those with a degree. Likewise, individuals with some college education (SOME_COLL), a college degree (COLL_DEG) or a graduate degree (GRAD_DEG) earn more than those with a high school degree. With respect to demographic characteristics, wage rates are positively correlated with age (AGE), being married (MARRIED), and being white (WHITE). They are negatively correlated with being female (FEMALE). The quadratic relationship between wage rates and age imply that more experienced workers earn higher wages and that the gains from experience occur primarily at the beginning of an individual's career. With respect to occupations, those employed in the service (OCC1), sales and administration (OCC2), construction (OCC3), and manufacturing (OCC4) sectors earn less than managers and professionals. Finally, the probability terms PMIG and PMIG_SQ, which control for non-random sorting across MSAs, are significant in several MSA-specific regressions.

2.6.3 Conditional Logit Regression

Table 2.9 presents results for several specifications of the conditional logit model. Under Specification 1, residential location choice is modeled as a function of predicted income and location-specific attributes. As previously discussed, the dataset is limited to household heads who have recently moved into their house or apartment and the choice set is limited to 19 randomly selected alternatives. Specification 2, the preferred model, incorporates the migration cost variables. Specification 3 examines the model's sensitivity to the randomly selected set of alternatives. It uses only 9 choice alternatives. Specification 4 examines whether the exclusion of non-movers (i.e. household heads who have not recently moved) drastically alters parameter estimates. The dataset utilized in this specification includes both movers and non-movers.

The probability of residing in a given MSA increases with predicted income and decreases with migration costs. The coefficient on income (lnI), interpreted as the marginal utility of income, is positive and significant in all specifications. Likewise, coefficients on the migration cost variables (Mig1, Mig2, and Mig3) are negative and significant in all specifications. The magnitude of these coefficients decreases with distance from the individual's place of birth—implying the marginal costs of migration are decreasing. This is consistent with the notion that migration costs are primarily physiological. Looking across the model specifications, neither the number of choice alternatives nor the inclusion of non-movers greatly affects estimates of migration cost. However, there is a marked increase in the marginal utility of income when non-movers are included in the analysis.

The conditional logit model also estimates a vector of 296 ASC-coefficients, which are interpreted as a quality-of-life index. The 25 highest and lowest ranking MSAs, as specified by this index, are listed in Table 2.10. These ranking are comparable to those published in the PRA and those produced by Bayer et al. (2009). A complete list of the ASC-coefficients, and corresponding MSA rankings, are reported in Appendix B.

2.6.4 ASC Decomposition

Results for the ASC decomposition are reported in Table 2.11. Model 1 regresses the housing-price-adjusted quality-of-life index against socioeconomic and environmental variables. MSA population (lnPOP) and census region dummy variables are incorporated in Models 2 and 3, respectively. MSA population is included in the model to capture the agglomeration effects, or, conversely, congestion effects of urban settings (Timmins, 2007; Bayer et al., 2009). Importantly, inclusion of MSA population introduces another source of endogeneity. To see this, recall that the sorting model utilizes residential location to reveal consumer preferences. The assumption is that MSAs with a large number of residents offer a better quality of life. Consequently, the quality-of-life index is itself a measure of population. Model 4 corrects for this endogeneity using two-stage least squares (2SLS) regression techniques. Both spatial and temporal lags are used as instruments for lnPOP. The spatial-lag variable is the average population density of each MSA's two nearest neighbors. The temporal-lag is each MSA's population in 1910.[22] Glaeser and Gottlieb (2008) use a similar temporal lag as an instrument for population.

Results are largely consistent across model specifications. With respect to socioeconomic characteristics, quality of life is positively correlated with population (lnPOP), healthcare services (PHYSICIANS), and entertainment resources (ARTINDEX). Quality of life is negatively correlated with the unemployment rate (UNEMP), per capita tax rate (PCTAX), violent crime rate (VCRIME), and student-teacher ratio (TEACHERS). With respect to environmental amenities, quality of life is positively correlated with temperature (TEMP) and annual precipitation (PRECIP); it is negatively correlated with ocean proximity (OCEAN), per capita emissions (EMISSIONS), the number of hazardous waste sites (NPLSITES), and natural hazard

[22] Historic populations are calculated using county-level population data from the 1910 U.S. Census.

risk (HRISK).[23] Notably, HRISK is highly skewed. This raises concerns that outliers unduly influence parameter estimates of hazard risk. Appendix C presents three alternate model specifications that address this issue. In each case there is a negative and significant relationship between quality-of-life and hazard risk.

WTP values are estimated using the formula given in Equation 2.9. The Krinsky-Robb method is used to obtain distributions, and 95% confidence intervals, for the ratio β_X/β_I (Krinsky and Robb, 1986). In particular, distributions are obtained using 5000 random drawings from a multivariate normal distribution; parameters for the distribution are based on the coefficients and variance-covariance matrix estimated in the regression model. Distribution and confidence intervals for TEMP, PRECIP, EMISSIONS, and NPLSITES are presented in Figure 2.2. The distribution and confidence interval for HRISK is presented in Figure 2.3. All distributions are normal, or approaching normal, and are significant at 95%. These distributions are then multiplied by various income levels to obtain a distribution of WTP values. Figures 2.4 (amenities) and 2.5 (disamenities) present WTP for individuals at the 25th, 50th, and 75th percentile of the income distribution. At the 50th percentile, households are willing to pay \$759 for a marginal increase in temperature and \$383 for an increase in precipitation. Likewise, marginal WTP for reductions in emissions per capita (i.e. reducing emissions by one

[23] An interesting extension of these results is to evaluate the relative contribution of each location-specific attribute to quality of life. Fields (2004) develops a method for calculating the relative importance of different explanatory variables. This method is applied to the ASC decomposition. Results are presented in Appendix D.

pound per resident), NPL sites, and the expected number of hazards per 1000 years are $102, $213, $275, respectively.

2.7 Discussion and Conclusion

When selecting a residential location, households trade off wages, prices, and location-specific attributes. Residential location models have been used to quantify these tradeoffs. In particular, these models have been used to recover preferences for school quality, racial composition, and green space within localized-markets. They have also been used to recover preferences for air quality and climate conditions across markets at a national level. This study extends these models, evaluating preferences for disaster risk-reduction. Results indicate, in line with hedonic-property literature, that households consider low-probability high-consequence events when making location decisions. This implies the widespread use of self-protection as a means of reducing hazard risk. It also highlights the need for estimating the value of risk reduction using non-market valuation techniques.

Previous studies have estimated non-market values of risk mitigation within localized housing markets. Marginal WTP values obtained from these studies, which range from 4% to 19% of housing price, primarily reflect the value placed on avoiding property damage. In contrast, by comparing preferences across MSAs it is possible to estimate the value placed on avoiding the broader consequences of natural disaster, such as the disruption of service flows from community and environmental amenities. Since this value pertains to self-protection measures, and does not include the cost of market insurance or self-insurance, is can be regarded as a lower bound.

Results from this study indicate an annual WTP of $275 for a marginal reduction in the expected number of earthquake, hurricane, and flood events per 1000 years.[24] From a policy perspective, this figure provides a point of departure for valuing the benefits of risk-mitigating interventions—and subsequently for conducting benefit-cost analyses. Consider a MSA that recently experienced severe flooding: Nashville-Davidson-Murfreesboro-Franklin, Tennessee. There are approximately 612,000 households in this MSA. Aggregating the value of risk mitigation across households, and adjusting for the MSA's median income, yields an estimated WTP of $172.7 million per year. Once aggregated, the WTP value can be compared to intervention costs. Arguably, interventions enacted at a municipal level will not lower the expected number of hazard events, particularly with regards to earthquakes and hurricanes. Nonetheless, WTP obtained from this study provide a useful indication of the value placed on risk reduction.

This study is also relevant to the valuation of hazard-related components of global climate change. Among other things, climate change is expected to alter the frequency, intensity, and duration of hazard events (Greenough et al., 2001; IPCC, 2011). In particular, a growing body of evidence suggests that climate change will increase the frequency of extreme precipitation events and the intensity of hurricane events

---

[24] For a couple of reasons this value cannot be directly compared to those obtained from localized housing markets. First, to some extent these studies value different components of the utility function (i.e. housing services and location-specific attributes). Second, they use different measures of disaster risk. Hedonic studies typically use a dummy variable to identify homes within high-risk areas (e.g. flood plains or earthquake zones). In contrast, this study defines hazard-risk as the expected number of hazard-events.

(Greenough et al., 2001; van Aalst, 2006; IPCC, 2011). This increase will, in turn, reduce expected utility through higher flood-event and major hurricane probabilities. The WTP values reported in this study can be used to estimate values for preventing these increases.

In addition to hazard risk, this analysis estimates WTP for marginal changes in temperature, precipitation, emissions, and NPL sites. Numerous non-market valuation studies have estimated the value of these amenities and disamenities. However, the values recovered in the analysis are notable for two reasons. First, this analysis explicitly controls for the cost of migration. Bayer et al. (2009) argue that WTP estimates are subject to considerable omitted-variable bias when migration costs are ignored. Second, use of disaster risk as an independent variable controls for another source of omitted-variable bias. Bin et al. (2008) note that environmental amenities are often highly correlated with natural hazards. For example, many coastal regions are characterized by desirable climate, proximity to water-related amenities, and high levels of disaster risk. If not accounted for, this correlation will bias estimates of amenity values downward. By way of illustration, Hoehn et al. (1987) estimate a national hedonic housing and wage model that does not control for migration costs or hazard levels. After controlling for socioeconomic and climate characteristics they estimate WTP values for a 10% reduction in Superfund sites. Adjusting their findings to $2009 indicates an annual household WTP of $24. The corresponding value for a 10% reduction in NPL sites, obtained from this analysis, is $65. The difference between these values is consistent with Bayer et al. (2009). They find that WTP to improve air quality is three times greater under the residential sorting model than a nation hedonic model.

Results for this analysis are affected by a number of restrictive assumptions, including use of the Cobb-Douglas utility (as opposed to a more flexible function) and independence of irrelevant alternatives. Despite these restrictions, however, this analysis offers valid information regarding consumer preferences for location-specific attributes. Results confirm that households consider low-probability high-consequence events when making location decisions. Marginal WTP values, which are not subject to omitted variable bias from migration costs, have practical implications for the development of risk-mitigating interventions.

Table 2.1 Migration Patterns Between Census Divisions
[Percent Birth Division by Division of Residence]

		Division of Residence								
		New England	Mid-Atlantic	East North Central	West North Central	South Atlantic	East South Central	West South Central	Mountain	Pacific
Birth Division	New England	64.79	5.66	2.38	0.7	14.35	0.91	2.24	3.11	5.88
	Mid-Atlantic	3.59	62.24	3.42	0.69	18.16	0.99	2.37	3.16	5.38
	East North Central	0.97	2.13	65.51	2.76	11.03	2.52	3.93	4.91	6.25
	West North Central	0.79	1.51	6.98	55.9	7.33	1.68	7.36	9.12	9.34
	South Atlantic	1.3	4.33	3.53	0.88	77.63	3.14	3.25	2.19	3.76
	East South Central	0.63	1.57	8.68	1.4	16.37	59.79	5.95	2.06	3.55
	West South Central	0.51	1.11	2.63	2.06	6.47	2.44	74.22	4.19	6.37
	Mountain	0.75	1.43	2.8	2.39	5.45	1.02	6.59	63.33	16.23
	Pacific	0.8	1.42	2.23	1.46	5.13	0.97	4.16	8.78	75.05

†Based on 2005-2009 PUMS

Table 2.2 Descriptive Statistics for User Cost Regression

Variable	Description	Units	Source	Mean	Std. Dev.
lnUC	Monthly user cost	ln($2009)	2005-2009 ACS	7.394	0.666
ACRE>1	Property size greater than 1 acre	0/1	2005-2009 ACS	0.130	0.336
UNITS1	Detached single family home	0/1	2005-2009 ACS	0.689	0.463
UNITS2	Attached single family home	0/1	2005-2009 ACS	0.063	0.243
UNITS3	Apartment with 2-4 units	0/1	2005-2009 ACS	0.068	0.252
UNITS4	Apartment with 5-19 units	0/1	2005-2009 ACS	0.080	0.271
UNITS5	Apartment with >19 units	0/1	2005-2009 ACS	0.070	0.256
NOHEAT	Dwelling does not contain heating system	0/1	2005-2009 ACS	0.005	0.073
NOKITCH	Dwelling does not contain complete kitchen facilities	0/1	2005-2009 ACS	0.003	0.056
NOPLUMB	Dwelling does not contain complete plumbing facilities	0/1	2005-2009 ACS	0.003	0.050
ROOMS	Number of rooms	ROOMS	2005-2009 ACS	5.960	1.875
BEDROOMS	Number of bedrooms	ROOMS	2005-2009 ACS	2.860	1.020
YBL1	Dwelling built between 2000-2004	0/1	2005-2009 ACS	0.094	0.292
YBL2	Dwelling built between 1990-1999	0/1	2005-2009 ACS	0.143	0.350
YBL3	Dwelling built between 1980-1989	0/1	2005-2009 ACS	0.139	0.346
YBL4	Dwelling built between 1970-1979	0/1	2005-2009 ACS	0.159	0.366
YBL5	Dwelling built between 1960-1969	0/1	2005-2009 ACS	0.119	0.324
YBL6	Dwelling built between 1950-1959	0/1	2005-2009 ACS	0.125	0.331
YBL7	Dwelling built between 1940-1949	0/1	2005-2009 ACS	0.057	0.232
YBL8	Dwelling built before 1940	0/1	2005-2009 ACS	0.126	0.332

Table 2.3 Descriptive Statistics for MSA-Specific Wage Regressions

Variable	Description	Units	Source	Mean	Std. Dev.
lnW	Hourly wage rate	ln($2009)	2005-2009 ACS	3.130	0.586
AGE	Age of household head	YRS	2005-2009 ACS	43.937	11.342
AGE_SQ	Squared age of household head	100 YRS	2005-2009 ACS	20.591	10.023
FEMALE	Household head is female	0/1	2005-2009 ACS	0.400	0.490
MARRIED	Household head is married	0/1	2005-2009 ACS	0.590	0.492
WHITE	Household head is white	0/1	2005-2009 ACS	0.806	0.395
NOHS_DEG	Highest level of education: no high school degree	0/1	2005-2009 ACS	0.059	0.400
SOME_COLL	Highest level of education: some college education	0/1	2005-2009 ACS	0.318	0.466
COLL_DEG	Highest level of education: college degree	0/1	2005-2009 ACS	0.262	0.440
GRAD_DEG	Highest level of education: graduate degree	0/1	2005-2009 ACS	0.161	0.368
OCC1	Occupation category: service	0/1	2005-2009 ACS	0.101	0.302
OCC2	Occupation category: sales or office administration	0/1	2005-2009 ACS	0.238	0.426
OCC3	Occupation category: construction and repair	0/1	2005-2009 ACS	0.084	0.277
OCC4	Occupation category: manufacturing	0/1	2005-2009 ACS	0.111	0.315
PMIG	Migration probability (Dahl correction)	PROB	2005-2009 ACS	0.441	0.297
PMIG_SQ	Squared migration probability (Dahl correction)	PROB	2005-2009 ACS	0.283	0.247

Table 2.4 Descriptive Statistics for Conditional Logit Analysis

Variable	Description	Units	Source	Mean	Std. Dev.
lnI	Predicted annual income	ln($2009)	2005-2009 ACS	10.672	0.447
MIG1	MSA located outside household head's birth state	0/1	2005-2009 ACS	0.938	0.241
MIG2	MSA located outside household head's birth division	0/1	2005-2009 ACS	0.845	0.362
MIG3	MSA located outside household head's birth region	0/1	2005-2009 ACS	0.711	0.453

Table 2.5 Descriptive Statistics for ASC Decomposition

Variable	Description	Units	Source	Mean	Std. Dev.
ASC	Alternative specific constant adjusted for housing costs	INDEX	NA	2.430	1.112
lnPOP	MSA Population	ln(POPULATION)	2005 CCDB	12.856	1.039
UNEMP	Unemployment rate	UNEMP/LF	2005 CCDB	0.052	0.016
PCTAX	Per capita tax rate	TH$/PER	2002 CCDB	1.075	0.368
VCRIME	Violent crime rate	CRM/100TH. PER	2005 CCDB	0.422	0.213
PHYSICIANS	Physicians rate	PHY/TH. PER	2005 CBP	3.024	1.762
ARTINDEX	Arts and entertainment index	INDEX	2005 CBP	0.000	1.314
SUBWAY	MSA maintains subway or light-rail system	0/1	2007 PRA	0.098	0.298
TEACHERS	Student teacher ratio	STU/TEACHER	2005 CCD	16.179	2.431
DROPOUT	High school dropout rate	DROPOUT/ENROLL	2005 CCD	0.037	0.020
VOTERS	Voter participation rate	VOTERS/ELIGIBLE	2004 CCDB	0.569	0.088
TEMP	Climate normal: average temperature	ºF	NCDC	56.351	8.169
PRECIP	Climate normal: average precipitation	INCHES	NCDC	39.585	14.195
OCEAN	MSA adjacent to ocean	0/1	ESRI	0.260	0.439
EMISSIONS	Annual emissions per capita	LBS/PER	2002 EPA	0.031	0.017
NPLSITES	National Priority List sites	SITES	2008 EPA	3.071	8.287
PARKS	Percentage of MSA area designated as local park	PARK AREA/AREA	ESRI	0.004	0.007
HRISK	Expected number of hazard events per 1000 years	EVENTS/TH YRS	GRDP	10.986	25.317
REG1	MSA located in Northeast census region	0/1	USCB	0.257	0.438
REG3	MSA located in South census region	0/1	USCB	0.409	0.492
REG4	MSA located in West census region	0/1	USCB	0.209	0.408

Table 2.6 Results for User Cost Regression

Variable	Coefficient	Robust Std. Err.
Constant	6.2975***	0.0160
ACRE>1	0.1519***	0.0009
UNITS1	0.7166***	0.0021
UNITS2	0.6491***	0.0025
UNITS3	0.7867***	0.0032
UNITS4	0.6719***	0.0034
UNITS5	0.8747***	0.0038
NOHEAT	-0.091***	0.0064
NOKITCH	-0.0544***	0.0107
NOPLUMB	-0.0026	0.0092
ROOMS	0.1063***	0.0003
BEDROOMS	0.0714***	0.0006
YBL1	-0.0169***	0.0018
YBL2	-0.0862***	0.0017
YBL3	-0.1872***	0.0017
YBL4	-0.2773***	0.0017
YBL5	-0.3061***	0.0018
YBL6	-0.3305***	0.0018
YBL7	-0.3563***	0.0021
YBL8	-0.3223***	0.0019

* p<0.1 ** p<0.05 *** p<0.01
N=1,599,627
R^2=0.6047

Table 2.7 Cost of Living Rankings for 25 Most and Least Expensive MSAs

Most Expensive MSA		Least Expensive MSA	
Rank	MSA	Rank	MSA
1	Santa Cruz-Watsonville, CA	272	Anderson, IN
2	San Jose-Sunnyvale-Santa Clara, CA	273	Alexandria, LA
3	San Francisco-Oakland-Fremont, CA	274	El Paso, TX
4	Salinas, CA	275	Altoona, PA
5	Napa, CA	276	Charleston, WV
6	Los Angeles-Long Beach-Santa Ana, CA	277	Clarksville, TN-KY
7	Santa Barbara-Santa Maria-Goleta, CA	278	Fort Wayne, IN
8	Oxnard-Thousand Oaks-Ventura, CA	279	Muncie, IN
9	Santa Rosa-Petaluma, CA	280	Kingsport-Bristol-Bristol, TN-VA
10	San Luis Obispo-Paso Robles, CA	281	Albany, GA
11	San Diego-Carlsbad-San Marcos, CA	282	Owensboro, KY
12	New York-Northern New Jersey-Long Island, NY-NJ-PA	283	Jackson, TN
13	Bridgeport-Stamford-Norwalk, CT	284	Brownsville-Harlingen, TX
14	Boston-Cambridge-Quincy, MA-NH	285	Sioux City, IA-NE-SD
15	Barnstable Town, MA	286	Decatur, AL
16	Washington-Arlington-Alexandria, DC-VA-MD-WV	287	Danville, VA
17	Vallejo-Fairfield, CA	288	Anniston-Oxford, AL
18	Naples-Marco Island, FL	289	Gadsden, AL
19	Ocean City, NJ	290	Sumter, SC
20	Sacramento--Arden-Arcade--Roseville, CA	291	Fort Smith, AR-OK
21	Trenton-Ewing, NJ	292	Terre Haute, IN
22	Seattle-Tacoma-Bellevue, WA	293	Florence-Muscle Shoals, AL
23	Stockton, CA	294	McAllen-Edinburg-Mission, TX
24	Miami-Fort Lauderdale-Pompano Beach, FL	295	Joplin, MO
25	Providence-New Bedford-Fall River, RI-MA	296	Johnstown, PA

Table 2.8 Results for MSA-Specific Wage Regressions
[Average Parameter Values]

Variable	Coefficient	Robust Std. Err.
Constant	1.5327	0.1448
AGE	0.0598	0.0064
AGE_SQ	-0.0596	0.0075
FEMALE	-0.2336	0.0228
MARRIED	0.1076	0.0217
WHITE	0.0894	0.0375
NOHS_DEG	-0.1637	0.0472
SOME_COLL	0.1421	0.0328
COLL_DEG	0.3433	0.0476
GRAD_DEG	0.4892	0.0620
OCC1	-0.3586	0.0363
OCC2	-0.2001	0.0292
OCC3	-0.1345	0.0380
OCC4	-0.2328	0.0366
PMIG	-0.4498	0.3871
PMIG_SQ	0.6022	0.5073

N: Mean=5405 Max.=99324 Min.=547
R^2: Mean=0.373 Max.=0.517 Min.=0.238
88.4% of coefficients are significant at $p<0.1$

Table 2.9 Results for Conditional Logit Analysis

	Specification 1: • Movers • w/o Migration Costs • Choice Set: 20		Specification 2: • Movers • w/ Migration Costs • Choice Set: 20		Specification 3: • Movers • w/ Migration Costs • Choice Set: 10		Specification 4: • Movers & Non-Movers • w/ Migration Costs • Choice Set: 20	
Variable	Coefficient	Robust Std. Err.	Coefficient	Robust Std. Err.	Coefficient	Robust Std. Err.	Coefficient	Robust Std. Err.
ln(I)	0.7835***	0.0883	1.1371***	0.1035	1.1641***	0.1147	1.5816***	0.1071
Mig1			-2.8127***	0.0216	-2.8498***	0.0255	-3.0251***	0.0225
Mig2			-0.7015***	0.0250	-0.6984***	0.0276	-0.6885***	0.0268
Mig3			-0.5054***	0.0211	-0.5271***	0.0227	-0.4965***	0.0225
Log-Likelihood	-113833.67		-72356.34		-49055.69		-67388.29	

* p<0.1 ** p<0.05 *** p<0.01
N=50,000

Table 2.10 Quality of Life Rankings for Top and Bottom 25 MSAs

Top MSA		Bottom MSA	
Rank	MSA	Rank	MSA
1	Phoenix-Mesa-Glendale, AZ	272	Saginaw-Saginaw Township North, MI
2	Atlanta-Sandy Springs-Marietta, GA	273	Yuba City, CA
3	Dallas-Fort Worth-Arlington, TX	274	Redding, CA
4	Washington-Arlington-Alexandria, DC-VA-MD-WV	275	Midland, TX
5	Denver-Aurora-Broomfield, CO	276	Atlantic City-Hammonton, NJ
6	Seattle-Tacoma-Bellevue, WA	277	Santa Cruz-Watsonville, CA
7	Los Angeles-Long Beach-Santa Ana, CA	278	Merced, CA
8	Las Vegas-Paradise, NV	279	Kokomo, IN
9	Houston-Sugar Land-Baytown, TX	280	Springfield, OH
10	Tampa-St. Petersburg-Clearwater, FL	281	Houma-Bayou Cane-Thibodaux, LA
11	Chicago-Joliet-Naperville, IL-IN-WI	282	Lebanon, PA
12	New York-Northern New Jersey-Long Island, NY-NJ-PA	283	Barnstable Town, MA
13	Orlando-Kissimmee-Sanford, FL	284	Mansfield, OH
14	Nashville-Davidson--Murfreesboro--Franklin, TN	285	Jackson, MI
15	Portland-Vancouver-Hillsboro, OR-WA	286	Vineland-Millville-Bridgeton, NJ
16	Boston-Cambridge-Quincy, MA-NH	287	Monroe, MI
17	Miami-Fort Lauderdale-Pompano Beach, FL	288	Hanford-Corcoran, CA
18	San Francisco-Oakland-Fremont, CA	289	Napa, CA
19	Charlotte-Gastonia-Rock Hill, NC-SC	290	Odessa, TX
20	Baltimore-Towson, MD	291	Glens Falls, NY
21	Raleigh-Cary, NC	292	Flint, MI
22	Austin-Round Rock-San Marcos, TX	293	Kingston, NY
23	Jacksonville, FL	294	El Centro, CA
24	Virginia Beach-Norfolk-Newport News, VA-NC	295	Madera-Chowchilla, CA
25	Minneapolis-St. Paul-Bloomington, MN-WI	296	Ocean City, NJ

Table 2.11 ASC Decomposition Results

Variable	Model 1: OLS		Model 2: OLS		Model 3: OLS		Model 4: 2SLS	
	Coefficient	Robust Std. Err.	Coefficient	Robust Std. Err.	Coefficient	Robust Std. Err.	Coefficient	Robust Std. Err.
Constant	-3.030**	1.300	-12.023***	0.977	-10.892***	1.024	-10.602***	1.076
lnPOP			0.977***	0.041	1.008***	0.036	0.974***	0.069
UNEMP	-18.802***	3.971	-12.674***	3.699	-9.632***	3.428	-9.821***	3.234
PCTAX	0.335*	0.192	-0.398***	0.120	-0.239**	0.100	-0.215**	0.103
VCRIME	0.550**	0.276	-0.256	0.169	-0.461***	0.137	-0.434***	0.137
PHYSICIANS	0.050	0.034	0.029	0.021	0.024	0.016	0.025	0.015
ARTINDEX	-0.039	0.053	0.100***	0.034	0.073***	0.027	0.069**	0.028
SUBWAY	1.357***	0.207	0.200*	0.119	0.067	0.096	0.108	0.125
TEACHERS	0.133***	0.026	0.056***	0.016	-0.049***	0.018	-0.045**	0.019
DROPOUT	0.301	2.602	2.970	1.809	-0.728	1.675	-0.795	1.601
VOTERS	0.326	0.683	0.064	0.498	0.465	0.485	0.486	0.472
TEMP	0.044***	0.013	0.025***	0.009	0.019*	0.010	0.019**	0.009
PRECIP	0.006	0.004	0.005*	0.003	0.010***	0.003	0.010***	0.003
OCEAN	-0.135	0.126	-0.177**	0.069	-0.125*	0.064	-0.124**	0.062
EMISSIONS	-0.007***	0.002	0.000	0.001	-0.002*	0.001	-0.003**	0.001
NPLSITES	0.014*	0.008	-0.010***	0.002	-0.006***	0.002	-0.005**	0.002
PARKS	12.876	9.998	-1.092	5.133	-4.076	4.384	-3.549	4.489
HRISK	-0.007***	0.002	-0.005***	0.002	-0.007***	0.002	-0.007***	0.002
REG1					-0.450***	0.097	-0.441***	0.099
REG3					0.239**	0.093	0.244***	0.091
REG4					0.940***	0.127	0.936***	0.125
Adjusted R^2	0.495		0.809		0.860		0.860	

* p<0.1 ** p<0.05 *** p<0.01
N=296

Figure 2.1 Expected Number of Hazard Events by MSA

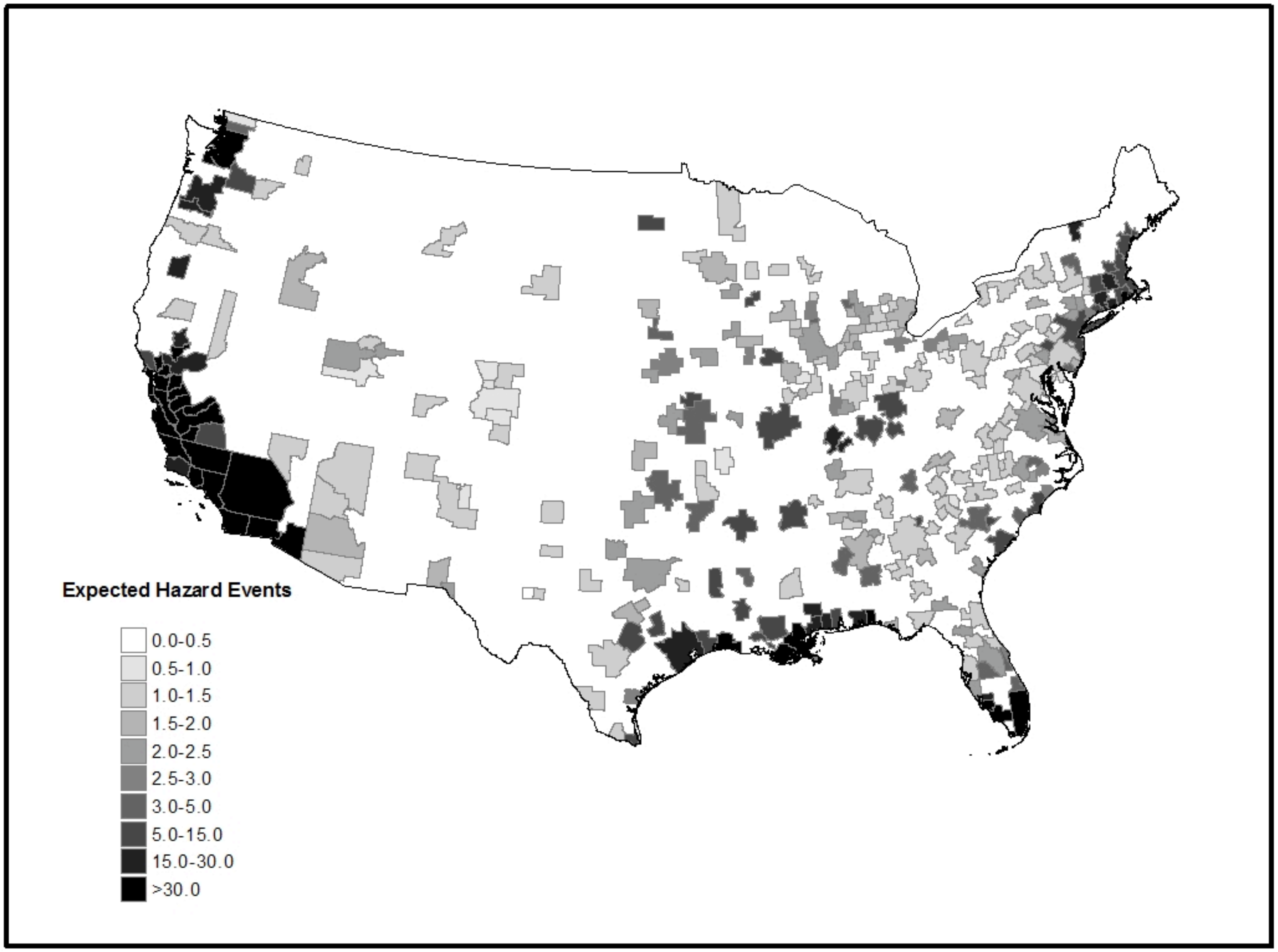


Figure 2.2 Distribution of Simulated WTP for Environmental Amenities

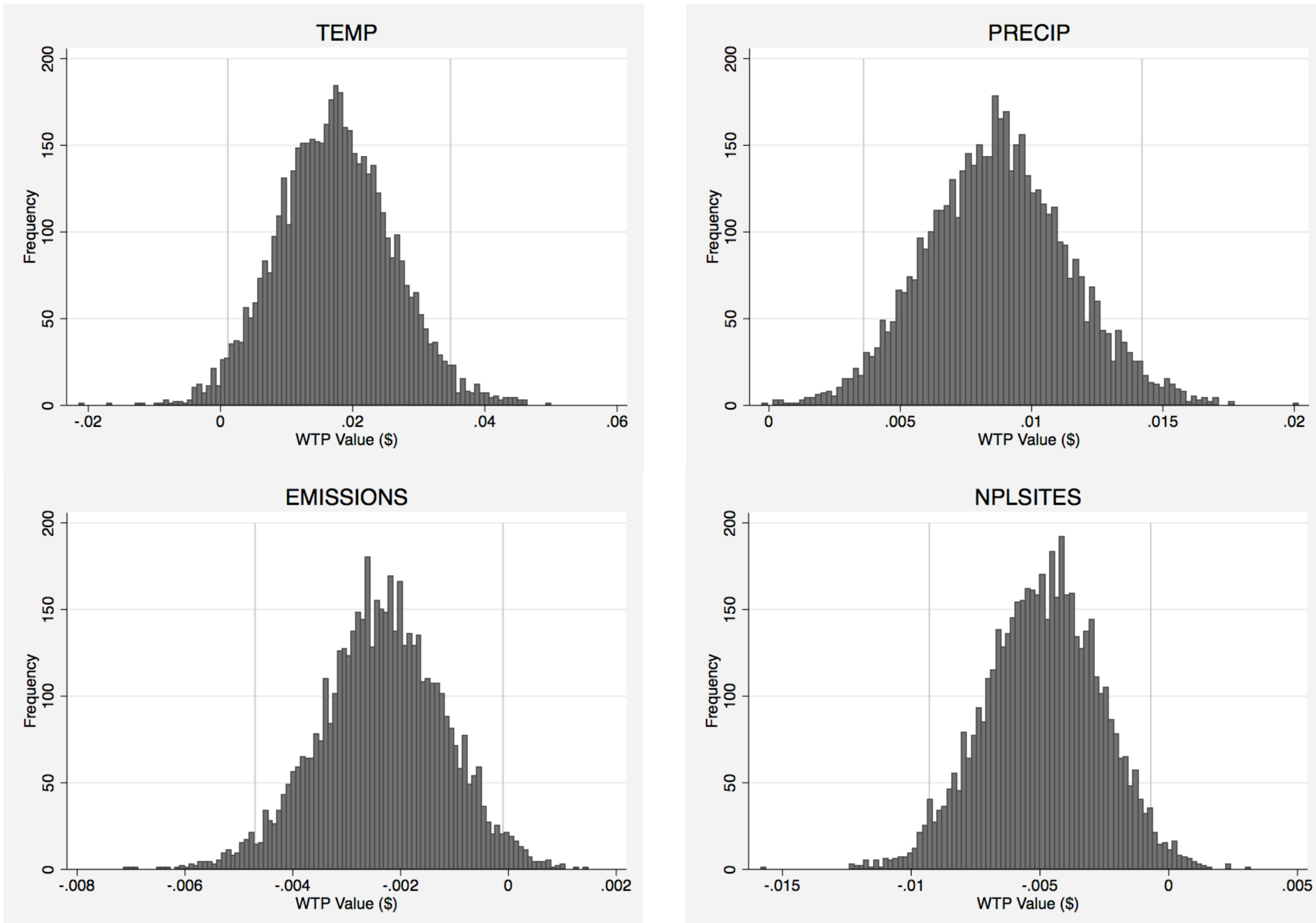


Figure 2.3 Distribution of Simulated WTP Values for Hazard Risk

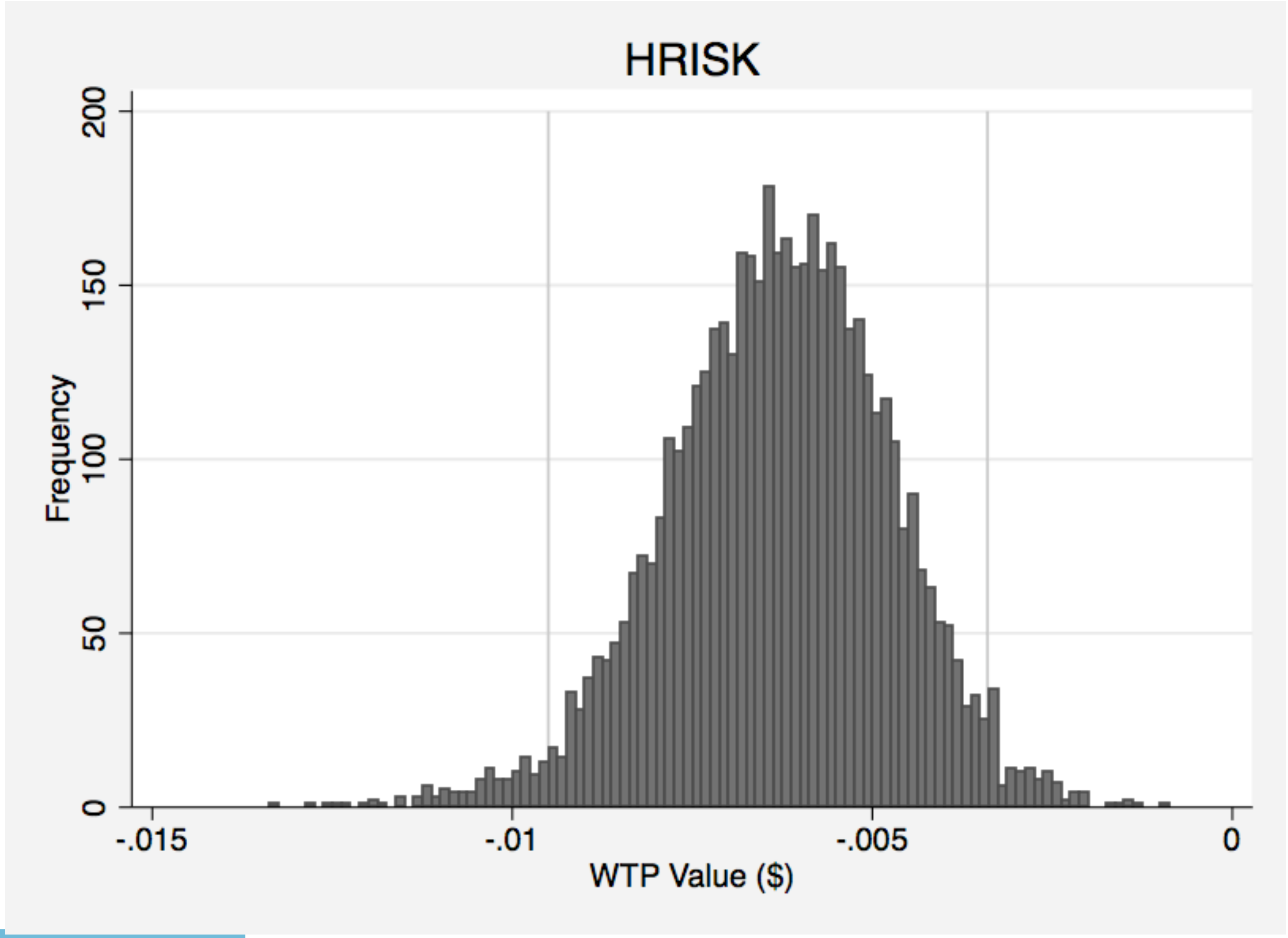


Figure 2.4 WTP at Select Income Percentiles (Amenities)

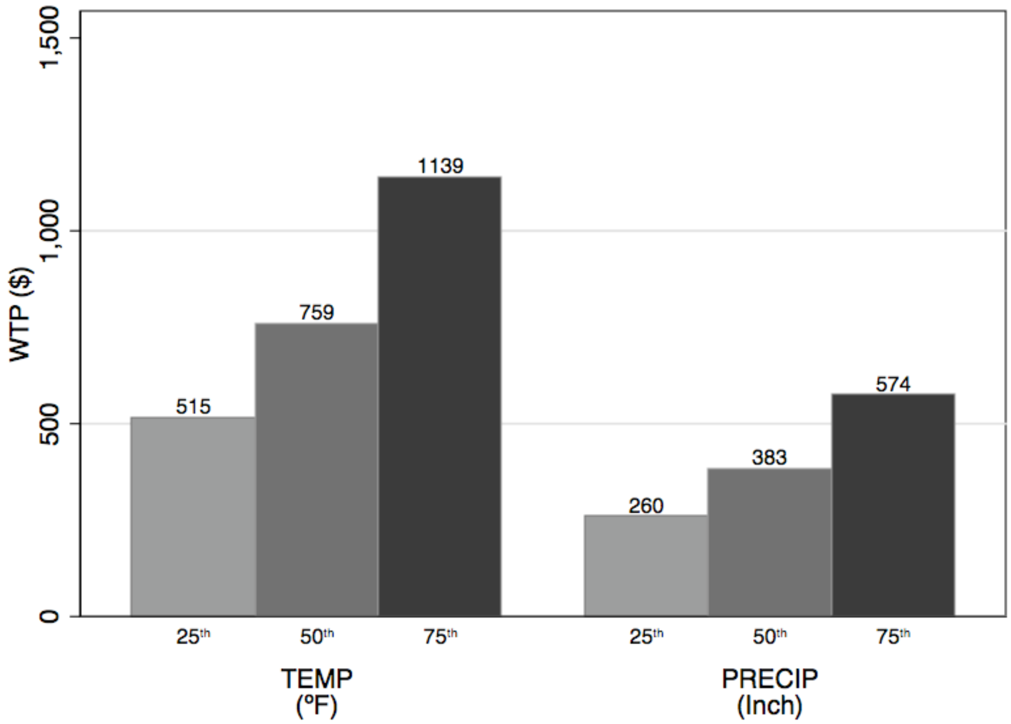
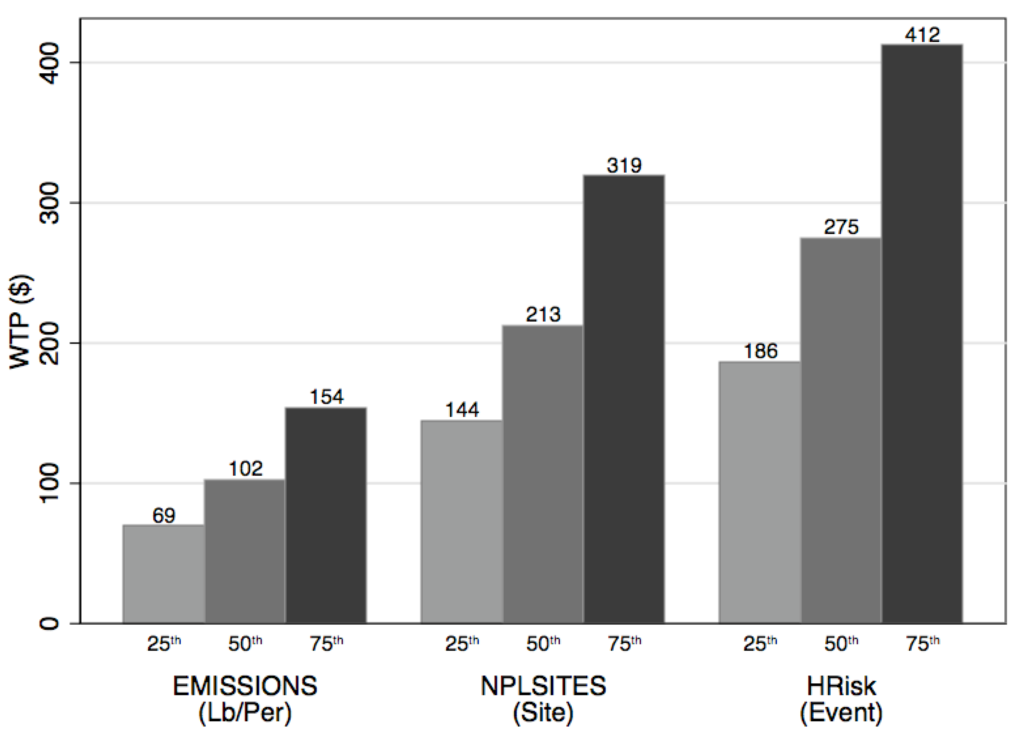


Figure 2.5 WTP at Select Income Percentiles (Disamenities)



Chapter 3: County Migration Patterns and the Risk of Natural Hazards

3.1 Introduction

The spatial distribution of population is affected by fertility, mortality, and migration rates. In the U.S., because fertility and mortality rates are low and fairly homogeneous, distributional changes are largely driven by migration.[25] In recent decades, domestic migration contributed to three distinct and ongoing population shifts (U.S. Census Bureau, 2006). First, migration is shifting population among broad geographic areas. Between 2001 and 2009, the Northeast, Midwest, and Pacific census geographies experienced substantial net out-migration (see Table 3.1). In contrast, the South and Mountain West census geographies experienced net in-migration. Second, migration is shifting population from rural to urban areas (e.g. metropolitan and micropolitan areas). Since 2001, population in rural counties declined by an average of 4.1% annually due to migration; population in urban counties increased by an average of 1.9%.[26] Thirdly, migration is shifting population from central urban areas to outlying urban areas. While

---

[25] Within U.S. counties the annual population change attributed to natural factors (i.e. fertility and mortality) ranges from -1.4 to 2.9% while the change attributed to migration ranges from -9.8 to 8.9 percent (Author's calculations based on U.S. Census Bureau's 2009 Components of Population Change).

[26] Author's calculations using the U.S. Census Bureau's 2000-2009 Components of Population Change.

both central and outlying counties typically experience net in-migration, the average rate of migration-related population growth is three times greater in outlying areas. Identifying and quantifying the determinants of these migration patterns is essential to explaining local area population growth and predicting future population distributions.

Domestic migration decisions are motivated by a number of economic, social, and environmental considerations. The relative importance of these factors has been the subject of debate. Partridge (2010) evaluates the two competing theoretical models that underlie this debate. The New Economic Geography developed by Krugman (1991) posits that migration is primarily determined by agglomeration economics (i.e. households are attracted to employment opportunities and product variety). In contrast, the spatial equilibrium model posits that households select their residential location so as to maximize utility, taking into account economic conditions and local amenities. This model is rooted in the sorting models developed by Tiebout (1956) and Roback (1982). Partridge (2010) concludes that, while empirical testing of the New Economic Geography is in the early stages, there is an overwhelming body of evidence supporting the spatial equilibrium model. In particular, evidence shows that environmental amenities (e.g. climate, recreational opportunities, and health risks) are important predictors of current and historic migration flows.

This analysis extends the amenity migration framework. It models the relationship between county-level net migration rates and earthquake, hurricane, and flood risk, while controlling for economic, demographic, and environmental characteristics. Empirical estimation is complicated by the presence of spatial dependency in migration patterns and preference heterogeneity across geographic locations. These issues are addressed using two separate regression techniques. The SAC method is used to control for spatial autocorrelation across observations—yielding unbiased and consistent parameter estimates. The GWR method is used to evaluate how the relationship between migration and amenities varies across counties.

3.2 Literature Review

Recent studies have evaluated the determinants of internal U.S. migration. Davies et al. (2001) and Poston et al. (2009) analyze inter-state migration patterns. Davies et al. (2001) analyze the effects of economic conditions on migration using state-to-state migration data from the Internal Revenue Service. Results indicate that in-migration increases with per capita income and decreases with unemployment. Moreover, there is a quadratic relationship between migration and distance, such that likelihood of migrating to the destination state decreases with distance from origin state, but at a decreasing rate. Poston et al. (2009) analyze the effects of climate on in-migration, out-migration, and net in-migration between 1995 and 2000, after controlling for economic and demographic characteristics. They use a principal component analysis to generate climate indices for temperature, humidity, and wind speed. Results indicate significant correlations between these indices and the various migration measures, suggesting that climate operates as both a push and pull factor in migration decisions.

Gawande et al. (2000), Rupasingha and Goetz (2004), McGranahan (2008), and Partridge et al. (2008) analyze inter-county migration patterns. Gawande et al. (2000) find a positive correlation between net out-migration and the number of hazardous waste sites, after controlling for socioeconomic characteristics. Inclusion of a cross term between hazardous waste sites and average per capita income allows for estimation of an

Environmental Kuznets Curve. Results indicate that net out-migration increases in response to the presence of hazardous waste sites when average per capita income reaches $16,932 ($1989). Below this threshold, hazardous waste sites have little impact on migration decisions. Rupasingha and Goetz (2004) examine the relationship between net in-migration and health risks. Measures of health risk include estimated cancer risks from air pollutants, the number of superfund sites, and the Hazard Risk Ranking for superfund sites. Findings show a negative correlation between net in-migration and health risk, after controlling for socioeconomic characteristics and spatial correlation. To control for spatial correlation, Rupasingha and Goetz (2004) employ a general spatial model that incorporates both spatial-lag and spatial-error terms. Results show that both spatial parameters are significant, suggesting interdependence between the migration rates of neighboring counties. McGranahan (2008) uses a simultaneous equation model (i.e. equations for migration and employment) to determine preferences for landscape characteristics in non-metropolitan counties. According to the model, migration is a function of employment opportunities, environmental amenities, and socioeconomic characteristics. Likewise, employment is a function of migration, population density, and socioeconomic characteristics. Findings indicate that individuals are attracted to areas with a mix of forest and open land, water, and topographic variation. Partridge et al. (2008) use GWR techniques to evaluate whether the relationship between county growth patterns and local amenities exhibit spatial heterogeneity. They find significant variation with regards to climate, topography, water area, human capital, and various demographic characteristics. This suggests that standard modeling techniques, which estimate global parameter values, mask important regional relationships.

With respect to natural hazards, several studies note the link between disaster occurrence and internal displacement (Saldaña-Zorrialla and Sandberg, 2009; Varano et al., 2010; Dun, 2011). Saldaña-Zorrialla and Sandberg (2009) conduct one of the few empirical analyses of this relationship, employing data from 2,443 municipalities in Mexico. Findings indicate a positive correlation between out-migration and the frequency of hazard occurrence, after controlling for socioeconomic characteristics and spatial correlation. Of those living in municipalities that experienced a high number of disasters, individuals with relatively high education attainment are the most likely to migrate. Finally, numerous studies have analyzed the adverse health and socioeconomic consequences (e.g. crime and unemployment) associated with displacement (Vigdor, 2007; Uscher-Pines, 2008; Davis et al., 2010; Varano et al., 2010).

3.3 Theoretical Considerations

The equilibrium model developed by Roback (1982), and extended to allow for spatial heterogeneity by Partridge et al. (2008), provides a theoretical framework for this analysis. Within this framework, migration results from the utility and profit-maximizing behavior of economic agents. Let $V_i(w_i, r_i, a_i)$ be the indirect utility function for a representative household in location i, where w denotes the wages, r the rents, and a the location-specific amenities. Likewise, let $C_i(w_i, r_i, a_i)$ be cost function for a representative firm. Both functions are assumed to be well-behaved (i.e. they are continuous, convex or concave in accordance with economic theory, and exhibit expected signs on partial derivatives). In equilibrium, utility and profits are constant across locations and neither consumers nor firms have incentive to migrate. Equilibrium conditions are presented in Equations 3.1 and 3.2, where P^* is a common product price.

$$V_i(w_i, r_i, a_i) = V^* \quad (3.1)$$

$$C_i(w_i, r_i, a_i) = P^* \quad (3.2)$$

Migration occurs as a result of market disequilibrium, which is brought about by exogenous shocks (e.g. shifts in demand, new production technologies, a change in location-specific amenities). Specifically, shocks alter product prices, wages, and rents and subsequently induce some utility and profit-maximizing agents to relocate. Migration alters labor and housing markets, through changes in supply and demand, such that the incentive to relocate gradually diminishes. Eventually labor and housing markets reach a new equilibrium.

According to this model, migration flows are driven by interactions between consumers and producers. This analysis, while recognizing the importance of these interactions, focuses solely on household migration. For analyses that estimate a reduced form of the Roback (1982) model see Mueser and Graves (1995) and Partridge et al. (2008). For a simultaneous equation treatment see Carlino and Mills (1987). Modifying Partridge et al. (2008) and Partridge (2010), net migration for location i is written as:

$$\begin{aligned} &NM_i = f\left[V_i(w_i, r_i, a_i) - \overline{V}(w, r, a) - \overline{M}\right] \\ &i = 1, 2, \ldots, n \end{aligned} \quad (3.3)$$

where V_i is indirect utility in location i, $\overline{V}$ average utility across all locations, and $\overline{M}$ the average cost of migration. As evident in Equation 3.3, net migration depends on relative utility levels. Areas characterized by above average utility experience net in-migration while areas characterized by below average utility experience net out-migration. The inclusion of migration costs in the net migration function ensures the transition between equilibriums occurs gradually (Mueser and Graves, 1995).

In the preceding net migration model, the functional relationship between utility and various location-specific attributes can be either homogeneous, as is typically assumed, or heterogeneous across locations. Heterogeneous relationships might occur as a result of differences in local market structure (e.g. the availability of human capital, access to infrastructure, and presence of agglomeration economies), social contexts (e.g. wealth and life-cycle stages), and histories (e.g. social customs and consumer tastes). As previously discussed, Partridge et al. (2008) demonstrate the presence of heterogeneous relationships for some determinants of U.S. migration. In this analysis, both homogeneous and heterogeneous interpretations are employed during empirical estimation.

3.4 Data

Data for this analysis is obtained from multiple sources. A complete list of regression variables, descriptive statistics, and corresponding data sources are presented in Table 3.2. The U.S. Census Bureau develops annual county-level estimates of births, deaths, net domestic migration and net international migration. Estimates are made using administrative records (e.g. birth certificates, death certificates, income tax returns, Medicare enrollment, and military records) and information obtained from the ACS (U.S. Census Bureau, 2009). For this analysis, the dependent variable is constructed using information on net domestic migration between 2001 and 2009. International migration is omitted from the analysis because the location decision of international migrants is largely based on social connections and ethnic concentrations (Bauer et al., 2005). The formula used to construct the dependent variable is presented in Equation 3.4.

$$Net\ Inmigration\ Rate_i = \left[\frac{\sum_{t=2001}^{2009} Net\ Domestic\ Migration_{it}}{\left(\frac{1}{9}\right) * \sum_{t=2001}^{2009} Population_{it}} \right] * 100 \qquad (3.4)$$

Here the net in-migration rate for county *i* is calculated as the ratio of total net migration during the study period to the average county population. The net migration rate ranges from -81.76 (St. Bernard, LA) to 55.32 (Flagler, FL), with a mean value of -0.65. The spatial distribution of the net migration rate is presented in Figure 3.1. Congruent with the previous discussion, counties in the Northeast and Midwest regions were more likely to experience out-migration. Counties in the South and Mountain West were more likely to experience in-migration. Economic and demographic information is obtained from the U.S. Census Bureau's USA Counties database. This database provides information on median household income, local taxes, unemployment, employment by sector, population density, educational attainment, median age, and violent crime. Following Rupasingha and Goetz (2004), migration is modeled as a function of expected income, which is calculated using median household income and unemployment rates. Equation 3.5 depicts this calculation for county *i*.

$$Expected\ Income_i = \left(1 - \frac{Unemployment\ Rate_i}{100}\right) * Median\ Income_i * COLIndex_s \qquad (3.5)$$

Here *COLIndex* is a state-specific cost of living index. This index is obtained from the American Chamber of Commerce Researchers Association (ACCRA).[27]

---

[27] The ACCRA publishes a quarterly cost of living index for major U.S. metropolitan areas. Cost of living values are aggregated to the state level using a population-weighted average.

The U.S. Census Bureau's CBP provides county-level information on the total number of employees and business establishments by NAICS classification. The 2000 CBP is used to develop a measure of entertainment opportunities. Specifically, this measure is the total number of performing arts (e.g. theaters, dance companies, and music groups), local attractions (e.g. museums, historical sites, zoos, botanical gardens, and amusement parks), and recreational (e.g. golf courses, skiing facilities, bowling centers, and recreational sport centers) establishments operating within the county. The variable is standardized by county population. It ranges from 0 (multiple counties) to 7.17 establishments per thousand residents (San Juan, CO), with a mean of 0.37 establishments.

Migration patterns respond to a number of climate characteristics, including winter temperature, summer temperature, precipitation, sunshine, humidity, and wind speed (Poston et al., 2009; McGranahan, 2008). However, high levels of colinearity between many climate variables complicate the regression analysis. To avoid this problem, the variables included in this analysis are limited to average winter temperature (December, January, and February), annual precipitation, and a measure of climate temperateness.[28,29] Climate variables are acquired from the NCDC, which contains

[28] Use of alternate climate variables (e.g. humidity, sunshine, summer temperature) does not substantially alter regression results. Drought indices, such as the Standardized Precipitation Index and Palmer Drought Index, were not evaluated.

[29] Following Rupasingha and Goetz (2004), climate variables enter the net migration function linearly. In reality, household preference for climate amenities will exhibit nonlinearities.

information on temperature and precipitation climate normals (1981-2010) for 7500 U.S. weather stations. Stations are matched to their corresponding county using GIS. An arithmetic average is used to calculate climate variables for counties with multiple weather stations. Counties without a weather station are assigned average values from nearby stations. Winter temperatures range from 7.9 (Cavalier, ND) to 68.3 ºF (Monroe, FL), with a mean of 34.9 ºF. Annual precipitation ranges from 4.5 (Yuma, AZ) to 87.0 inches (Grays Harbor, WA), with a mean of 38.1 inches. Following McGranahan (2008), a measure of temperateness is constructed by regressing winter temperatures on summer temperatures. The residual from this regression, which represents temperateness, indicates the degree to which summer temperatures are higher or lower than predicted by winter temperatures. Temperateness ranges from -18.3 (Lincoln, OR) to 8.6 (Inyo, CA). Spatial distributions for these variables, as well as other environmental amenities, are presented in Appendix E.

Emissions data are retrieved from the EPA's Air Quality System database. This database contains county-level information on total 2002 emissions of hazardous air pollutants, as defined by the 1990 Clean Air Act. On average, counties produced 2.7 million lbs. of hazardous pollutants. The NPL provides the location of hazardous waste sites. As of 2001 the NPL included 1139 sites. The number of sites within each county, determined using GIS, ranges from 0 (multiple counties) to 23 (Santa Clara, CA), with an average of 0.4 sites. In addition, GIS is used to calculate the percentage of county area classified as water and classified as national or state park. Within the contiguous U.S., the percentage of land area classified as water ranges from 0 (multiple counties) to 56.9 (Grand Isle, VT), with an average of 1.8%. Likewise, the percentage of land area classified as either national or state park ranges from 0 (multiple counties) to 99.8 (Hamilton, NY), with an average of 1.7%. These area calculations are based on the StreetMap Pro atlas produced by ESRI. Previous studies have shown that local topography, or more precisely the aesthetic and recreational opportunities associated with local topography, are a significant determinant of net migration (McGranahan, 1999; Rupasingha and Goetz, 2004; McGranahan, 2008). Moreover, given potential correlations between topography and natural disaster risk it is necessary to control for landform variation. Following Rupasingha and Goetz (2004), this study uses the topography scale developed by McGranahan (1999). This scale is compiled from the National Atlas of the United States of America, which distinguishes between 21 types of land formations: plains (categories 1-4), tablelands (categories 5-8), plains with hills and mountains (categories 9-12), open hills and mountains (categories 13-17), and hills and mountains (categories 18-21). Each county is assigned the highest landform category that comprises at least 25% of its total land area. These values are then scaled based on their standard deviation from the mean. The scale ranges from -1.2 to 1.8, where higher values indicate more rugged terrain.

Finally, hazard risk data is obtained from the GRDP. The GRDP is a database developed and maintained by the United Nations Environment Programme and the UNISDR, in conjunction with numerous partner organizations. It provides spatial information on past natural hazard events and estimates of hazard frequency, hazard exposure and hazard risk. Information is available for various hazard types: earthquakes, cyclones, floods, landslides, tsunamis and volcanic eruptions. Dataset resolution, while inadequate for local-area planning, is sufficient to capture hazard variation across counties. Three hazard variables are constructed for this analysis: earthquake frequency, hurricane frequency, and flood frequency.[30] Specific to earthquakes, frequency is the expected number of earthquakes classified as 5 (strong) or greater on the Modified Mercalli Intensity scale per 1000 years. Specific to hurricanes, frequency is the expected number of hurricanes categorized as 3 (major) or greater on the Saffir-Simpson scale per 1000 years. Both measures pertain to major hazard events that would likely damage housing structures. There are no intensity scales associated with flood data. Consequently, flood frequency is the expected number of events per 1000 years.

Hazard variables are calculated, using GIS, as the spatially weighted average of the expected number of events within county boundaries. All three hazard-risk variables are highly skewed. Earthquake frequency ranges from 0 (multiple counties) to 206.3 (Mono, CA), with an average of 1.7 events. Hurricane frequency ranges from 0 (multiple counties) to 55.1 (Broward, FL), with an average of 0.6 events. Flood frequency ranges from 0 (multiple counties) to 86.6 (Plaquemines, LA), with an average of 2.7 events. The spatial distribution of these variables is presented in Figure 3.2. Unsurprisingly, earthquake risk is highest along the Pacific Coast while hurricane risk is concentrated along the Atlantic and Gulf coasts. Flood risk is distributed throughout the contiguous U.S. but is highest within the Midwest.

---

[30] The GRDP does not contain frequency data for tornados, wildfire, drought, and blizzards. Exclusion of these hazards may bias parameter estimates (i.e. omitted variable bias). An important extension of this analysis would be to include risk measures for the omitted hazards.

3.5 Empirical Model and Methodologies

Spatial econometric techniques are increasingly used to address issues of spatial dependence and spatial heterogeneity. This analysis employs two such techniques: SAC and GWR. The SAC model controls for spatial autocorrelation in net migration rates; it estimates global parameter values that are unbiased and efficient. The GWR model estimates separate regression coefficients for each county. By mapping these coefficients it is possible to evaluate heterogeneous effects of environmental attributes on migration patterns. A more detailed description of each technique is provided below.

3.5.1 Empirical Model

Net in-migration is modeled as a function of location-specific attributes. Equation 3.6 depicts the function in general notation.

$$M_i = f(E_i, D_i, A_i) \tag{3.6}$$

Here *M* denotes county *i*'s net in-migration rate between 2001 and 2009, *E* is a vector of county-specific economic characteristics, *D* is a vector of demographic and social characteristics, and *A* is a vector of environmental amenities and disamenities—including the risk of natural disaster. The model contains several potential endogenous relationships. In particular, many economic and demographic characteristics are likely to be affected by the inflow or outflow of migrants. Following Rupasingha and Goetz (2004), endogeneity problems are avoided by using beginning-of-period values for all time-varying explanatory variables. With the exception of per capita taxes the economic and demographic variables are 2000 values. Per capita taxes are 2002 values.

Economic characteristics included in the model are expected income (EXPINC), per capita local taxes (PCTAX), and the portion of employees working in construction

(IND_CONST), manufacturing (IND_MNF), trade (IND_TRADE), transportation (IND_TRANS), or finance (IND_FIN). The portion of employees working in agriculture, fishing, forestry, or mining serves as the base category. Demographic and social characteristics include population density (POPDEN), median age (MEDAGE), the portion of residents 25 years and older with a high school degree (HSEDU), violent crime (VCRIME), and the number of entertainment and recreation establishments (ARTREC). Following previous studies the model assumes a quadratic relationship between net migration and population density. Environmental amenities and disamenities include winter temperature (WINTEMP), annual precipitation (PRECIP), climate temperateness (TEMPERATE), hazardous air pollutants (EMISSONS), hazardous waste sites (NPLSITES), topography (TOPO), portion of county area classified as water (WATER), portion classified as a national or state park (PARKS), and the expected frequency of earthquakes (EFREQ), hurricanes (HFREQ), and floods (FFREQ).

Dummy variables indicating central (URBAN) and outlying (SUBURBAN) urban areas are incorporated into the model. These variables control for unobserved characteristics of urban and suburban settings. Rural areas serve as the base category. Additional unobserved characteristics are controlled for using census division and state fixed-effects.[31] The model also controls for the effects of Hurricanes Katrina and Rita. In 2005, Katrina and Rita caused severe damaged in several Gulf Coast counties. There is a

[31] Census divisions are New England, Mid-Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific. New England serves as the base category. When state fixed-effects are employed, Alabama serves as the base category.

well-founded concern that the massive out-migration that followed these events will dominate the estimated coefficients on hurricane and flooding risk. To address this issue the analysis incorporates a dummy variable (DISAREA) indicating the 117 counties declared disaster areas by the Federal Emergency Management Agency (FEMA).[32]

3.5.2 Spatial Simultaneous Autoregressive (SAC) Regression

Ruspasingha and Goetz (2004) demonstrate the presence of significant spatial dependency in net migration rates. Parameter estimates obtained without controlling for

---

[32] These counties and parishes are: in Alabama (Baldwin, Choctaw, Clarke, Greene Hale, Marengo, Mobile, Pickens, Sumter, Tallapoosa, Washington); in Louisiana (Acadia, Allen, Ascension, Assumption, Beauregard, Calcasieu, Cameron, East Baton Rouge, East Felicia, Evangeline, Iberia, Iberville, Jefferson, Jefferson Davis, Lafayette, Lafourche, Livingston, Orleans, Plaquemines, Pointe Coupe, Sabine, St. Bernard, St. Charles, St. Helena, St. James, St John Baptist, St. Landry, St. Martin, St. Mary, St. Tammany, Tangipahoa, Terrebonne, Vermilion, Vernon, Washington, West Baton Rouge, West Felicia); in Mississippi (Adams, Amite, Attala, Choctaw, Claiborne, Clarke, Copiah, Covington, Forrest, Franklin, George, Greene, Hancock, Harrison, Hinds, Jackson, Jasper, Jefferson, Jefferson Davis, Jones, Kemper, Lamar, Lauderdale, Lawrence, Leake, Lincoln, Lowndes, Madison, Marion, Neshoba, Newton, Noxubee, Oktibbeha, Pearl River, Perry, Pike, Rankin, Scott, Simpson, Smith, Stone, Walthall, Warren, Wayne, Wilkinson, Winston, Yazoo; and in Texas (Angelina, Brazoria, Chambers, Fort Bend, Galveston, Hardin, Harris, Jasper, Jefferson, Liberty, Montgomery, Nacogdoches, Newton, Orange, Polk, Sabine, San Augustin, San Jacinto, Shelby, Trinity, Tyler, Walker)

this dependency are biased and inefficient (Anselin, 2001). The SAC model, presented using vector notation in Equation 3.7, controls for spatial dependence in net migration patterns.

$$\begin{aligned} \mathbf{M} &= \rho\mathbf{WM} + \mathbf{E}\beta_E + \mathbf{D}\beta_\mathbf{D} + \mathbf{A}\beta_\mathbf{A} + \mathbf{u} \\ \mathbf{u} &= \lambda\mathbf{Wu} + \mathbf{e} \end{aligned} \tag{3.7}$$

As before **M**, **E**, **D**, and **A** denote county-level net migration rates, economic characteristics, demographic and social characteristics, and environmental amenities and disamenities, respectively. Spatial dependence is incorporated into the model with the terms $\rho\mathbf{WM}$ and $\lambda\mathbf{Wu}$, where ρ and λ are estimated parameters and **W** is a spatial weight matrix.[33] The term $\rho\mathbf{WM}$, referred to as a spatial-lag, controls for spatial correlation in the dependent variable (i.e. that a county's net migration rate is affected by net migration rates in surrounding counties). The term $\lambda\mathbf{Wu}$, referred to as a spatial-error, controls for spatial correlation in the error term (i.e. that a county's net migration rate is affected by random shocks in neighboring counties). The spatial weight matrix is n×n, where n is the number of observations. The diagonal elements are set to zero and the off-diagonal elements, which identify the neighbor set for each county, are row standardized such that their sum is one. More formally,

[33] Spatial dependency models can be estimated with the spatial-lag term, the spatial-error term, or both the spatial-lag and spatial-error terms. The Lagrange Multiplier test statistic is used to evaluate the presence of spatial dependence. Congruent with Ruspasingha and Goetz (2004) spatial dependence is identified in both spatial-lag and spatial-error specification.

$$W_{ij} = \frac{d_{ij}}{\sum_{i=1}^{n} d_{ij}} \qquad where \ \ d_{ij} = \begin{cases} 1 & \text{if counties } i \text{ and } j \text{ are neighbors} \\ 0 & \text{otherwise.} \end{cases}$$

Under this construction, county migration rates are partially determined by the weighted average of migration rates and unexplained variation in neighboring counties, but are unaffected by non-neighboring counties. Following Ruspasingha and Goetz (2004) neighbors are defined as the set of adjacent counties.

Estimation of the SAC model is complicated by inclusion of the spatial-lag and spatial-error terms. The spatial-lag, which introduces an endogenous relationship to the model, violates the OLS assumption that regressors are uncorrelated with the error. Likewise, the spatial-error term violates the OLS assumption of uncorrelated errors. Under these circumstances, parameter estimates obtained from OLS are biased and inconsistent. Several studies have addressed these issues through the use of maximum likelihood (ML) estimation (Lu and Zhang, 2010). This method, first proposed by Ord (1975), produces consistent and unbiased parameter estimates given that the error term is normally distributed (i.e. $\mathbf{e}$~i.i.d. N[0,$\sigma^2\mathbf{I}$]). The reduced form of the model is given in Equation 3.8.

$$\mathbf{M} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{Z}\delta + (\mathbf{I} - \rho\mathbf{W})^{-1}(\mathbf{I} - \lambda\mathbf{W})^{-1}\mathbf{e} \tag{3.8}$$

Here $\mathbf{Z}$=($\mathbf{E}$,$\mathbf{D}$,$\mathbf{A}$) and δ=(β_E',β_D',β_A')′. The log-likelihood, as presented Drukker et al. (2011), is presented in Equation 3.9.

$$\begin{aligned} \ln L = & -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) + \ln\|\mathbf{I} - \rho\mathbf{W}\| + \ln\|\mathbf{I} - \lambda\mathbf{W}\| \\ & - \frac{1}{2\sigma^2}\left[(\mathbf{I} - \rho\mathbf{W})\mathbf{M} - \mathbf{Z}\delta\right]'(\mathbf{I} - \lambda\mathbf{W})'(\mathbf{I} - \lambda\mathbf{W})\left[(\mathbf{I} - \rho\mathbf{W})\mathbf{M} - \mathbf{Z}\delta\right] \end{aligned} \tag{3.9}$$

In practice the model is estimated using a concentrated log-likelihood function (i.e. the quasi maximum likelihood estimator). To derive the concentrated function, Equation 3.9 is maximized with respect to δ and σ^2. Results are presented in Equations 3.10 and 3.11.

$$\delta^*(\rho,\lambda) = \left[\mathbf{Z}'(\mathbf{I}-\lambda\mathbf{W})'(\mathbf{I}-\lambda\mathbf{W})\mathbf{Z}\right]^{-1}\mathbf{Z}'(\mathbf{I}-\lambda\mathbf{W})'(\mathbf{I}-\lambda\mathbf{W})(\mathbf{I}-\rho\mathbf{W})\mathbf{M} \qquad (3.10)$$

$$\sigma^{2*}(\rho,\lambda) = \left(\frac{1}{n}\right)\left[(\mathbf{I}-\rho\mathbf{W})\mathbf{M}-\mathbf{Z}\delta^*(\rho,\lambda)\right](\mathbf{I}-\lambda\mathbf{W})'(\mathbf{I}-\lambda\mathbf{W})\left[(\mathbf{I}-\rho\mathbf{W})\mathbf{M}-\mathbf{Z}\delta^*(\rho,\lambda)\right] \qquad (3.11)$$

Substituting these expressions into Equation 3.9 yields the concentrated log-likelihood function, presented in Equation 3.12.

$$\ln L_c(M|\rho,\lambda) = -\frac{n}{2}(\ln(2\pi)+1) - \frac{n}{2}\ln\left(\sigma^{2*}(\rho,\lambda)\right) + \ln\|\mathbf{I}-\rho\mathbf{W}\| + \ln\|\mathbf{I}-\lambda\mathbf{W}\| \qquad (3.12)$$

The concentrated function is maximized to produce parameter estimates ρ^* and λ^*. Substituting these parameters into Equations 3.10 and 3.11 yields estimated values for δ^* and σ^{2*}.

The ML estimates exhibit two major limitations. First, the estimation procedure is complex, particularly with respect to evaluation of the Jacobian determinant, and computationally burdensome with large datasets. Second, ML produces inconsistent parameter estimates when the error term is heteroskedastic (Arraiz et al., 2008). Despite these limitations, this analysis employs the ML estimator for the majority of regression specifications—primarily because it is the most common method of estimating the SAC model (Lu and Zhang, 2010).

Kelejian and Prucha (1998) propose an alternate method, the generalized spatial two-stage least squares (GS2SLS) estimator, which employs both instrumental variable (IV) and generalized-method-of-moments (GMM) techniques. Compared to ML, this method is more tractable and, as demonstrated by Arraiz et al. (2010), produces consistent estimates when errors are heteroskedastic. The GS2SLS estimation proceeds in three steps. First, initial estimates of δ are obtained using an IV approach, where the spatially lagged independent variables are used as instruments for the spatially lagged dependent variable (i.e. $\mathbf{WM}$=f[$\mathbf{X}$, $\mathbf{WX}$, $\mathbf{W^2X}$]). Second, estimates of λ and σ are obtained using a GMM procedure and initial δ values. Third, estimated values for λ and σ are used to perform a spatial Cochrane-Orcut transformation of the data. Final estimates for δ are then obtained by estimating the transformed model with instrumental variables. A detailed description of this procedure is available in Kelejian and Prucha (1998), Arraiz et al. (2010), and Drukker et al. (2011). This analysis employs the GS2SLS estimator in a single regression specification to evaluate whether heteroskedasticity is a major concern.

3.5.3 Geographically Weighted Regression (GWR)

GWR allows for spatial heterogeneity in the parameter values by estimating a separate regression model for each location. Regressions are conducted using information from the primary location and neighboring locations within a specified bandwidth. As a result, each regression uses a unique sub-sample of the global dataset. Locations within the bandwidth are weighted according to their inverse distance from the primary location. Under this construction, nearby locations have a greater influence on parameter estimates than distant locations. The GWR is presented in Equation 3.13.

$$\begin{aligned} &M_i = \beta_{iE} E_i + \beta_{iD} D_i + \beta_{iA} A_i + \varepsilon_i \\ &\varepsilon \sim \text{i.i.d. N}(0, \sigma^2) \\ &i = 1, 2, \ldots, n \end{aligned} \tag{3.13}$$

Here the subscripts indicate separate data sub-samples and parameter estimates at every location. Parameter estimates are obtained, as depicted in Equation 3.14, using a variant of weighted least squares.[34]

$$\begin{aligned} &\beta_i = (\mathbf{X}_i'\mathbf{W}_i\mathbf{X}_i)^{-1}\mathbf{X}_i'\mathbf{W}_i\mathbf{Y}_i \\ &i = 1,2,\ldots,n \end{aligned} \tag{3.14}$$

The weight matrix **W**, which is n×n, is specific to location *i*. The off-diagonal elements are zero and the diagonal elements are location weights. The weight matrix for location *i* is calculated using Equation 3.15.

$$W_{ik} = \exp\left(\frac{-d_{ik}}{b^2}\right) \tag{3.15}$$

Here the weight for diagonal element *k* is a function of the distance *d* between location *i* and *k* and the bandwidth *b*. As evident from Equation 3.15, the primary location is assigned a weight of one and neighboring locations are assigned relatively lower weights based on distance. This analysis uses a fixed bandwidth, implying that the bandwidth distance is constant across location but the number of neighbors varies. The optimal bandwidth is determined using a cross-validation technique that minimizes:

$$\sum_{i=1}^{n}\left[M_i - M^*_{\neq i}(b)\right]^2$$

where $M^*(b)$ is the fitted value of M_i using the weighted regression for location *i* and bandwidth *b*, with the observation for point *i* excluded from the calibration process.

---

[34] GWR regressions are conducted using the user-written State code developed by Mark S. Pearce. Documentation is available online at http://www.staff.ncl.ac.uk/m.s.pearce/ stbgwr.htm.

The GWR approach offers several advantages over global OLS models.[35] The primary advantage, as previously discussed, is the ability to evaluate spatial heterogeneity in parameter estimates. By mapping parameter values it is possible to identify geographic variation in behavior and preferences. In some instances, it may be possible to identify significant correlations using GWR that are obscured (i.e. averaged out) in a global model. Another advantage of GWR is improved model performance. In particular, estimating location-specific parameter values leads to lower residuals and more accurate predicted values. Finally, as with the SAC model, GWR reduces spatial dependency. In this case, spatial dependency is reduced, although not necessarily eliminated, by using weighted sub-samples (Partridge et al. 2008).

3.6 Results

3.6.1 SAC Regression Results

Results for the SAC model are reported in Table 3.3.[36] Model 1 regresses the domestic migration rate against economic, demographic and social, and environmental characteristics using the ML estimator. Subsequent models incorporate census division fixed-effects (Model 2), state fixed-effects (Model 3), and address potential heteroskedasticity problems using the GS2SLS estimator (Model 4). The latter model also includes state fixed-effects.

---

[35] A key disadvantage of the GWR, compared to the global models, is that regressions are conducted with smaller sample sizes. This may produce less efficient parameter estimates (Partridge et al. 2008).

[36] Stata codes for the SAC and GWR analyses are presented in Appendix I.

Regression coefficients are robust across model specifications. With respect to economic variables, the net in-migration rate is positively correlated with expected income and the percentage of the labor force employed in construction, manufacturing, and finance. Likewise, net migration is negatively correlated with the per capita tax rate. With respect to demographic characteristics, net migration is positively correlated with the county's median age and the percentage of residents with a high school degree. The positive coefficient on median age contradicts several previous findings but is consistent with Rupasingha and Goetz (2004). As discussed by Rupasingha and Goetz (2004), this may reflect the increasing importance of retirement migration in which elderly households are attracted to locations with older age distributions. As expected there is a quadratic relationship between net migration and population density. Counties with higher densities attract fewer migrants but at a decreasing rate. Neither violent crime nor the number of entertainment and recreation opportunities has a significant effect on migration.

Compared to rural counties, urban and suburban areas experienced net in-migration. The slightly larger coefficient on suburban counties is consistent with the notion of migration from central to outlying urban areas. The census division and state fixed-effects, while not reported, are jointly significant. Compared to New England, migration is positively correlated with the Pacific census region and negatively correlated with the East North Central, West North Central, and West South Central. Among states, the largest fixed-effects are associated with Wyoming, Washington, Arizona, Maine, and Delaware. The smallest fixed-effects are associated with Kansas, Nebraska, Louisiana, Iowa, and Arkansas.

Net migration rates, in accordance with the spatial equilibrium model, are significantly correlated with environmental amenities. Migration is positively correlated with the three climate variables: winter temperatures, precipitation, and temperateness. Positive coefficients on temperature and precipitation are in line with previous studies. They imply, all else equal, that households are attracted to counties with warmer winters and higher amounts of precipitation. In contrast, the positive coefficient on temperateness contradicts previous studies. It suggests that households are attracted to counties with larger annual temperature variation. One possible explanation, supported by the negative and insignificant coefficient in Model 1, is that the effects of climate temperateness are being absorbed in census division and state fixed-effects. In terms of non-climate amenities, households are attracted to counties with greater surface water area, more rugged terrain, fewer hazardous waste sites, and lower per capita emissions. These findings support those of Gawande et al. (2000), Rupasingha and Goetz (2004), and McGranahan (2008).

Earthquake, hurricane, and flood risk are negatively and significantly correlated with net migration rates—even after controlling for counties affected by Hurricanes Katrina and Rita. The indicator variable for Katrina- and Rita-affected counties is negative and significant in Models 1 and 2, but insignificant when state fixed-effects are included. These findings imply that U.S. migration and population growth patterns are partially determined by high-risk low-probability disaster events. Specifically, a marginal increase in expected frequency of earthquake, hurricane, and flood events reduces net in-migration (for the nine year study period) by 0.03, 0.23, and 0.11%, respectively. While small, these effects are not necessarily trivial. Reestimating the model using standardized variables (i.e. variables normalized so that their mean=0 and standard deviation=1) provides an indication of each regressors relative importance. The standardized coefficients, using Model 4, for the disaster risk variables are -0.03 (earthquake), -0.07 (hurricane), and -0.06 (flood). These variables can be interpreted in terms of standard deviation. For example, a marginal increase in the standard deviation of earthquake frequency reduces net in-migration by 0.03 standard deviations. Among the disaster risk variables, expected hurricane and flood frequency have approximately twice the effect on net migration as earthquake frequency. Moreover, hurricane and flood frequency have a greater or similar effect on migration as water area (0.01), topography (0.07), hazardous waste sites (-0.08), and emissions (-0.09). Their effect is roughly one-half to two-thirds that of the three climate variables: winter temperature (0.14), temperateness (0.11), and precipitation (0.10).[37]

3.6.2 GWR Results

Several changes are made to the regression model in order to conduct the GWR analysis. Explanatory variables that are either highly skewed or geographically concentrated result in excessively large bandwidths—and dramatically reduce spatial variation in parameter estimates. To prevent this problem the following explanatory variables are excluded: URBAN, SUBURBAN, OCEAN, DISAREA, and the fixed-effects. Due to the moderately high correlation between urban areas and population density exclusion of URBAN and SUBURBAN indicator variables does not substantially

[37] Standardized coefficients have similar units (i.e. standard deviation units) and are directly comparable. Care should be taken, however, in that a standard deviation change in one variable is not necessarily equivalent to a standard deviation change in another.

alter the model's explanatory power or coefficients on the remaining variables. In addition, the model uses a combined disaster risk variable (MultiFreq) that avoids the highly concentrated nature of earthquake and hurricane risk. MultiFreq is the unweighted sum of the three disaster frequency variables. It represents the expected number of hazard events per 1000 years.

The GWR produces a unique set of regression coefficients for each geographic location. The average of the location-specific coefficients, referred to as the global model, is presented in Table 3.4. Results are similar to those obtained from the SAC analysis. The unexpected signs on violent crime and recreational opportunities are likely a result of omitted variable bias—an issue that is not present in the SAC model because of the fixed-effects and spatial autocorrelation parameter. The global model, however, masks spatial heterogeneity in the regression coefficients. To illustrate this, Table 3.4 also reports results at select percentiles of the coefficient distribution. All variables exhibit considerable variation across counties, implying the effects of a particular regressor are stronger in some counties than in others.[38] Furthermore, there are often a few counties where the function relationship is reversed (i.e. a positive relationship is negative with a few counties). Evaluating the characteristics of these counties may yield further insight into household preferences. For example, the coefficients for WINTEMP range from -0.42 to 0.64, with a mean value of 0.14. In the majority of counties there is a positive correlation between net migration and temperature, signifying that households

[38] Monte Carlo simulations indicate that most regression variables, and all environmental variables, exhibit statistically significant spatial variation. The bandwidth is also statistically significant, indicating the GWR estimator outperforms OLS.

prefer moderate winters. There are, however, a number of counties where net migration is negatively correlated with temperature.[39] Partridge et al. (2008) suggest this correlation is due to the recreation opportunities, and winter tourism industries, available in these locations.

By mapping the estimated parameter values it is possible to identify spatial patterns otherwise ignored in standard regression techniques. Figure 3.3 presents a map of estimated coefficients for MultiFreq. Maps for the remaining environmental amenities and disamenities (not discussed) are provided in Appendix F. GWR maps are shaded in accordance with standard deviation from the mean. The lightest shade denotes coefficients over one standard deviation below the mean and the darkest shade denotes coefficients over one standard deviation above the mean. Figure 3.3 shows distinct regional patterns in MultiFreq coefficients. The effect of hazard risk on migration rates is strongest along the Gulf Coast and at the border of Wyoming, Montana, North Dakota, and South Dakota.[40] The strong negative effect in the latter area is surprising, as the region exhibits only moderate risk levels. Slightly weaker effects are present throughout the South, Midwest, and West. The weakest effects, and occasionally positive effects,

---

[39] These counties are primarily located in Alabama, Maine, Mississippi, Minnesota and Wisconsin.

[40] The GWR model does not control for counties declared disaster areas in the wake of hurricanes Katrina and Rita. As a result, the coefficients obtained for the Gulf Coast are partially driven by these results. An alternate model that excluded these counties was also estimated. Spatial patterns were not substantially different.

occur in the Northeast and in parts of the Mountain West (namely, Arizona, Colorado, New Mexico, and Texas).

3.7 Discussion and Conclusion

Climate change will impose considerable, if uncertain, costs on society. Mitigating these costs will require an improved understanding of the physical processes and behavior responses of households and firms. Among other things, climate change is expected to increase the frequency and intensity of hazard events—specifically those related to extreme precipitation (Greenough et al., 2001; van Aalst, 2006; IPCC, 2011). This analysis examines the effects of hazard frequency on internal U.S. migration. In doing so, it offers insight on how changing climate conditions will affect migration patterns.

This analysis employs SAC and GWR techniques to estimate the effects of earthquake, hurricane, and flood risk on the net in-migration rates of U.S. counties between 2000 and 2009. The SAC model controls for spatial dependency using a spatial-lag and spatial-error term. Results show, consistent with Chapter 2, that residential location decisions are partially determined by the risk of high-intensity low-probability hazard events. Specifically, a marginal increase in the expected number of earthquake events reduces net in-migration by 0.03%. Likewise, a marginal change in the expected number of hurricane and flood events reduces net in-migration by 0.22 and 0.11%, respectively. While small, the marginal effects of disaster risk are comparable to several environmental amenities. Standardizing the coefficients allows for direct comparison across variables with different units. Hurricane and flood risk influence migration in a manner comparable to water area, topography, hazardous waste sites, and emissions.

The GWR regression estimates separate regression coefficients for each location, allowing for spatial heterogeneity in parameter values. Results display significant spatial variation in estimated parameter values. Mapping the coefficients for hazard risk shows clear regional patterns. The spatial patterns depicted in Figure 3.3 lend themselves to three observations. First, coefficients in the bottom half of the distribution are located in areas that are most likely to witness increased hurricane intensity due to climate change (i.e. the South and Midwest). Second, coefficients on the Pacific Coast are on the upper half of the distribution. This is consistent with findings from the SAC model that earthquakes have less of an impact than hurricanes and floods. Third, there is no obvious similarity between the two regions with the weakest migration-hazard relationship (i.e. Northeast and Mountain West). The two regions exhibit different climates, demographic characteristics, and disaster risk levels. This implies that there may be two underlying explanations for the weak migration-hazard relationship.

Understanding household responses to changing environmental amenities is vital to anticipating and mitigating the costs of climate change. This work demonstrates that global parameter estimates obscure regional differences in how environmental amenities affect migration, including hazard risk. Incorporating spatial heterogeneity into population growth models will improve estimates of amenity migration and migration related to climate change.

Table 3.1 Net Domestic Migration by Census Geography: 2001-2009

Region/Division	Total Migration	Average Annual Migration	Average Annual Rate (Migration/1000 People)
Northeast	-2,488,084	-276454	-5.1
New England	-353,914	-39324	-2.8
Middle Atlantic	-2,134,170	-237130	-5.9
Midwest	-1,719,445	-191049	-2.9
East North Central	-1,546,573	-171841	-3.7
West North Central	-172,872	-19208	-1.0
South	3,803,776	422642	3.9
South Atlantic	2,767,011	307446	5.5
East South Central	392,560	43618	2.4
West South Central	644,205	71578	2.1
West	403,753	44861	0.7
Mountain	1,513,828	168203	8.3
Pacific	-1,110,075	-123342	-2.6

† Values calculated using the U.S. Census Bureau's Annual Components of Population Change

Table 3.2 Descriptive Statistics for Net Migration Analysis

Variable	Description	Units	Source	Mean	Std. Dev.
NETMIG	Net in-migration	PERCENT	2000-09 USCB	-0.653	9.819
EXPINC	Expected income	TH$	2000 CCCB	32.515	7.946
PCTAX	Local taxes	TH$/PER	2002 CCDB	0.997	0.795
IND_CONST	Employment classified as construction	PERCENT	2000 CCDB	7.715	2.378
IND_MNF	Employment classified as manufacturing	PERCENT	2000 CCDB	15.919	9.084
IND_TRADE	Employment classified as trade	PERCENT	2000 CCDB	14.493	2.495
IND_TRANS	Employment classified as transportation	PERCENT	2000 CCDB	5.452	1.839
IND_FIN	Employment classified as finance	PERCENT	2000 CCDB	4.570	1.863
POPDEN	Population density	TH PERSON/SQM	2000 CCDB	0.242	1.668
URBAN	Classified as central urban area	0/1	2009 USCB	0.402	0.490
SUBURBAN	Classified as outlying urban area	0/1	2009 USCB	0.169	0.375
HSEDU	Population >25 that completed high school	PERCENT	2000 CCDB	77.335	8.729
MEDAGE	Median age of population	YEARS	2000 CCDB	37.376	3.967
VCRIME	Violent crime	CRM/100TH PERSON	2000 CCDB	2.486	3.019
ARTREC	Art and recreation related businesses	FIRMS/TH PERSON	2000 CBP	0.340	0.369
WINTEMP	Climate normal: winter temperature	DEGREE	NCDC	34.925	11.037
TEMPERATE	Climate normal: temperate climate	INDEX	NCDC	0.00	3.282
PRECIP	Climate normal: annual precipitation	INCHES	NCDC	38.143	13.651
OCEAN	Adjacent to ocean or Great Lake	0/1	ESRI	0.093	0.291
lnWATER	Water area	lnPERCENT	ESRI	0.657	0.726
lnPARKS	State and national parks area	lnPERCENT	ESRI	0.385	0.756
TOPO	Topography scale	INDEX	1970 McGranahan	0.002	1.000
EMISSIONS	Annual HAP 188 emissions	TH LBS/PERSON	2002 EPA	0.065	0.172
NPLSITES	National Priority List sites	SITES	2008 EPA	0.394	1.279
DISAREA	Katrina or Rita disaster area	0/1	FEMA	0.038	0.190
EFREQ	Expected frequency of earthquake events	EVENTS/TH YREARS	GRDP	1.730	11.305
HFREQ	Expected frequency of hurricane events	EVENTS/TH YREARS	GRDP	0.563	3.269
FFREQ	Expected frequency of flood events	EVENTS/TH YREARS	GRDP	2.705	5.575
MultiFreq	Expected frequency of hazard events	EVENTS/TH YREARS	GRDP	5.000	13.184

Table 3.3 Regression Results for SAC Model

Variable	Model 1: ML	Model 2: ML DIV Fixed-Effects	Model 3: ML ST Fixed-Effects	Model 4: GS2SLS ST Fixed Effects
Constant	-47.163*** (3.21)	-46.952*** (3.14)	-54.928*** (3.28)	-51.699*** (3.16)
EXPINC	0.305*** (0.03)	0.319*** (0.03)	0.293*** (0.03)	0.268*** (0.03)
PCTAX	-2.007*** (0.21)	-1.937*** (0.21)	-1.765*** (0.21)	-1.671*** (0.21)
IND_CONST	0.935*** (0.07)	0.931*** (0.07)	0.950*** (0.06)	0.916*** (0.07)
IND_MNF	-0.007 (0.02)	-0.011 (0.02)	0.013 (0.02)	0.010 (0.02)
IND_TRADE	0.570*** (0.06)	0.570*** (0.06)	0.577*** (0.05)	0.583*** (0.05)
IND_TRANS	-0.076 (0.07)	-0.063 (0.07)	-0.018 (0.07)	-0.018 (0.07)
IND_FIN	0.395*** (0.10)	0.406*** (0.10)	0.411*** (0.10)	0.384*** (0.09)
POPDEN	-0.989*** (0.19)	-1.079*** (0.19)	-1.100*** (0.19)	-1.136*** (0.19)
POPDEN^2	0.019*** (0.00)	0.020*** (0.00)	0.021*** (0.00)	0.022*** (0.00)
URBAN	1.051*** (0.36)	0.853** (0.36)	0.764** (0.35)	0.766** (0.35)
SUBURBAN	1.097*** (0.38)	0.947** (0.38)	0.947** (0.37)	0.914** (0.37)
HSEDU	0.155*** (0.03)	0.180*** (0.03)	0.200*** (0.03)	0.188*** (0.03)
MEDAGE	0.083** (0.04)	0.094** (0.04)	0.075** (0.04)	0.081** (0.04)
VCRIME	0.001 (0.05)	-0.012 (0.05)	-0.012 (0.05)	-0.022 (0.05)
ARTREC	-0.273 (0.40)	-0.197 (0.40)	-0.003 (0.39)	-0.058 (0.38)
WINTEMP	0.141*** (0.02)	0.078*** (0.03)	0.134*** (0.04)	0.127*** (0.04)
TEMPERATE	-0.074 (0.06)	0.305*** (0.09)	0.383*** (0.08)	0.342*** (0.08)
PRECIP	0.064*** (0.02)	0.073*** (0.02)	0.098*** (0.03)	0.076*** (0.02)
OCEAN	-0.632 (0.54)	-0.318 (0.56)	-0.644 (0.55)	-0.505 (0.53)
lnWATER	0.591*** (0.21)	0.415* (0.22)	0.255 (0.21)	0.238 (0.20)
lnPARKS	0.048 (0.18)	-0.014 (0.19)	0.006 (0.19)	0.018 (0.18)
TOPO	0.836*** (0.18)	0.840*** (0.19)	0.778*** (0.19)	0.680*** (0.17)
EMISSIONS	-4.905*** (1.35)	-5.510*** (1.35)	-5.725*** (1.34)	-5.700*** (1.32)
EMISSIONS^2	1.298*** (0.30)	1.428*** (0.30)	1.429*** (0.30)	1.399*** (0.30)
NPLSITES	-0.564*** (0.11)	-0.610*** (0.11)	-0.602*** (0.11)	-0.578*** (0.11)
DISAREA	-3.082*** (0.86)	-2.059** (0.92)	-0.611 (0.92)	-0.631 (0.83)
EFREQ	-0.027** (0.01)	-0.024* (0.01)	-0.027* (0.01)	-0.027** (0.01)
HFREQ	-0.199*** (0.04)	-0.209*** (0.05)	-0.242*** (0.04)	-0.226*** (0.04)
FFREQ	-0.118*** (0.03)	-0.098*** (0.03)	-0.115*** (0.03)	-0.112*** (0.02)
Rho	0.048*** (0.01)	0.034*** (0.01)	0.038*** (0.01)	0.057*** (0.01)
Lambda	0.035*** (0.01)	0.047*** (0.01)	0.022** (0.01)	0.000 (0.01)
N	3107	3107	3107	3107
AIC	20591.170	20526.255	20450.292	NA

* p<0.1 ** p<0.05 *** p<0.01
Standard Errors in Parenthesis

Table 3.4 Regression Results for GWR Model

Variable	Global Model	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
Constant	-53.139*** (2.50)	-140.16	-54.348	-36.573	-31.172	-5.311
EXPINC	0.358*** (0.03)	-0.649	0.257	0.426	0.597	0.933
PCTAX	-2.425*** (0.21)	-7.14	-2.158	-1.039	-0.051	2.131
IND_CONST	1.202*** (0.06)	0.038	0.74	0.9	1.069	2.608
IND_MNF	0.058*** (0.02)	-0.23	-0.089	-0.049	0	2.837
IND_TRADE	0.685*** (0.06)	-0.129	0.336	0.511	0.609	1.531
IND_TRANS	-0.165** (0.08)	-3.443	-0.145	-0.065	0.041	1.519
IND_FIN	0.525*** (0.10)	-1.656	0.261	0.502	0.793	4.002
POPDEN	-0.944*** (0.20)	-18.687	-6.955	-3.857	-2.176	82.001
POPDEN^2	0.018*** (0.00)	-25.287	0.03	0.096	0.785	3.802
HSEDU	0.160*** (0.03)	-0.187	0.022	0.126	0.193	0.712
MEDAGE	0.067* (0.04)	-0.501	-0.112	0.011	0.138	1.024
VCRIME	0.109** (0.05)	-0.949	-0.108	-0.04	0.066	1.692
ARTREC	-0.851** (0.42)	-14.857	-1.264	0.023	1.078	4.568
WINTEMP	0.160*** (0.02)	-0.418	0.015	0.143	0.272	0.644
TEMPERATE	-0.196*** (0.05)	-1.808	-0.725	-0.39	0.238	2.877
PRECIP	0.044*** (0.01)	-0.783	-0.039	0.056	0.142	0.36
lnWATER	0.658*** (0.20)	-2.196	-0.036	0.455	1.171	5.333
lnPARKS	0.352* (0.19)	-4.538	-0.408	0.032	0.656	3.515
TOPO	1.348*** (0.16)	-5.548	0.188	0.73	1.194	2.798
EMISSIONS	-5.534*** (1.46)	-49.984	-15.579	-7.953	-1.506	21.751
EMISSIONS^2	1.637*** (0.33)	-38.818	-0.619	2.883	5.875	128.763
NPLSITES	-0.496*** (0.11)	-3.428	-0.307	-0.189	-0.012	1.236
MultiFreq	-0.083*** (0.01)	-0.335	-0.117	-0.063	-0.038	0.806
N	3107					
AIC	21087.41					

* p<0.1 ** p<0.05 *** p<0.01
Standard Errors in Parenthesis

Figure 3.1 Net In-Migration Rate by County: Standard Deviation from Mean
[Note: Lightest color denotes values <1 s.d. below mean and darkest color denotes values >1 s.d. above mean.]

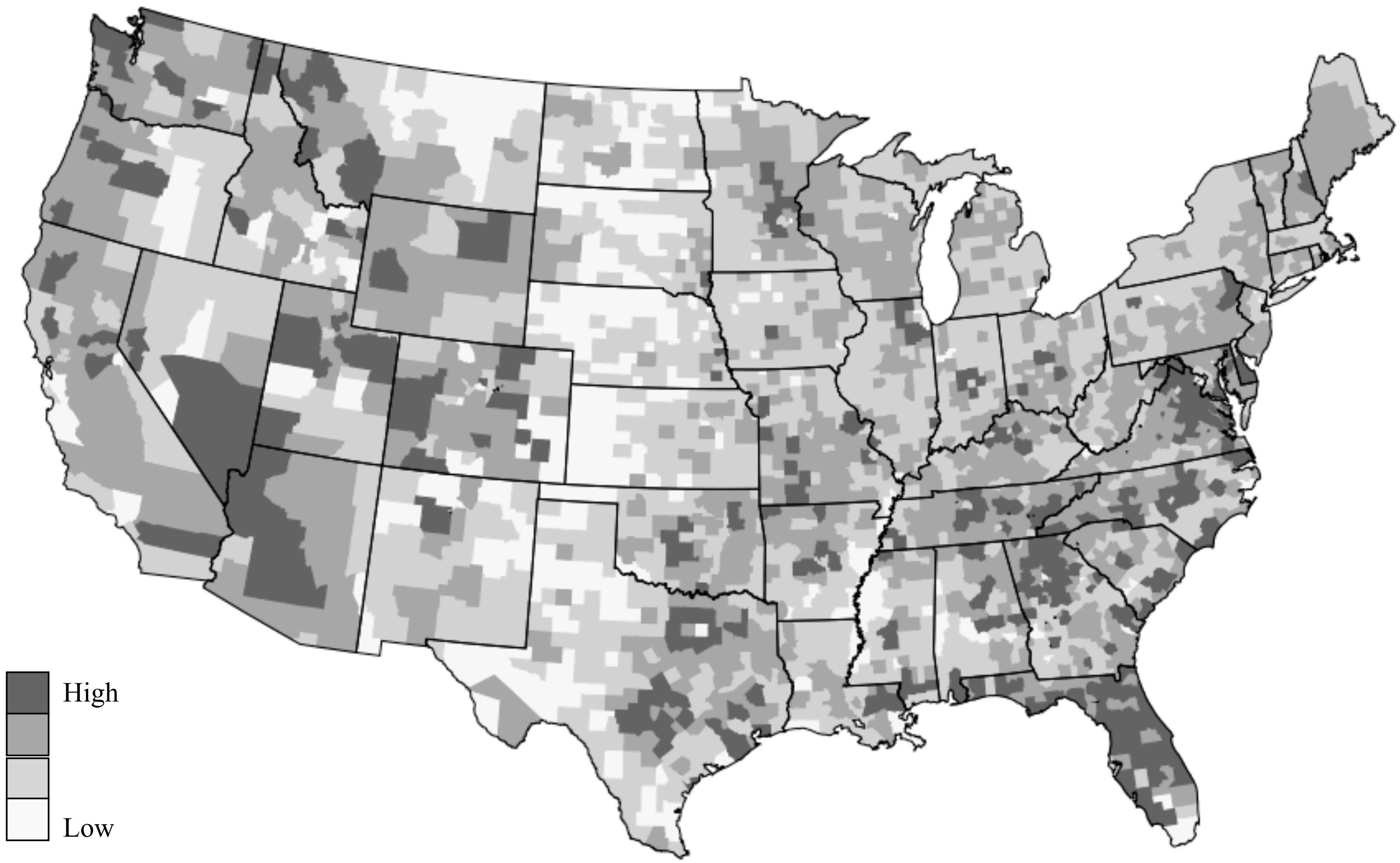


Figure 3.2 Spatial Distribution of Hazard Risk
[Note: Shading values vary by hazard variable]

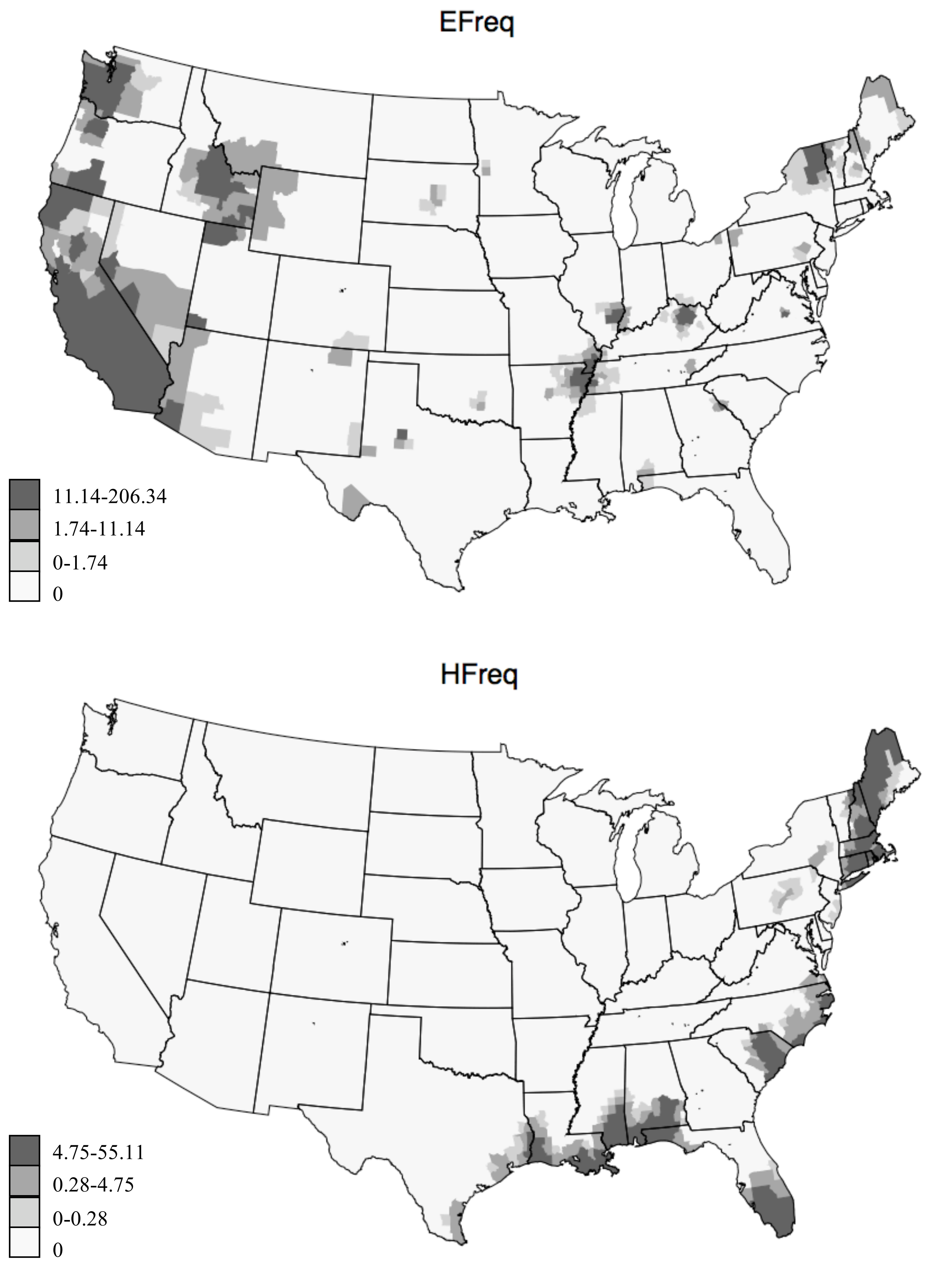
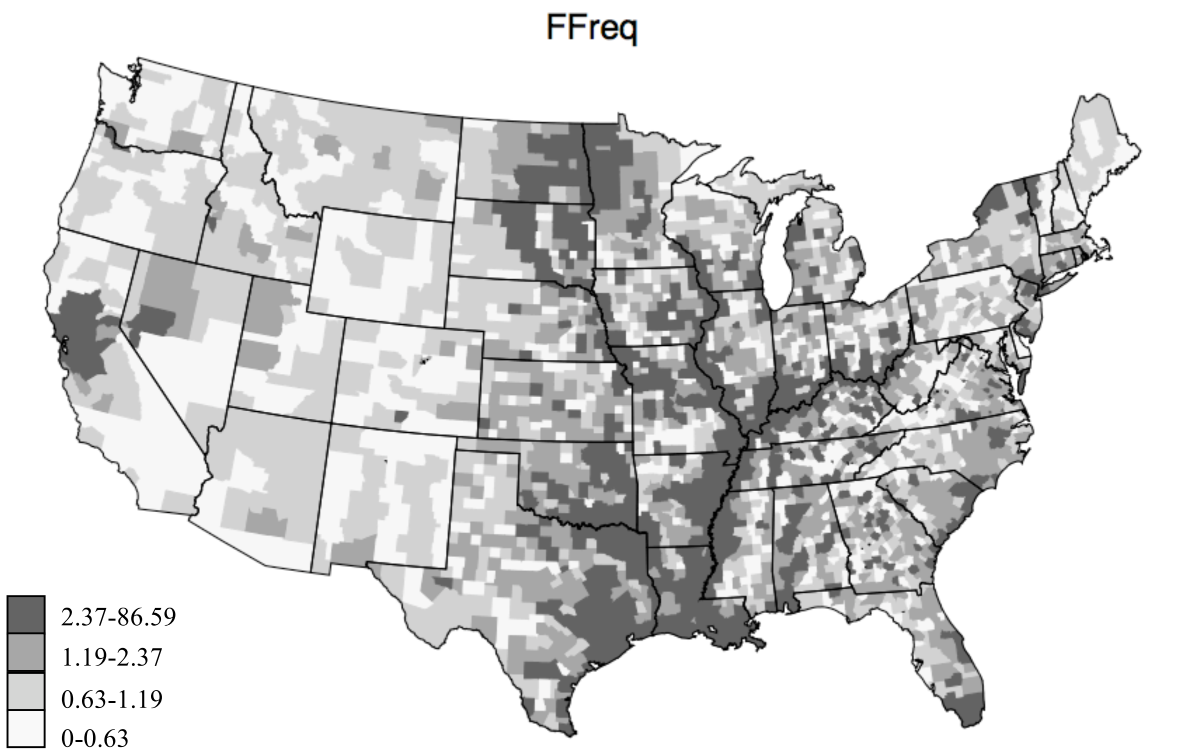


Figure 3.2 (cont.) Spatial Distribution of Hazard Risk
[Note: Shading values vary by hazard variable]



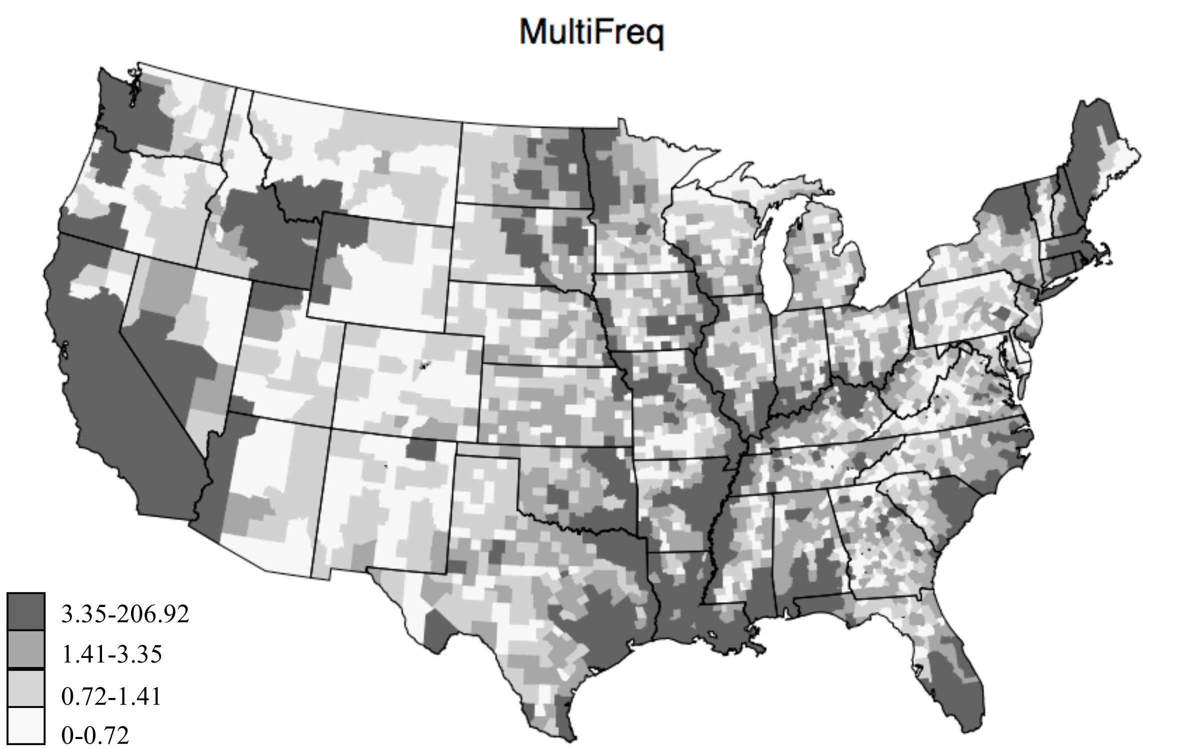
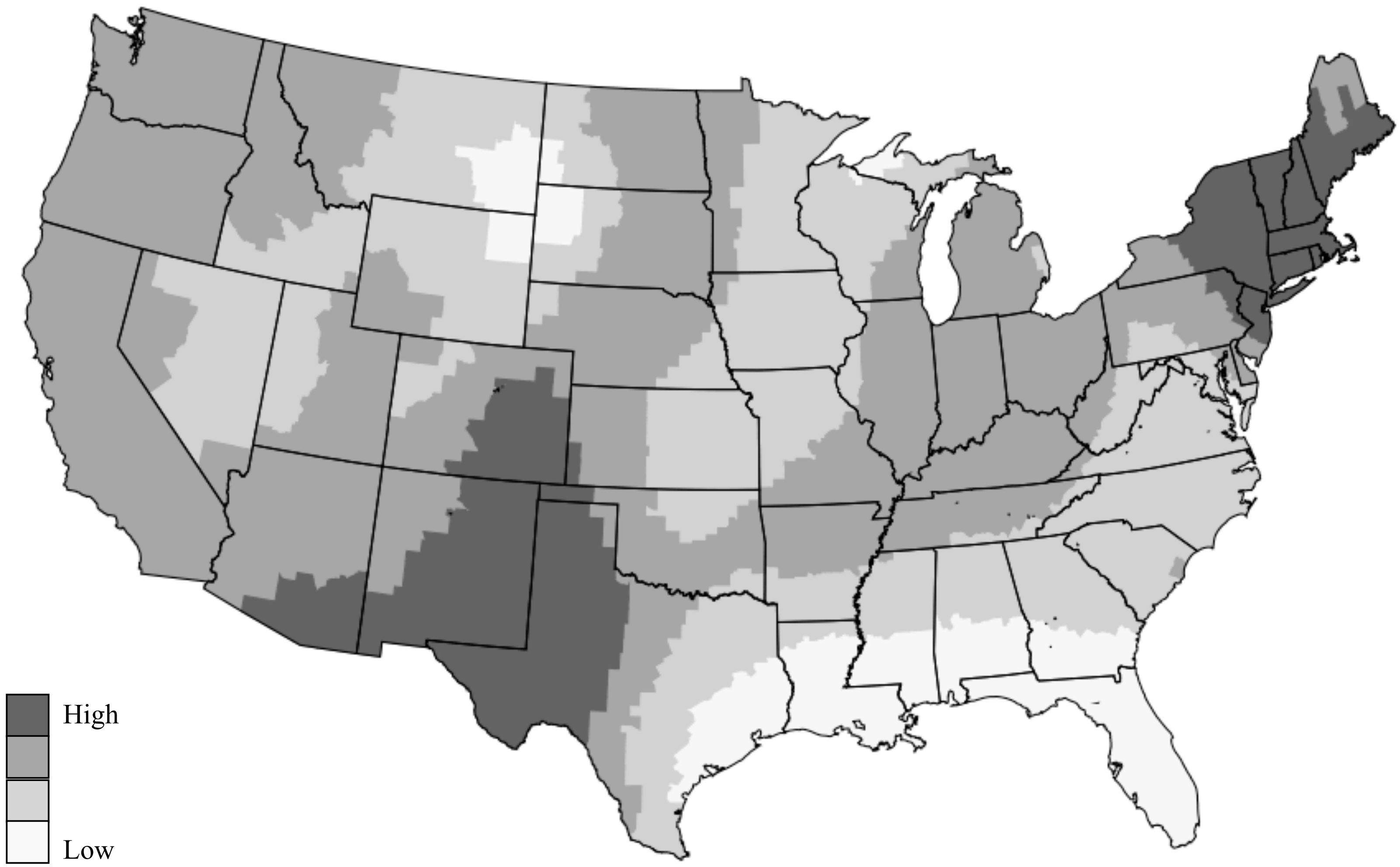


Figure 3.3 Distribution of GWR Coefficients for MultiFreq: Standard Deviation from Mean
[Note: Lightest color denotes values <1 s.d. below mean and darkest color denotes values >1 s.d. above mean.]



Chapter 4: Determinants of Mental Health and Displacement Following Hurricanes Katrina and Rita

4.1 Introduction

In 2005 Hurricanes Katrina and Rita devastated parts of Mississippi, Alabama, and Louisiana. The most severe damage occurred in New Orleans, where storm surge from Hurricane Katrina breached multiple levees flooding 80% of the city and rendering many neighborhoods uninhabitable (Beaudoin, 2007). Other Gulf Coast communities were damaged due to intense winds and heavy rainfall. Ultimately, Hurricane Katrina, which is often cited as the worst natural disaster in recent U.S. history, was responsible for 1500 deaths, mass displacement, and an estimated $108 billon in property damage (Beaudoin, 2007; Knabb et al., 2011). Hurricane Rita, while less destructive, was responsible for $11 billion in damages (National Hurricane Center, 2007). In the aftermath of these hazards, survivors experienced a myriad of physical and psychological stressors. These stressors included increased exposure to disease, inadequate access to basic necessities (e.g. water, food, medical care), isolation, and prolonged social and community disruption. Coupled with the loss of income and property, these stressors posed a substantial risk to the health and wellbeing of hurricane victims. Quantifying the adverse effects of these stressors, and identifying the factors that reduce or enhance these effects, is essential to mitigating their impact.

This study conducts two separate analyses regarding the wellbeing of individuals affected by Hurricanes Katrina and Rita. The first analysis evaluates long-term mental health status. A system of simultaneous equations is developed in order to identify the determinants of posttraumatic stress, depression, and anxiety disorder. Mental health status is assumed to vary according to each individual's personal experiences. The more distressing the individual's experiences the more likely they are to suffer from posttraumatic stress, depression, or anxiety. Variation in personal experiences is captured using the post-disaster vulnerability index (PDVI). This index measures the vulnerability of survey respondents to major stressors in the immediate aftermath of Katrina and Rita. These stressors include displacement, property damage, food shortages, water shortages, exposure to unsanitary conditions, and electricity shortages. In addition to post-disaster vulnerability, emphasis is placed on the role of social support in reducing adverse mental health outcomes.

The second analysis evaluates household displacement. Thousands of households were displaced as a result of Hurricanes Katrina and Rita. As with the loss of life and property damage, displacement was most pronounced in New Orleans where mandatory evacuation reduced the city's population to a few thousand residents (Fussell et al., 2010). Following the hazard event, the city's population gradually increased—reaching 210,00 in mid-2006, 288,000 in mid-2007, and 312,000 in mid-2008 (Fussell et al., 2010). The mid-2008 estimates are approximately 68% of the pre-Katrina population. Previous studies have identified significant correlations between disaster-related displacement and adverse health and economic outcomes. Displaced households experience greater incidence of mental illness, have reduced access to primary healthcare, have lower incomes, are more likely to be unemployed, and are less likely to be homeowners (Hori and Schafer, 2010). This analysis, taking a slightly different approach, evaluates the determinants of displacement and displacement duration. Specifically, a hurdle model is used to simultaneously estimate the relationship between household characteristics, displacement, and displacement duration. As in the mental health analysis, emphasis is placed on the role of social support. For the displacement analysis social support is decomposed into emotional support, financial support from relatives, and housing support from relatives (i.e. living with relatives while displaced). Results from both the mental health and displacement analysis offer insight into the impact of post-disaster vulnerability and social support on adverse hazard-related outcomes.

The remainder of this chapter is divided into two major sections: the mental health analysis (Section 4.2) and the displacement analysis (Section 4.3). Each section consists of several sub-sections, including a discussion of relevant literature, description of the data, presentation of the empirical model, and results. A joint discussion and conclusion are provided in Section 4.

4.2 Mental Health Analysis

4.2.1 Literature Review (Mental Health)

Previous studies have identified a correlation between disasters and long-term physical, behavioral, and mental health outcomes. Hazard events have been linked to higher incidence of cardiovascular disease (Trevisan et al, 1986; Kario et al., 2003; Steinberg et al., 2004), strained interpersonal relationships (Norris et al., 2002) and mental illness (Norris et al., 1993; Norris et al., 2002; Galea et al. 2003; Cao et al., 2003; Qouta et al., 2003; Bödvarsdóttir and Elklit, 2004; Norris et al., 2004; Miguel-Tobal et al., 2006; Galea et al., 2007; Galea et al., 2008; Adeola, 2009). Among children, disasters have also been linked to higher incidence of adjustment disorders and academic difficulties (Madrid et al., 2006). Specific to mental health, studies have found a strong positive correlation between posttraumatic stress disorder (PTSD) and exposure to a variety of hazard-types, including earthquakes (Cao et al., 2003; Bödvarsdóttir and Elklit, 2004), flooding (Norris et al., 2004), hurricanes (Norris et al., 1993; Galea et al., 2007; Galea et al., 2008; Adeola, 2009), and terrorist attacks (North et al., 1999; Qouta et al., 2003; Miguel-Tobal et al., 2006).

A few studies specifically evaluate the mental health effects of Hurricane Katrina. Beaudoin (2007) assesses the role that news information and social capital, defined by the frequency and nature of social interactions, had on the occurrence of depression and illness. Data for the analysis was obtained from a survey of hurricane-shelter residents a few weeks after the disaster occurred. Results suggest that depression is more common among respondents who relied heavily on news information and respondents with relatively little social capital. Likewise, results indicate a positive relationship between illness and reliance on news information.

Gelea et al. (2007) estimate the prevalence of PTSD and anxiety disorder among Katrina-affected individuals in March 2006 (i.e. six month after the event). Findings indicate a high occurrence of both PTSD (16.3%) and anxiety disorder (31.2%). Logistic regression techniques are used to examine the relationship between mental health indicators and various explanatory variables. These variables include socioeconomic characteristics, exposure to traumatic events, financial loss, and disaster-related stressors. Regression results indicate significant correlation between disaster-related stressors and the likelihood of being diagnosed with PTSD or anxiety disorder. These findings are reinforced by Gelea et al. (2008), which evaluates the impact of Hurricane Katrina on the mental health status of Mississippian households between March and September 2007 (i.e. between eighteen and twenty-four months after the event). As before, regression results indicate significant correlation between disaster-related stressors and the likelihood of being diagnosed with PTSD. Moreover, results suggest that ongoing stressors and traumatic events (i.e. lower income, unemployment, housing damage, and displacement) are central determinants of long-term PTSD.

Adeola (2009) evaluates the factors affecting acute (short-term) and chronic (long-term) stress using survey data from households requesting Red Cross assistance. Data is obtained from two telephone surveys conducted by The Gallup Organization in collaboration with several partner organizations. The first survey was conducted in the weeks following Hurricane Katrina and a follow-up, of the same households, was conducted eleven months later. Respondents who were female, unemployed, and had lower incomes exhibited higher levels of acute stress. Respondents who were female, had been evacuated from their home, and had incurred housing damage exhibited higher levels of chronic stress. Findings also suggest significant correlations between mental health and financial support from friends and family. In particular, respondents who received financial support experienced higher levels of acute and chronic stress. Adeola (2009) suggests this finding may reflect the additional stress of relying on friends and family who were also affected by the hazard event. It may also reflect a general discomfort with receiving financial aid from friends or family. Finally, results show that aid from state and local governments had no significant impact on stress levels.

4.2.2 Data (Mental Health)

Data for the mental health analyses is largely obtained from the 2005 and 2007 PSID. The PSID is an ongoing longitudinal survey of U.S. individuals and family units. The survey is conducted by the Institute for Social Research at the University of Michigan and is primarily funded by the National Science Foundation. Respondents are interviewed every two years (1997 to present) on a wide range of household characteristics and behaviors—including employment status, income, health, wealth, education, housing, expenditures, marital and fertility behavior, and philanthropy. With respect to health the survey contains information on health status, the onset of health conditions, health behaviors (e.g. smoking and exercise), health insurance coverage, health expenditures, and mental health status.

The 2005 PSID included approximately five hundred families that resided in Louisiana, Alabama, or Mississippi at the time of Hurricanes Katrina and Rita. In 2007, a supplemental questionnaire was administered to these families regarding the physical, psychological, and economic impacts of the hurricanes. These questions are used to construct social support and post-disaster vulnerability variables. A detailed discussion of these variables is presented in Sections 4.2.3.3 and 4.2.3.4. The supplement also contains well-established instruments for identifying PTSD, depression, and anxiety disorder. These instruments are used to construct dependent variables for the mental health analysis.

In addition to the PSID, this analysis employs publicly available data from FEMA and the Hurricane Katrina and Rita Clearinghouse Cooperative. Data obtained from FEMA is used to determine the extent of hurricane-related damage within census tracts.

FEMA maps indicate areas that were flooded or classified as damaged by the two hurricanes. These maps are used, in conjunction with GIS, to calculate the percent of each census tract that was flooded or classified as damaged. The Hurricane Katrina and Rita Clearinghouse Cooperative, maintained by Louisiana State University, is comprised of several hurricane-related databases. For this analysis, information is obtained on the estimated cost of building replacement. These estimates indicate the total cost of repairing damage to the commercial, residential, industrial, and governmental buildings within each zip code.

4.2.3 Empirical Model and Index Variables (Mental Health)

4.2.3.1 Empirical Model (Mental Health)

This study employs a simultaneous equation model to estimate the effects of post-disaster vulnerability and social support on mental health outcomes, after controlling for relevant socioeconomic and behavioral characteristics. Equations 4.1 and 4.2 depict the system of equations.

$$MH_i = f(E_i, B_i, SS_i, PDVI_i) \qquad (4.1)$$

$$PDVI_i = f(E_i, B_i, SS_i, DS_i) \qquad (4.2)$$

Here *MH* is a binary variable indicating whether respondent *i* is diagnosed with PTSD, depression, or anxiety disorder. This diagnosis is a function of the respondent's socioeconomic characteristics (*E*), behavioral and health attributes (*B*), social support (*SS*), and post-disaster vulnerability (*PDVI*). The PDVI is a function of socioeconomic characteristics (*E*), behavioral and health attributes (*B*), social support (*SS*), and disaster severity (*DS*). Disaster severity measures the extent of hurricane damage within the respondent's area of residence—accounting for the relationship between localized hazard damage and the respondent's post-disaster vulnerability. Previous analyses estimate the relationship between hazard events and mental health using a series of stress indicators. Examples of these stressors include: presence at the time of the event, displacement, housing damage, and loss of relative or friend. Use of the PDVI is a departure from these analyses. The PDVI emphasizes the compounding effects, rather than the singular effects, of post-disaster stressors. This view of vulnerability is supported by a number of recent studies, including Morrow (1999), Jaspers et al. (1999), and Curtis et al. (2007). In addition, the PDVI deemphasizes spatial proximity to the hazard event. Instead it posits that mental heath outcomes depend on the hazard's direct impact on the respondent's health, property, and social status. As indicated in the simultaneous equation model, spatial proximity indirectly affects mental health through the disaster severity measure.

Mental health is modeled as a function of the individual's current socioeconomic, behavioral, and health characteristics. Socioeconomic variables include the respondent's age (AGE), educational attainment (EDU), household income (HHINC) and dummy variables indicating whether the respondent is female (FEMALE), black (BLACK), married (MARRIED), or unemployed (UNEMPLOY). Also included are indicators for homeownership (OWNHOME) and the presence of children in the household (CHILD). Behavioral and health variables indicate whether the respondent smokes cigarettes (SMOKE), drinks alcoholic beverages on a regular basis (DRINK), is physically inactive (INACTIVE), has been diagnosed with a chronic health problem (CHRONIC), or has experienced significant trauma prior to Hurricane Katrina or Rita (TRAUMA). The remaining independent variables (e.g. the PDVI, social support measure, and disaster severity measure) are constructed indexes. A detailed discussion of these variables, as well as mental health instruments, is proved below. Summary statistics for all variables are presented in Table 4.1.

4.2.3.2 Mental Health Measures (Mental Health)

The supplemental questionnaire to the 2007 PSID contains three widely used instruments for diagnosing mental health: the PTSD Checklist (civilian version), the Patient Health Questionnaire (PHQ-9), and the Generalized Anxiety Disorder Survey (GAD-7). All three of these measures are employed as dependent variables. The PTSD Checklist is a 17-item inventory of PTSD-related symptoms. Each of the 17 items is scored using a Likert scale from 1 (not at all) to 5 (extremely). The sum of these items produces a score, ranging between 17 and 85, which represents the respondent's level of post-traumatic stress. Following Smith et al. (1999) and Walker et al. (2002) two diagnostic cutoffs are used to identify respondents with PTSD.[41] Under the first definition, respondents with a checklist score greater than or equal to 50 are classified as having PTSD (PTSD50). As evident from Table 4.1, only 3% of the sample population exhibits symptoms of PTSD using this definition. Under the second definition, respondents with a score greater than or equal to 30 are classified as having PTSD (PTSD30). Based on this definition, 15% of the sample exhibits symptoms of PTSD.

The PHQ-9, developed by Kroenke et al. (2001), is a 9-item inventory of depression related symptoms. Each of the 9 items is scored on a scale from 0 (not at all) to 3 (nearly every day). The sum of these items produces a score, ranging from 0 to 27,

---

[41] Diagnostic cutoffs convert continuous mental-health measures into binary variables. This conversion is employed, despite the loss of information, in order to maintain consistency with prior research.

which represents the respondent's level of depression. Following recommendations by Kroenke et al. (2001) diagnostic cutoffs of 10 and 5 are used to identify respondents with moderate and mild symptoms of depression. Approximately 6% of the sample exhibits symptoms of depression using the moderate cutoff value (PHQ10) and 11% using the mild cutoff value (PHQ5).

Spitzer et al. (2006) develop a 7-item inventory of anxiety-related symptoms, referred to as the GAD-7. Each of the 7 items is scored from 0 (not at all) to 3 (nearly every day). The sum of these items provides a score, ranging from 0 to 21, which represents the respondent's level of anxiety. Similar to PHQ-9, and following the recommendations by Spitzer et al. (2006), diagnostic cutoffs of 10 and 5 are used to identify respondents with moderate and mild symptoms of anxiety. Approximately 5% of the sample exhibits symptoms of depression using the moderate cutoff value (GAD10) and 12% using the mild cutoff value (GAD5). Although depression and anxiety are closely related, there is evidence that the PHQ-9 and GAD-7 measure distinct dimensions of mental health. Spitzer et al. (2006), for instance, find that depression and anxiety have differing effects on functional impairment and disability.

4.2.3.3 Post-Disaster Vulnerability Index (Mental Health)

Previous studies have employed various definitions of vulnerability to describe the potential for disaster-related loss for a particular group or household (Adger, 1999). In this analysis, using a slightly different construct, vulnerability refers to an individual's susceptibility to adverse economic, health, and social outcomes following the hazard event. The PDVI is designed to provide a single measure of post-disaster vulnerability. The index is constructed using a principle component analysis and six measures obtained from the 2007 PSID. Specifically, the PDVI is calculated using the formula presented in Equation 4.3.

$$A_{1f} = \frac{f_{11}(a_{1j} - a_1)}{s_1} + \frac{f_{12}(a_{2j} - a_2)}{s_2} + \ldots + \frac{f_{1n}(a_{nj} - a_n)}{s_n} \qquad (4.3)$$

Here A represents the first principle component, a_{nj} a set of n vulnerability measures for respondent j, a_n the sample mean, s_n the standard deviation, and f_{1n} is the inverse of the eigenvectors—which are scaled such that the sum of squares equals one. A detailed explanation of the principle component method can be found in Filmer and Pritchett (2001). Measures used to construct the PDVI pertain to the length of the displacement, personal property damage, food shortages, water shortages, unsanitary conditions, and electricity loss experienced following the hazard event. Each measure is scored on a scale of 1 (not displaced/not at all) to 4 (long-term displacement/a lot). The survey questions used to construct the PDVI are reproduced in Appendix G. The PDVI ranges from -1.94 to 4.87, with a mean of 0 (by construction) and a standard deviation of 1.77.

4.2.3.4 Social Support Index (Mental Health)

Several studies find negative and significant correlations between an individual's level of social support and the probability of being diagnosed with PTSD (Galea et al., 2008; Galea et al., 2007; Miguel-Tobal et al., 2006). Social support is incorporated into this analysis using the social support index (SSI) included in the PSID hurricane supplement. This index is a variant of the Crisis Support Scale developed by Joseph et al. (1992). It has been used in several trauma and disaster studies, including those evaluating earthquakes (Bödvarsdóttir et al., 2004), hurricanes (Galea et al., 2008), abuse (Bal et al., 2009), and premature births (Elkit et al., 2007). Elkit et al. (2007) note that the index has demonstrated good internal consistency and good discriminatory power. The index employed by the PSID is a 6-item inventory of social interactions that occurred in the three months following the hazard event. These questions, which are not specific to the disaster event, aim to proxy the quantity and quality of social interaction. Each item is scored using a scale from 1 (never) to 7 (always). The sum of these items produces a score, ranging from 7 to 42, that represents the respondent's level of social support. For respondents included in this analysis, the average social support score is 27.0 and the standard deviation is 9.0.

4.2.3.5 Disaster Severity Index (Mental Health)

As previously mentioned, disaster severity measures the extent of hurricane damage within the respondent's area of residence. Two variables are used to control for disaster severity: a disaster severity index (DSI) and distance to the nearest hurricane path (DISTANCE). The severity index is a composite of three variables: total building replacement costs per capita by zip code (i.e. the cost of repairing and replacing buildings damaged by Katrina and Rita), percentage of census tract classified as flooded by Katrina and Rita, and percentage of census tract classified as damaged by Katrina and Rita. To construct the index each variable is standardized using the formula given in Equation 4.4.

$$z_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{4.4}$$

Here z represents the standardized value for respondent i, x the variable of interest (i.e. building replacement costs, percent of census tract flooded, percent of census tract damaged), x_{min} is the minimum value of x, and x_{max} the maximum of x. Once standardized, the three variables are summed to create the DSI. The DSI ranges from 0 to 2.83, with a mean value of 0.89. The distance from the respondent's census tract (i.e. the centroid of the census tract) to the nearest hurricane path is also included as a measure of disaster severity. It is hypothesized that hurricane damage decreases with distance from the hurricane path.

4.2.3.6 Maximum Likelihood Function (Mental Health)

Parameter estimates are obtained using full-information maximum likelihood techniques. An econometric version of the model is presented in Equations 4.5 and 4.6.[42]

$$MH_i = X_i\alpha + \varepsilon_i \tag{4.5}$$

$$PDVI_i = Z_i\beta + \mu_i \tag{4.6}$$

Explanatory variables are denoted by the vectors X and Z, estimated parameters by the vectors α and β, and random error terms by ε and μ. The error terms are assumed i.i.d. normally distributed with mean 0, respective variances σ_{11} and σ_{22}, and covariance σ_{21}. The normalized covariance matrix is presented in Equation 4.7.

$$Var(\varepsilon,\mu) = \Omega = \begin{bmatrix} 1 & \sigma'_{21} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \tag{4.7}$$

The log likelihood function is derived from the joint density function $f(Y_i, PDVI_i \mid Z_i)$, which can be written as $f(Y_i \mid PDVI_i, Z_i)\, f(PDVI_i \mid Z_i)$. The log-likelihood function is presented in Equation 4.8.

$$\begin{aligned} \ln L_i &= MH_i \ln\Phi(m_i) + (1 - MH_i)\{1 - \Phi(m_i)\} + \ln\phi(PDVI_i|Z_i) \\ \ln\phi(PDVI_i|Z_i) &= -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln|\sigma_{22}| - \frac{1}{2}(PDVI_i - Z_i\beta)\sigma_{22}^{-1}(PDVI_i - Z_i\beta)' \\ m_i &= (1 - \sigma'_{21}\sigma_{22}^{-1}\sigma_{21})^{-\frac{1}{2}}\{X_i\alpha + (PDVI_i - Z_i\beta)\sigma_{22}^{-1}\sigma_{21}\} \end{aligned} \tag{4.8}$$

Here $\Phi(\bullet)$ denotes the cumulative normal distribution function, $\phi(\bullet)$ the normal density distribution function, and n the number of observations. In practice, the variance and

[42] The model is estimated using Stata's ivprobit command. The log-likelihood functions are modified from StataCorp LP (2009).

covariance parameters are not estimated directly. Instead the natural-log of variance and the inverse hyperbolic tangent of covariance are estimated.

4.2.4 Results (Mental Health)

Regression results for the simultaneous equations analysis of PTSD are presented in Table 4.2a (MH equation) and Table 4.2b (PDVI equation).[43] Results for the PHQ analysis (depression) are presented in Table 4.3a and 4.3b while results for the GAD analysis (anxiety) are presented in Tables 4.4a and 4.4b. Each table presents four model specifications. Model 1 regresses mental health status on the socioeconomic characteristics and the PDVI. Subsequent models incorporate behavior and health characteristics (Model 2) and the social support index (Model 3). These specifications use the narrower definition of mental disorder (i.e. PTSD50, PHQ10 and GAD10). Model 4 reestimates Model 3 using the broader definition of mental disorder (i.e. PTSD30, PHQ5 and GAD5).

The significance of socioeconomic, behavioral, and health characteristics vary across mental health measures and model specifications. Despite this variation, nearly all significant-variables exhibit the expected sign. Under Model 3, for example, the probability of being diagnosed with PTSD increases with respondent's age and whether the respondents is female, smokes cigarettes, and experienced a traumatic event prior to Katrina or Rita. The probability of being diagnosed with PTSD decreases with unemployment, household income, homeownership, and physical inactivity. Many of these variables, however, are insignificant under the broader definition of PTSD.

---

[43] Stata codes for the mental health and displacement duration analyses are provided in Appendix I.

According to Model 4, the probability of being diagnosed with PTSD is negatively correlated with educational attainment, marriage or cohabitation, and the presence of children in the household.

Slightly more consistent results are observed for measures of depression and anxiety disorder. PHQ10 is positively and significantly correlated with whether the respondent is female, black, and diagnosed with a chronic health problem; it is negatively correlated with educational attainment, marriage or cohabitation, and homeownership. PHQ5 is positively correlated with household income and physical inactivity. It is negatively correlated with educational attainment, marriage or cohabitation, and children in the household. With respect to anxiety disorder, GAD10 is positively correlated with respondent's age, whether the respondent is black, and physical inactivity. There is a negative correlation with homeownership. GAD5 is negatively correlated with educational attainment and marriage or cohabitation. Despite inconsistencies across mental health measures and diagnostic cutoffs, these results demonstrate the importance of socioeconomic and behavioral characteristics on adverse mental health outcomes. It is important to note, however, that previous studies often find more robust correlations between mental health and socioeconomic characteristics (Galea et al., 2008; Migueal-Tobal et al., 2006). Moreover, studies that have evaluated multiple mental health measures typically find a greater degree of commonality among significant explanatory variables (Averina et al., 2005).

Coefficients for the PDVI are consistently positive and significant, implying that respondents with higher levels of post-disaster vulnerability are more likely to be diagnosed with PTSD, depression, and anxiety disorder two years after Katrina and Rita.

The robustness of this result suggests that post-disaster vulnerability is a crucial indicator of long-term mental health. In addition, consistent with Galea et al (2008) and Galea et al. (2007), there is evidence of an inverse relationship between social support and adverse mental health outcomes. In particular, findings indicate that the probability of being diagnosed with symptoms of depression or anxiety disorder decrease as the social support index increases, all else equal. There is no significant correlation between PTSD and social support.

Results for PDVI regressions are consistent with expectations. Post-disaster vulnerability is significantly correlated with a number of respondent characteristics, including racial classification, educational attainment, marriage or cohabitation, household income, and chronic health conditions. As expected, the PDVI is positively and significantly correlated with severity of damage within the respondent's area of residence, as measured by the DSI, and negatively correlated with distance to the nearest hurricane path.

4.3 Displacement Analysis

4.3.1 Literature Review (Displacement)

Few empirical analyses have been conducted regarding the determinants of disaster-related displacement and return migration. Recent studies conducted within the U.S. focus on displacement caused by Hurricane Katrina. Landry et al. (2007) evaluate return migration decisions for Katrina-displaced households from Louisiana and Texas. Among households originating from Louisiana, the probability of return is greater for those with higher income, and lower for senior citizens and residents of metropolitan New Orleans. Among households from Texas, the probability of return is higher for homeowners, those who are married, and those employed prior to Katrina. The probability of return is lower for those with greater educational attainment. Groen and Polivka (2010) extend this research using information obtained from the Current Population Survey. They find significant correlation between the probability of return and several demographic variables, including age, racial classification, gender, educational attainment, and income. More specifically, the probability of return is positively correlated with age, educational attainment, income and being male. It is negatively correlated with being black or Hispanic. In addition, they find a significant and negative correlation between the probability of return and the amount of damage in their area of residence. Fussell et al. (2010), adopting a different approach, evaluate the duration of displacement using a proportional hazard model. Findings show that the severity of housing damage and age categories are significant predictors of displacement duration. Surprisingly, the hazard rate is not significantly correlated with the remaining socioeconomic characteristics of the household (i.e. race, gender, education, employment and homeownership).

Hori and Schafer (2010) evaluate the effects of displacement, rather than the determinants, on housing, economic, and health outcomes. The analysis is based on a survey of Louisianan households conducted between June and December 2006 (i.e. 10 to 16 months after Hurricane Katrina). It distinguishes between three groups based on their place of residence at the time of the survey: non-displaced households, internal displacement (i.e. relocation within the initial parish), and external displacement (i.e. relocation outside the initial parish). Displaced individuals, compared to non-displaced, are less likely to be employed, own their home, live in detached housing or have access to primary health care. They also have lower incomes and are more likely to exhibit signs of serious mental illness. Results suggest the adverse effects of displacement are greater for individuals who are externally displaced than those internally displaced.

4.3.2 Data (Displacement)

As with the mental health analysis, data for the displacement analyses is obtained from the 2005 and 2007 PSID. Information obtained from the 2007 supplemental questionnaire is used to construct displacement, duration of displacement, and housing damage variables. Indexes regarding post-disaster vulnerability, social support, and disaster severity in the household's area of residence are identical to those employed in the mental health analysis. Among households included in this analysis, nearly 26% were displaced as a result of Hurricanes Katrina and Rita. The duration, measured in days, ranges from 1 to 837, with a mean of 141.9. Approximately, 28% of the sample experienced minimal housing damage, 16% moderate housing damage, and 12% severe housing damage. The remaining 44% experienced no housing damage.

4.3.3 Empirical Model (Displacement)

A hurdle model is used to evaluate displacement. Hurdle models are modified survival (or count) models that assume different processes govern zero outcomes and positive outcomes. More specifically, it is assumed that a binary process determines whether an outcome is zero or positive while a truncated-at-zero survival process determines the value of positive outcomes. Hurdle models have been used within a variety of contexts, including migration frequency (Bohara and Krieg, 1996), demand for bass fishing (Bilgic and Florkowski, 2007) and repeat instances of self-harm (Bethell et al., 2010). In this analysis, the hurdle model evaluates the determinants of displacement (i.e. the binary component) and displacement duration (i.e. the survival component). The hurdle model is made operational using the probit distribution for the binary component and the Weibull distribution for the survival component. The Weibull distribution is a two-parameter distribution commonly employed in parametric survival analysis (Greene, 2003).[44] It assumes a monotonic hazard function and encompasses, as special cases, the

[44] The survival component of this analysis is referred to as an accelerated failure time (AFT) model. These models, which are fully parametric, offer an alternative to the more widely used proportional hazard models, such as the semi-parametric Cox model. Under a proportional hazard model, the effect of a covariate is to multiply the hazard by some constant. In contrast, under an AFT model the effect of a covariate is to multiply the predicted event time by some constant (Greene, 2003). The straightforward interpretation of coefficients is a key advantage of the AFT model.

exponential and Rayleigh distribution. Validity of the regression results depends on proper selection of the survival distribution. The Weibull distribution is selected from among several possible distributions through visual examination of the survival function and comparison of AIC values. The Kaplan-Meier estimator is used to visually examine the survival function.[45] Results, presented in Figure 4.1, imply that over time households gradually return home—but at a decreasing rate. This pattern is indicative of a Weibull, generalized gamma, and log-logistic distribution. The AIC is used to distinguish between these distributions. While values are similar across possibilities and model specifications, the Weibull distribution consistently exhibits the lowest AIC.[46]

The joint log likelihood function of the hurdle-Weibull model is presented in Equation 4.9. A detailed derivation this function is provided in Appendix H.

$$
\begin{aligned}
\ln L = {} & y_{1j} \ln\left[\Phi(-x_{1j}\beta_1)\right] + (1 - y_{1j}) \ln\left[\Phi(1 - x_{1j}\beta_1)\right] \\
& - \left(\frac{y_{2j}}{\lambda}\right)^{\kappa} + y_{3j}\left(\ln(\kappa) - \ln(\lambda) + (\kappa - 1)\left[\ln(y_{2j}) - \ln(\lambda)\right]\right)
\end{aligned}
\tag{4.9}
$$

[45] The Kaplan-Meier estimator is a nonparametric maximum likelihood estimate of the survival function. With large samples the estimator approaches the true survival function. See Kaplan and Meier (1958) for a detailed description of the estimator.

[46] The generalized gamma distribution is highly flexible and encompasses the exponential, Weibull, and lognormal distributions. As a result, formal tests of the parameter values provide a subsequent method of identifying the proper parametric distribution. Results from the generalized gamma distribution reject the exponential and lognormal distributions but not the Weibull distribution—confirming findings from the AIC comparison.

Here $y1$, $y2$, $y3$ are dependent variables that respectively indicate whether the household j was displaced, the duration of displacement, and whether the household is right-censored (i.e. has yet to return home after being displaced). The term Φ represents the cumulative normal distribution function, κ the shape parameter of the Weibull distribution, λ the scale parameter of the Weibull distribution, x_1 the vector of explanatory variables that determine the binary process, and β_1 the corresponding vector of estimated parameters. Determinants of displacement duration are incorporated into the model through the scale parameter, which, along with the shape parameter, is positively restricted by log-link functions. The functions are presented in Equations 4.10 and 4.11:

$$\kappa = \exp(\gamma) \tag{4.10}$$

$$\lambda = \exp(\eta) = \exp(x_{2j}\beta_2) \tag{4.11}$$

where, x_2 is the vector of explanatory variables that determine the survival process and β_2 is the corresponding vector of estimated parameters. Maximum likelihood estimates are obtained using a modified Newton-Raphson algorithm.

The probability of displacement and duration of displacement are modeled as a function of housing damage, community damage, socioeconomic characteristics, behavior and health characteristics, and social support. Equations 4.12 and 4.13 present these functions in general notation.

$$x_{1j} = f(H_j, DS_j, E_j^h, B_j^h, SS_j) \tag{4.12}$$

$$x_{2j} = f(H_j, DS_j, E_j^h, B_j^h, SS_j) \tag{4.13}$$

Here H is a vector of dummy variables indicating whether housing damage for household j is minimal (MIN_DMG), moderate (MOD_DMG), or severe (SVR_DMG). The base category is households that had no hazard-related housing damage. As in the mental health analysis, *DS* is a vector of disaster severity within the household's area of residence. It contains the disaster severity index (DSI) and distance to the nearest hurricane path (DISTANCE). It is assumed that displacement and return migration decisions depend on socioeconomic characteristics of the household head at the time of the hazard event. It is important to note the difference between the socioeconomic variables used in the mental health and the displacement analyses. Socioeconomic variables used in the mental health analysis pertain to the survey respondent, regardless of the respondent's status in the household. Socioeconomic variables in the displacement analysis pertain to the household head. The head is often, but not always, the survey respondent. The superscript *h* is used to distinguish between the two sets of variables. These variables include the head's age (AGE^h), educational attainment (EDU^h), and dummy variables indicating whether the head is female ($FEMALE^h$), black ($BLACK^h$), married ($MARRIED^h$), unemployed ($UNEMPLOY^h$), or is in poor health ($UNHEALTHY^h$).[47] Also included is household income (HHINC) and indicators for homeownership (OWNHOME), mobile homes (MOBILE), and the presence of children (CHILD). Summary statistics for these variables are presented in Table 4.5.

Similar to the mental health analysis *SS* is a vector of social support variables. For the binary component of the displacement analysis, *SS* consists of the previously discussed SSI. In the mental health analysis, this index measures the respondent's level of social support in the months immediately following the hazard event. Within the current

[47] The PSID records the self-reported health status of household heads using a five-category Likert scale. Heads with a self-reported health status in the bottom two categories (i.e. fair and poor) are considered to have poor health.

context, however, this index takes on a slightly differently interpretation. It serves as a proxy for the strength of a household's social network. This network may be imposed upon to provide assistance in the time leading up to the hazard event, immediately following the event, and throughout the recovery process. For the survival component of the displacement analysis, *SS* consists of the SSI, the amount of insurance payments received during 2006 (INSURANCE), remittances from relatives in 2005 and 2006 (REMIT), and a dummy variable indicating whether the household lived with family members while displaced (DIS_FAMILY). Although not perfect measures of assistance, these variables do capture variation in social-based resources across household.

4.3.4 Results (Displacement)

Results for the hurdle-Weibull model are provided in Table 4.6a (binary component) and Table 4.6b (survival component). Four model specifications are presented in order to illustrate robustness of regression results. Model 1 includes the vector of housing damage and disaster severity variables. Subsequent models incorporate socioeconomic characteristics (Model 2), social support measures (Model 3), and state-fixed effects (Model 4). Fixed-effects, where Louisiana is the base category, are employed to capture variation in disaster response and aid across states. The AIC, also reported in Table 4.6b, is used to compare goodness-of-fit between model specifications.

Coefficient values and significance levels are largely consistent across the model specifications. Model 4 exhibits the lowest AIC and is the preferred specification. The probability of being displaced increases with housing damage and disaster severity. Regression coefficients for the binary component represent effects on a cumulative normal distribution, and are easily converted into probabilities. Converting the damage coefficients to probabilities, and evaluating other covariates at their mean, implies the probability of displacement increases by 18% with minimal housing damage. Likewise probability of displacement increases by 19% with moderate housing damage and 40% with severe housing damage. The likelihood of displacement also increases with damage to the area of residence. In particular, the probability of displacement increases by 18% for every unit increase in the DSI and 0.03% for every 10 miles closer to the nearest hurricane path.

Socioeconomic characteristics of the household head, with the exception of BLACK[h] in Model 4, have no apparent effect on the probability of being displaced. Among housing characteristics, residing in a mobile home prior to the hazard event is positively correlated with displacement. Specifically, mobile homes increase the probability of displacement by nearly 12%, after controlling for housing damage, residential area damage, and household income. The probability of being displaced is also positively and significantly correlated with the SSI. Assuming the SSI accurately measures the strength of social connections, this result implies that households with relatively stronger connections are more likely to be displaced.

Consistent with Fussell et al. (2010), results show significant correlation between the duration of displacement and housing damage. As an accelerated constant model, all respondents are assumed to follow the same baseline survival curve. The effect of covariates serves to accelerate or decelerate the rate of movement along the curve. Coefficients are interpreted as the change in the log of time (i.e. length of displacement) given a marginal change in the dependent variable. For instance, the coefficient for MOD_DAM is 1.14, indicating that the length of displacement is greater for households with moderate housing damage. In practice, the coefficients may be exponentiated to simplify their interpretation (Cleves et al., 2008). Taking the exponent of the MOD_DAM coefficient yields 3.15, which implies that moderate housing damage lengthens displacement by a factor of 3.15 compared to displaced households without damage (i.e. for every day households without damage are displaced, households with moderate damage are displaced 3.15 days, all else equal). Severe housing damage lengthens displacement by a factor of 6.13. Damage within the household's area of residence also increases the duration of displacement. A marginal increase in the DSI lengthens displacement by a factor of 3.67.

In contrast to the binary component and Fussell et al. (2010), socioeconomic characteristics of the household head significantly affect the duration of displacement. Duration is positively correlated with the head's age and educational attainment, and the presence of children in the household. It is negatively correlated with being married or cohabitating and residing in a mobile home prior to the hazard event. More specifically, for every ten years of age and educational attainment the length of displacement increases by a factor of 1.22 and 4.02, respectively. As a point of reference, the average household head is 41.62 years of age and has completed 12.5 years of education. The presence of children has a considerable impact on displacement. For every day households without children are displaced, households with children are displaced 3.33 days. Households with either cohabitating or married head's return to their homes more quickly, by a factor of 0.39, than those with unmarried heads. Likewise, households residing in mobile homes return more quickly, by a factor of 0.49, than those residing in single-family homes or apartments.

Displacement duration also depends on social support, where higher levels of support reduce duration. Households with relatively higher SSI, for example, return to their homes more quickly—conditional on having been displaced. Moreover, for every $1000 received in remittances the length of displacement decreases by a factor of 0.83. Insurance payments and staying with family while displaced have no significant impact. Finally, households residing in Mississippi returned to their homes more quickly than those residing in Louisiana.

The natural log of the shape parameter, denoted by $\ln(\kappa)$, is significantly different from 0 at standard significance levels in Models 1 and 2. It is significant at the $p<0.15$ in Models 3 and 4. This suggests that the shape parameter is significantly less than one and that the probability of retuning home after being displaced decreases over time. Because the rate of return decreases, rather than being constant, the Weibull distribution is preferred to the exponential.[48] This being said, results from Models 3 and 4 suggest that the survival component could be adequately modeled with the exponential distribution.

4.4 Discussion and Conclusions

This study conducts two separate analyses regarding the wellbeing of those affected by Hurricanes Katrina and Rita. The first analysis evaluates the effects of post-disaster vulnerability on long-term mental health status. Three measures of mental health are employed: PTSD, depression, and anxiety disorder. The second analysis employs a hurdle-Weibull model to estimate the relationship between household characteristics and displacement duration, conditional on being displaced. Displacement, which separates households from familiar social and institutional settings, has previously been associated

[48] The Weibull is equivalent to the exponential distribution when $p=1$.

with adverse outcomes. Consequently, displacement duration has important implications for the wellbeing of those affected by natural hazards. Both analyses are based on a previously unexplored database, namely the supplemental questionnaire to the 2007 PSID, and offer insights into the long-term effects of natural disasters in the US.

The mental health analysis is conducted using a simultaneous equations model. The first equation models the respondent's mental health status as a function of socioeconomic characteristics, behavioral and health attributes, social support, and post-disaster vulnerability. The latter is measured using the uniquely constructed PDVI, which quantifies the respondent's exposure to major stressors in the immediate aftermath of Katrina and Rita. The PDVI is continuous and allows for greater flexibility in econometric analysis than the binary variables employed in similar studies. The second equation models the PDVI as a function of disaster severity in the respondent's area of residence. Under this construction, mental health is not directly affected by the hazard event but rather the stress imposed on the respondent. Results from this analysis indicate a positive correlation between post-disaster vulnerability and the probability of exhibiting signs of PTSD, depression, or anxiety disorder. These results are robust across most model specification and suggest that Katrina and Rita have had a lasting impact, of at least two years, on the mental health status of affected individuals. Results also suggest, congruent with previous studies, that respondents with high levels of social support are less likely to be diagnosed with adverse mental health conditions.

The displacement analysis is conducted using a hurdle-Weibull model. Findings from this study are particularly relevant since few analyses have evaluated displacement duration using survival or hazard regression techniques. Unsurprisingly, housing and community damage are the most important predictors of displacement and displacement duration. Greater damage is associated with longer periods of displacement. Socioeconomic factors such as age, educational attainment, marital status, and the presence of children also have significant effects. In particular, duration is positively correlated with age, educational attainment, and the presence of children. It is negatively correlated with marriage or cohabitation. The most interesting findings pertain to the SSI, which is positively correlated with the probability of being displaced and negatively correlated with the length of displacement. This suggests that households with stronger social networks rely on their connections to provide safe accommodations during the hazard event. Once the hazard occurs, these networks provide support to the displaced household, allowing them to return more quickly. Direct financial support from relatives is shown to reduce the length of displacement.

Providing effective assistance in the aftermath of major hazard event is a difficult but important task. This work suggests that post-disaster vulnerability is a key determinant of mental health disorders. DRM interventions that reduce post-disaster vulnerability (i.e. facilitate return migration and mitigate food, water, and electricity shortages) could drastically improve short- and long-term mental health. To the extent possible, assistance designed to address lasting psychological effects of traumatic events, should be targeted towards those with high post-disaster vulnerability. One way of improving wellbeing, as evidenced in previous studies, is for households to return to familiar social and institutional settings. While research in this area is limited, results from this work indicate that social support reduces displacement duration. Accordingly, DRM interventions that foster social capital may facilitate return migration.

Table 4.1 Descriptive Statistics: Mental Health Analysis†

Variable	Description	Mean	Std. Dev.
PTSD50	Indicator of PTSD using a cutoff value of 50	0.03	0.171
PTSD30	Indicator of PTSD using a cutoff value of 30	0.151	0.358
PHQ10	Indicator of depression using a cutoff value of 10	0.058	0.234
PHQ5	Indicator of depression using a cutoff value of 5	0.114	0.318
GAD10	Indicator of anxiety disorder using a cutoff value of 10	0.053	0.225
GAD5	Indicator of anxiety disorder using a cutoff value of 5	0.123	0.329
FEMALE	Gender (female=1, male=0)	0.71	0.454
BLACK	Race and ethnicity (black=1, otherwise=0)	0.8	0.4
AGE	Age (10 years)	4.195	1.421
EDU	Educational attainment (10 years)	1.268	0.214
MARRIED	Marital status (married or cohabitating=1, otherwise=0)	0.439	0.497
CHILD	Children in household (children=1, otherwise=0)	0.536	0.499
UNEMPLOY	Employment status (unemployed=1, otherwise=0)	0.086	0.28
HHINC	Household income in 2006 (ln$)	10.202	1.115
OWNHOME	Household tenure status (homeowner=1, otherwise=0)	0.568	0.496
SMOKE	Smokes cigarettes (yes=1, no=0)	0.227	0.42
DRINK	Regularly drinks alcohol (yes=1, no=0)	0.074	0.262
INACTIVE	Physically inactive (yes=1, no=0)	0.232	0.423
CHRONIC	Chronic health condition (yes=1, no=0)	0.181	0.385
TRAUMA	Experienced prior traumatic event (yes=1, no=0)	0.03	0.171
SSI	Social support index	2.697	0.902
PDVI	Post-disaster vulnerability index	0	1.77
DSI	Disaster severity index	0.147	0.428
DISTANCE	Distances to nearest hurricane path (10 miles)	7.315	4.325

†Socioeconomic, behavioral, and health characteristics refer to the respondent in 2007.

Table 4.2a PTSD Regression Results: Mental Health Status Equation

	Model 1: PTSD50	Model 2: PTSD50	Model 3: PTSD50	Model 4: PTSD30
Constant	-2.9267	-3.3818**	-4.0816**	-0.0393
	(1.788)	(1.675)	(1.800)	(0.978)
FEMALE	1.0648*	1.5074***	1.4350***	0.1937
	(0.630)	(0.554)	(0.528)	(0.213)
BLACK	-0.4624	0.1493	0.2876	-0.0649
	(0.524)	(0.519)	(0.585)	(0.250)
AGE	0.3122**	0.5623**	0.5992**	0.0080
	(0.154)	(0.232)	(0.264)	(0.070)
EDU	0.3768	0.8546	1.0558	-0.8504**
	(0.704)	(0.777)	(0.856)	(0.368)
MARRIED	0.2130	0.6223	0.5077	-0.4097**
	(0.419)	(0.427)	(0.397)	(0.204)
CHILD	-0.9083**	-0.8939*	-0.9597*	-0.3019*
	(0.385)	(0.496)	(0.552)	(0.171)
UNEMPLOY	-3.9330***	-3.8335**	-3.9970**	-0.1434
	(0.746)	(1.195)	(1.290)	(0.279)
HHINC	-0.1209	-0.3889**	-0.3604***	0.0616
	(0.174)	(0.135)	(0.135)	(0.093)
OWNHOME	-1.3585***	-1.6639**	-1.6976**	-0.1873
	(0.451)	(0.687)	(0.764)	(0.189)
SMOKE		0.9084***	0.9652***	-0.0082
		(0.303)	(0.331)	(0.188)
DRINK		1.0855*	1.0634	-0.5558
		(0.644)	(0.680)	(0.355)
INACTIVE		-0.8818*	-0.8729*	-0.0326
		(0.471)	(0.490)	(0.172)
CHRONIC		-0.2396	-0.3264	0.1683
		(0.532)	(0.549)	(0.219)
TRAUMA		1.8085***	1.8776**	0.1350
		(0.680)	(0.807)	(0.415)
SSI			-0.0322	-0.1010
			(0.251)	(0.097)
PDVI	0.2439	0.2745	0.3207	0.5407***
	(0.218)	(0.271)	(0.330)	(0.084)
atanh(Rho)	0.5723	0.8532**	0.7852*	-0.5548***
	(0.357)	(0.412)	(0.420)	(0.205)
ln(Sigma)	0.4069***	0.3934***	0.3781***	0.3782***
	(0.036)	(0.037)	(0.037)	(0.037)
N	418	418	418	418
AIC	1623.7175	1619.7635	1610.8042	1846.9797

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.2b PTSD Regression Results: PDVI Equation

	Model 1: PTSD50	Model 2: PTSD50	Model 3: PTSD50	Model 4: PTSD30
Constant	0.9677 (0.978)	0.7197 (0.950)	-0.2612 (0.989)	-0.2667 (0.990)
FEMALE	0.1482 (0.161)	0.1601 (0.161)	0.1546 (0.158)	0.1574 (0.158)
BLACK	0.3639** (0.175)	0.4023** (0.181)	0.4019** (0.180)	0.4073** (0.182)
AGE	0.0584 (0.073)	0.0276 (0.073)	0.0506 (0.071)	0.0493 (0.071)
EDU	0.9026** (0.393)	0.9253** (0.395)	0.9021** (0.391)	0.8940** (0.389)
MARRIED	0.3956** (0.183)	0.4083** (0.181)	0.4597*** (0.176)	0.4598*** (0.176)
CHILD	0.1642 (0.174)	0.1633 (0.176)	0.1432 (0.172)	0.1424 (0.172)
UNEMPLOY	0.1096 (0.300)	0.0409 (0.302)	0.0537 (0.299)	0.0556 (0.298)
HHINC	-0.2209** (0.087)	-0.1942** (0.085)	-0.1909** (0.085)	-0.1914** (0.085)
OWNHOME	-0.0307 (0.194)	-0.0625 (0.196)	-0.1005 (0.192)	-0.0971 (0.191)
SMOKE		0.2006 (0.184)	0.2409 (0.182)	0.2407 (0.182)
DRINK		-0.2543 (0.306)	-0.3023 (0.308)	-0.3021 (0.309)
INACTIVE		-0.2837 (0.197)	-0.2243 (0.193)	-0.2237 (0.193)
CHRONIC		0.5416** (0.215)	0.4854** (0.212)	0.4857** (0.212)
TRAUMA		0.0221 (0.303)	0.0066 (0.294)	0.0037 (0.294)
SSI			0.3025*** (0.079)	0.3026*** (0.079)
DSI	1.0191*** (0.204)	0.9984*** (0.199)	0.9064*** (0.199)	0.9264*** (0.180)
DISTANCE	-0.0129*** (0.002)	-0.0129*** (0.002)	-0.0117*** (0.002)	-0.0115*** (0.002)

Robust standard errors in parentheses
$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table 4.3a PHQ Regression Results: Mental Health Status Equation

	Model 1: PHQ10	Model 2: PHQ10	Model 3: PHQ10	Model 4: PHQ5
Constant	-1.2522	-2.3138	-1.7452	-1.0090
	(1.407)	(1.529)	(1.735)	(1.075)
FEMALE	0.5213	0.6111*	0.6523*	0.0120
	(0.343)	(0.357)	(0.352)	(0.214)
BLACK	0.7312	0.7387	0.8002*	0.2518
	(0.534)	(0.486)	(0.461)	(0.271)
AGE	0.1270	0.0369	0.0161	0.0128
	(0.103)	(0.123)	(0.124)	(0.073)
EDU	-1.6943**	-1.7986***	-1.8833***	-0.9174**
	(0.675)	(0.668)	(0.642)	(0.433)
MARRIED	-0.5312	-0.7207*	-0.8702**	-0.6150***
	(0.339)	(0.391)	(0.377)	(0.200)
CHILD	-0.3029	-0.3494	-0.4165	-0.3328*
	(0.267)	(0.284)	(0.277)	(0.184)
UNEMPLOY	0.0229	0.1153	0.1360	0.1619
	(0.462)	(0.436)	(0.440)	(0.293)
HHINC	0.0429	0.1463	0.2017	0.1621*
	(0.102)	(0.133)	(0.136)	(0.090)
OWNHOME	-0.6490**	-0.7325**	-0.6284**	-0.2928
	(0.266)	(0.298)	(0.316)	(0.224)
SMOKE		0.2240	0.2030	0.2085
		(0.286)	(0.305)	(0.189)
DRINK		0.2164	0.3055	0.0173
		(0.577)	(0.500)	(0.332)
INACTIVE		0.4283	0.3722	0.3963**
		(0.275)	(0.280)	(0.180)
CHRONIC		0.8400**	0.8480**	0.2440
		(0.348)	(0.368)	(0.244)
TRAUMA		0.5879	0.5824	0.5932
		(0.590)	(0.603)	(0.454)
SSI			-0.3643**	-0.2664**
			(0.170)	(0.101)
PDVI	0.3483**	0.3920**	0.5170***	0.5562***
	(0.153)	(0.176)	(0.176)	(0.090)
atanh(Rho)	-0.0622	-0.0989 (0.315)	-0.2480	-0.5754**
	(0.254)		(0.330)	(0.229)
ln(Sigma)	0.4068***	0.3933***	0.3781***	0.3786***
	(0.035)	(0.037)	(0.037)	(0.037)
N	418	418	418	418
AIC	1700.0382	1693.4798	1681.6736	1789.2212

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.3b PHQ Regression Results: PDVI Equation

	Model 1: PHQ10	Model 2: PHQ10	Model 3: PHQ10	Model 4: PHQ5
Constant	0.9747	0.7271	-0.2541	-0.2815
	(0.978)	(0.950)	(0.989)	(0.991)
FEMALE	0.1441	0.1560	0.1506	0.1643
	(0.161)	(0.161)	(0.158)	(0.158)
BLACK	0.3555**	0.3945**	0.3942**	0.4204**
	(0.176)	(0.182)	(0.180)	(0.181)
AGE	0.0603	0.0294	0.0524	0.0463
	(0.073)	(0.073)	(0.071)	(0.071)
EDU	0.9148**	0.9371**	0.9137**	0.8751**
	(0.395)	(0.396)	(0.392)	(0.390)
MARRIED	0.3955**	0.4083**	0.4598***	0.4603***
	(0.183)	(0.181)	(0.176)	(0.176)
CHILD	0.1653	0.1645	0.1444	0.1405
	(0.174)	(0.176)	(0.172)	(0.172)
UNEMPLOY	0.1072	0.0381	0.0510	0.0603
	(0.300)	(0.302)	(0.299)	(0.298)
HHINC	-0.2201**	-0.1936**	-0.1903**	-0.1924**
	(0.087)	(0.085)	(0.085)	(0.085)
OWNHOME	-0.0361	-0.0677	-0.1055	-0.0894
	(0.194)	(0.196)	(0.192)	(0.192)
SMOKE		0.2008	0.2411	0.2402
		(0.184)	(0.182)	(0.181)
DRINK		-0.2544	-0.3025	-0.3016
		(0.305)	(0.307)	(0.311)
INACTIVE		-0.2847	-0.2253	-0.2222
		(0.197)	(0.193)	(0.193)
CHRONIC		0.5412**	0.4850**	0.4865**
		(0.215)	(0.212)	(0.212)
TRAUMA		0.0262	0.0105	-0.0035
		(0.301)	(0.293)	(0.297)
SSI			0.3026***	0.3030***
			(0.079)	(0.080)
DSI	0.9886***	0.9686***	0.8769***	0.9716***
	(0.210)	(0.205)	(0.203)	(0.183)
DISTANCE	-0.0132***	-0.0131***	-0.0120***	-0.0111***
	(0.002)	(0.002)	(0.002)	(0.002)

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.4a GAD Regression Results: Mental Health Status Equation

	Model 1: GAD10	Model 2: GAD10	Model 3: GAD10	Model 4: GAD5
Constant	-2.4499	-4.0944**	-3.8494*	-0.3187
	(1.726)	(1.837)	(2.006)	(0.921)
FEMALE	-0.0630	-0.0144	0.0165	0.1281
	(0.314)	(0.343)	(0.343)	(0.184)
BLACK	0.8064	0.9433*	0.9772**	0.2513
	(0.504)	(0.482)	(0.476)	(0.221)
AGE	0.3561***	0.3090**	0.3044**	0.0554
	(0.130)	(0.145)	(0.150)	(0.069)
EDU	-0.2251	-0.0035	0.0223	-0.7630*
	(0.811)	(0.864)	(0.894)	(0.392)
MARRIED	-0.0152	-0.1535	-0.2619	-0.3212*
	(0.313)	(0.331)	(0.342)	(0.191)
CHILD	0.2148	0.1640	0.1316	0.0570
	(0.268)	(0.288)	(0.294)	(0.170)
UNEMPLOY	0.2443	0.3708	0.3728	-0.0140
	(0.434)	(0.406)	(0.424)	(0.268)
HHINC	-0.1190	-0.0215	0.0158	0.0526
	(0.120)	(0.134)	(0.137)	(0.081)
OWNHOME	-1.4250***	-1.4300***	-1.4096***	-0.2203
	(0.346)	(0.351)	(0.361)	(0.197)
SMOKE		0.2996	0.2813	0.2228
		(0.321)	(0.348)	(0.186)
DRINK		0.4406	0.4738	0.0443
		(0.440)	(0.422)	(0.324)
INACTIVE		0.5262*	0.5076*	0.1058
		(0.284)	(0.289)	(0.174)
CHRONIC		0.5821	0.6015	-0.0916
		(0.355)	(0.368)	(0.222)
TRAUMA		0.3594	0.3414	0.6909
		(0.606)	(0.666)	(0.426)
SSI			-0.2558	-0.2837**
			(0.193)	(0.087)
PDVI	0.2664*	0.2872*	0.3806**	0.6214**
	(0.150)	(0.169)	(0.189)	(0.059)
atanh(Rho)	0.1552	0.1384	0.0282	-0.8636**
	(0.240)	(0.278)	(0.302)	(0.205)
ln(Sigma)	0.4068***	0.3933***	0.3781**	0.3782***
	(0.035)	(0.037)	(0.037)	(0.037)
N	418	418	418	418
AIC	1676.1551	1676.5832	1666.0870	1792.5393

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.4b GAD Regression Results: PDVI Equation

	Model 1: GAD10	Model 2: GAD10	Model 3: GAD10	Model 4: GAD5
Constant	0.9734 (0.977)	0.7260 (0.950)	-0.2540 (0.988)	-0.2665 (0.991)
FEMALE	0.1449 (0.161)	0.1566 (0.161)	0.1505 (0.158)	0.1573 (0.158)
BLACK	0.3571** (0.176)	0.3957** (0.182)	0.3942** (0.180)	0.4072** (0.181)
AGE	0.0599 (0.073)	0.0291 (0.073)	0.0524 (0.071)	0.0493 (0.071)
EDU	0.9125** (0.394)	0.9352** (0.396)	0.9138** (0.392)	0.8942** (0.388)
MARRIED	0.3955** (0.183)	0.4083** (0.181)	0.4598*** (0.176)	0.4598*** (0.176)
CHILD	0.1651 (0.174)	0.1643 (0.176)	0.1444 (0.172)	0.1424 (0.171)
UNEMPLOY	0.1077 (0.300)	0.0385 (0.302)	0.0510 (0.299)	0.0556 (0.298)
HHINC	-0.2203** (0.087)	-0.1937** (0.085)	-0.1903** (0.085)	-0.1914** (0.085)
OWNHOME	-0.0350 (0.194)	-0.0668 (0.196)	-0.1056 (0.192)	-0.0972 (0.190)
SMOKE		0.2008 (0.184)	0.2411 (0.182)	0.2407 (0.182)
DRINK		-0.2544 (0.305)	-0.3025 (0.307)	-0.3021 (0.309)
INACTIVE		-0.2846 (0.197)	-0.2253 (0.193)	-0.2237 (0.193)
CHRONIC		0.5413** (0.215)	0.4850** (0.212)	0.4857** (0.212)
TRAUMA		0.0255 (0.302)	0.0105 (0.293)	0.0038 (0.294)
SSI			0.3026*** (0.079)	0.3026*** (0.079)
DSI	0.9945*** (0.210)	0.9734*** (0.207)	0.8768*** (0.208)	0.9257*** (0.174)
DISTANCE	-0.0131*** (0.002)	-0.0131*** (0.002)	-0.0120*** (0.002)	-0.0116*** (0.002)

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.5 Descriptive Statistics: Displacement Analysis†

Variable	Description	Mean	Std. Dev.
DISPLACE	Indicator of household displacement	0.03	0.171
DURATION	Duration of displacement (days)	0.151	0.358
MIN_DMG	Indicator of minimal housing damage (Base: no damage)	0.058	0.234
MOD_DMG	Indicator of moderate housing damage (Base: no damage)	0.114	0.318
SVR_DMG	Indicator of severe housing damage (Base: no damage)	0.053	0.225
DSI	Disaster severity index	0.123	0.329
DISTANCE	Distances to nearest hurricane path (10 miles)	0.71	0.454
$FEMALE^h$	Gender (female=1, male=0)	0.8	0.4
$BLACK^h$	Race and ethnicity (black=1, otherwise=0)	4.195	1.421
AGE^h	Age (10 years)	1.268	0.214
$MARRIED^h$	Marital status (married or cohabitating=1, otherwise=0)	0.439	0.497
EDU^h	Educational attainment (10 years)	0.536	0.499
$UNHEALTHY^h$	Health status (unhealthy=1, otherwise=0)	0.086	0.28
$UNEMPLOY^h$	Employment status (unemployed=1, otherwise=0)	10.202	1.115
CHILD	Children in household (children=1, otherwise=0)	0.568	0.496
HHINC	Household income in 2004 (ln$)	0.227	0.42
OWNHOME	Household tenure status (homeowner=1, otherwise=0)	0.074	0.262
MOBILE	Household home status (mobile home=1, otherwise=0)	0.232	0.423
SSI	Social support index	0.181	0.385
DIS_FAMILY	Reside with family while displaced (yes=1, no=0)	0.03	0.171
INSURANCE	Insurance receipts in 2006 ($10,000)	2.697	0.902
REMIT	Remittance from relative in 2005 and 2006 ($1,000)	0	1.77
ST_AL	Resided in AL at time of hazard event (yes=1, no=0)	0.147	0.428
ST_MS	Resided in AL at time of hazard event (yes=1, no=0)	7.315	4.325

†Socioeconomic, behavioral, and health characteristics refer to the household head in 2005, distinguished by their counterparts in the mental health analysis with the superscript h.

Table 4.6a Displacement Regression Results: Binary Component

	Model 1	Model 2	Model 3	Model 4
Constant	-0.7426***	-0.3934	-1.1213	-0.339
	(0.226)	(1.321)	(1.371)	(1.394)
MIN_DMG	0.5968***	0.6313***	0.6307***	0.7285***
	(0.191)	(0.201)	(0.200)	(0.213)
MOD_DMG	0.6468***	0.6688***	0.6163***	0.7409***
	(0.216)	(0.22)	(0.222)	(0.242)
SVR_DMG	1.4871***	1.5163***	1.4647***	1.5557***
	(0.292)	(0.306)	(0.313)	(0.315)
DSI	0.9654***	0.9814***	0.9291***	0.5921*
	(0.324)	(0.335)	(0.328)	(0.31)
DISTANCE	-0.0878***	-0.1002***	-0.0942***	-0.1219***
	(0.023)	(0.026)	(0.025)	(0.032)
FEMALE[h]		0.359	0.308	0.3078
		(0.229)	(0.234)	(0.246)
BLACK[h]		-0.3003	-0.3336	-0.446*
		(0.237)	(0.235)	(0.233)
AGE[h]		-0.0803	-0.0537	-0.0528
		(0.072)	(0.074)	(0.075)
MARRIED[h]		0.2374	0.2038	0.1688
		(0.235)	(0.237)	(0.243)
EDU[h]		0.3157	0.2411	0.2132
		(0.403)	(0.404)	(0.417)
UNHEALTHY[h]		0.2695	0.2145	0.1891
		(0.19)	(0.191)	(0.195)
UNEMPLOY[h]		-0.0561	-0.0476	0.0091
		(0.243)	(0.242)	(0.242)
CHILD		-0.1368	-0.135	-0.1259
		(0.178)	(0.18)	(0.184)
HHINC		-0.0308	-0.0331	-0.0394
		(0.124)	(0.128)	(0.126)
OWNHOME		-0.2171	-0.2639	-0.2538
		(0.189)	(0.192)	(0.192)
MOBILE		0.3747*	0.4172**	0.402*
		(0.201)	(0.205)	(0.214)
SSI			0.2889***	0.3102***
			(0.105)	(0.106)
ST_AL				-0.2901
				(0.292)
ST_MS				-0.7449***
				(0.221)

Robust standard errors in parentheses
$p < 0.10$, ** $p < 0.05$, *** $p < 0.0$

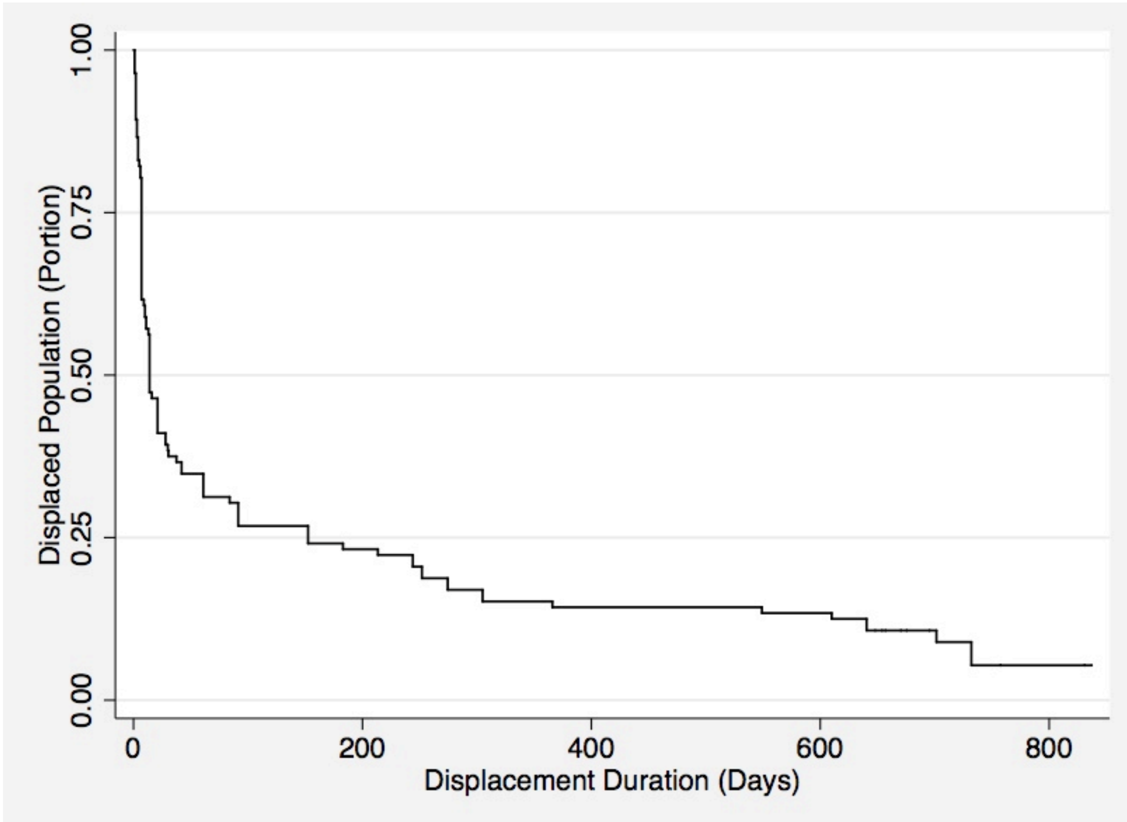
Table 4.6b Displacement Regression Results: Survival Component

	Model 1	Model 2	Model 3	Model 4
Constant	2.6833***	2.7722	4.0269*	3.9015*
	(0.582)	(2.039)	(2.204)	(2.194)
MIN_DMG	-0.4841	-0.3967	-0.4396	-0.3682
	(0.605)	(0.389)	(0.355)	(0.366)
MOD_DMG	0.8507	0.9978**	1.0901**	1.0517**
	(0.708)	(0.496)	(0.454)	(0.461)
SVR_DMG	1.5249**	1.4903***	1.6184***	1.6609***
	(0.727)	(0.501)	(0.479)	(0.514)
DSI	1.4205***	1.3632***	1.1908***	1.3773***
	(0.368)	(0.324)	(0.317)	(0.325)
DISTANCE	0.0008	-0.0206	-0.0523	-0.0486
	(0.065)	(0.067)	(0.064)	(0.083)
FEMALE[h]		-0.4292	-0.3102	-0.3107
		(0.381)	(0.345)	(0.378)
BLACK[h]		0.2981	0.2814	0.3948
		(0.307)	(0.290)	(0.291)
AGE[h]		0.1935*	0.1937**	0.2012**
		(0.102)	(0.095)	(0.100)
MARRIED[h]		-0.7516*	-0.9613**	-0.8648**
		(0.396)	(0.373)	(0.380)
EDU[h]		1.118	1.4417*	1.3376*
		(0.750)	(0.736)	(0.734)
UNHEALTHY[h]		-0.266	-0.3556	-0.3833
		(0.345)	(0.337)	(0.370)
UNEMPLOY[h]		0.1627	0.1155	0.0689
		(0.448)	(0.376)	(0.385)
CHILD		1.1118***	1.0601***	1.1233***
		(0.292)	(0.260)	(0.299)
HHINC		-0.2331	-0.2068	-0.2298
		(0.212)	(0.196)	(0.195)
OWNHOME		-0.2496	-0.3649	-0.374
		(0.350)	(0.319)	(0.314)
MOBILE		-0.4512	-0.5642	-0.585
		(0.414)	(0.394)	(0.417)
SSI			-0.4633*	-0.4575*
			(0.237)	(0.235)
DIS_FAMILY			-0.0794	-0.2295
			(0.252)	(0.270)
INSURANCE			0.0211	0.0212
			(0.034)	(0.033)
REMIT			-0.1898***	-0.1954***
			(0.048)	(0.050)
ST_AL				0.16
				(0.615)
ST_MS				0.5551*
				(0.301)
$\ln(\kappa)$	-0.3543***	-0.2147**	-0.1424	-0.1251
	(0.074)	(0.083)	(0.094)	(0.092)
N	425	424	422	422
AIC	1369.806	1381.366	1373.492	1365.956

Robust standard errors in parentheses

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.0$

Figure 4.1 Kaplan-Meier Estimator Plot for Displacement Duration: Zero-Truncated



Chapter 5: Concluding Remarks

5.1 Dissertation Summary

In the U.S., as well as globally, exposure to natural hazards is rapidly increasing, primarily due to growing urban populations within flood plains and along hazard-prone coastlines. Improved infrastructure, urban planning, and disaster preparedness have reduced mortality risk in recent decades. In contrast, greater levels of exposure have raised the risk of economic loss—and the adverse economic and health outcomes that accompany the large-scale destruction of assets. This trend, coupled with potential increases in the frequency and intensity of extreme weather events from climate change, underscores the importance of disaster research and continued advancements in hazard risk mitigation.

Thousands of hazard-events occurred in the U.S. between 2000 and 2011. Among these events were 246 major disasters, which resulted in 4293 deaths, $350 billion in damage, and directly or indirectly affected 2.1 million Americans.[49] Hurricane Katrina, the single most destructive hazard event to occur during this period, accounted for approximately 35% of these deaths and 31% of damages (Beaudoin, 2007; Knabb et al., 2011). High-severity low-probability disasters, such as Katrina, naturally receive a great deal of media attention and disaster aid (both federal assistance and aid from non-

[49] U.S. hazard statistics are calculated from the Emergency Events Database and pertain to earthquakes, floods, storms, volcanic activity, and wildfire.

governmental organizations). However, this emphasis masks the high cost of low-severity high-probability (i.e. extensive-risk) events. In the U.S., extensive-risk disasters account for the majority of hazard-related deaths and one-third of property damage since 1960 (UNISDR, 2011). While the underlying causes of hazard risk are complex, income levels and geographic location appear to be key determinants of extensive risk. In particular, areas with high mortality rates tend to have low average incomes and be located in rural areas of the Midwest, West, and South.

Natural hazards, either extensive or intensive, have severe consequences for affected populations. Immediate effects include displacement, physical injury, property damage, and the disruption of service flows from community and environmental amenities. In many cases, immediate effects have lasting impacts on wellbeing. Adverse mental health outcomes are among the most well documented long-term effects. Numerous studies have identified a positive correlation between hazard-related stressors, or proximity to a hazard event, and likelihood of being diagnosed with PTSD. Long-term effects also strain interpersonal relationships, lower income, and reduce access to primary health care. Hori and Schafer (2010) suggest these effects are especially prominent for those experiencing long-term displacement, and the associated stress of rebuilding livelihoods within new communities.

Households mitigate the effects of hazard events through investments in market insurance, self-insurance, and self-protection. Evidence indicates that, despite increasing availability of market insurance, a substantial share of households located in hazard prone areas remain uninsured. In the areas affected by Hurricane Katrina, for example, the portion of households with flood insurance ranged from 57.7% in St. Bernard Parish to

7.3% in Tangipahoa Parish (Kunreuther, 2006). Analyses of localized housing markets also reveal the widespread use of self-insurance (e.g. reinforced construction, storm shutters, fire sprinkler systems) and self-protection (e.g. migration to lower-risk areas) measures. In addition to households, government agencies mitigate the effects of hazard events though a variety of DRM interventions. Common interventions including disaster insurance programs, building codes, land use restrictions, disaster preparedness measures, and restoration of critical natural resources.[50]

As hazard risk continues to increase and policymakers seek to enhance DRM, it becomes increasingly important to quantify the costs and benefits of risk mitigation. This dissertation addresses the issue of benefits by estimating household WTP to live in MSAs with lower hazard event probabilities. Increasing risk also underscores the need for continued research on the behavior and health effects of hazard events. To this end, the dissertation conducts three analyses evaluating the effects on natural hazards on migration patterns, long-term mental health status, and displacement.

The WTP analysis, presented in Chapter 2, employs a residential sorting model in which households select the MSA that maximizes their indirect utility. When selecting a residential location, households trade off wages, prices, and location-specific attributes. These tradeoffs reflect household preferences and can be used to derive implicit prices. Coefficients for the sorting model are obtained using a two-stage estimation process that exploits spatial variation in labor markets, housing markets, and environmental amenities. In the first stage, a set of ASC is estimated using a conditional logit model and

---

[50] DRM interventions may incentivize risky behavior if government acts as an insurer of last-resort—reducing household investment in self-insurance and self-protection.

controlling for income and migration costs. The vector of ASC reflects household preferences for MSAs and is interpreted as a quality-of-life index. In the second stage, the vector of ASC is regressed against location-specific attributes. Coefficients from this regression represent parameters of the indirect utility function, and are easily manipulated to obtain marginal rates of technical substitution (i.e. marginal implicit prices). Results indicate, in line with hedonic-property literature, that households consider high-severity low-frequency events when making location decisions. More specifically, findings show an annual WTP of $275 per household for a marginal reduction in the expected number of earthquake, hurricane, or flood events per 1000 years. This value, in contrast to hedonic studies conducted in localized housing markets, incorporates WTP to avoid the broader consequences of natural disaster, such as disruption of service flows from community and environmental amenities. Results provide a point of departure for conducting benefit-cost analyses of risk-mitigating interventions, including efforts to mitigate climate change.

Chapter 3 explores the relationship between hazard risk and county migration patterns. Domestic migration decisions are motivated by a number of economic, social, and environmental considerations. Among these considerations is the potential for high-risk low-probability hazard events. This analysis estimates the relationship between county-level net in-migration rates and the expected frequency of earthquakes, hurricanes, and floods. Empirical estimation is complicated by the presence of spatial dependency and heterogeneity. These issues are addressed using SAC and GWR estimation metods. The SAC model controls for spatial dependency using a spatial-lag and spatial-error term. Findings indicate a negative correlation between net in-migration rates and hazard risk, after controlling for economic characteristics, demographic characteristics, environmental amenities, and counties affected by hurricanes Katrina and Rita. This result, consistent with Chapter 2, implies that residential location decisions are partially determined by hazard risk. Moreover, comparing standardized coefficients indicates that marginal impacts from hurricane and flood risk are comparable to other environmental amenities. In particular, they are similar to water area, topography, hazardous waste sites, and emissions. The marginal impact from earthquake risk is considerably smaller. The GWR regression estimates separate regression coefficients for each location, allowing for spatial heterogeneity in parameter values. Mapping regression coefficients reveals significant spatial variation in relationship between migration and hazard risk. The greatest impact occurs along the Gulf Coast. This suggests that regions where hazard risk has the greatest impact on migration are also the regions most susceptible to increased hurricane intensity from climate change.

Chapter 4 conducts two separate analyses regarding the wellbeing of those affected by Hurricanes Katrina and Rita. Both analyses are based on the supplemental questionnaire to the 2007 PSID, which is a previously unexplored database. The first analysis evaluates the effects of post-disaster vulnerability on long-term mental health status. Specifically, a simultaneous equations model is used to determine the effects of post-disaster vulnerability on incidents of PTSD, depression, and anxiety disorder. The first equation models the respondent's mental health status as a function of socioeconomic characteristics, behavioral and health attributes, social support, and post-disaster vulnerability. The latter is measured using the uniquely constructed PDVI, which quantifies the respondent's exposure to major stressors in the immediate aftermath of

Katrina and Rita. The second equation models the PDVI as a function of socioeconomic characteristics, behavioral and health attributes, social support, and disaster severity in the respondent's area of residence. Results from the analysis indicate a positive correlation between the PDVI and the probability of being diagnosed with PTSD, depression, or anxiety disorder. These findings are robust across most model specification and suggest that Katrina and Rita have had a lasting impact, of at least two years, on the mental health status of affected individuals.

Chapter 4 also evaluates the determinants of household displacement, which has been linked to adverse health and economic outcomes. A hurdle-Weibull model is developed in order to estimate the relationship between household characteristics and displacement duration, conditional on being displaced. Unsurprisingly, regression results show that housing damage is the most important predictor of displacement and displacement duration. Greater damage is associated with longer periods of displacement. Socioeconomic factors also have significant effects. In particular, duration is positively correlated with age, educational attainment, and the presence of children. It is negatively correlated with marriage or cohabitation. Finally, measures of social support are significantly correlated with displacement and displacement duration.

The social support index (SSI) has positive impact on displacement but a negative impact on the duration of displacement, implying that households rely on social networks to provide accommodations during and immediately following the hazard event as well as aid in repairing and returning to their residence. Remittances from relatives are shown to reduce the length of displacement. With respect to policy implications, these findings suggest that hazard assistance should be targeted towards those with high post-disaster vulnerability. One way of improving wellbeing, as evidenced in previous studies, is for households to return to familiar social and institutional settings. Policy interventions that foster social capital may improve the rate of return migration following a hazard event.

5.2 Avenues for Future Research

This work presents several avenues for further research. Specific to the residential sorting model (Chapter 2), improving the geographic precision of the choice set (i.e. from MSAs to counties or sub-counties) would allow for more accurate estimation of household preferences. A more precise choice set would allow for a larger vector of ASC-coefficients that, in turn, could be used to estimate more accurate parameter values. It would also improve the model's ability to identify nonlinear relationships between location-specific attributes and quality of life. The model could also be improved by relaxing a number of restrictive assumptions, namely use of the Cobb-Douglas utility function, independence of irrelevant alternatives, and the current specification of migration costs. In addition, the model could be extended, following Timmins (2007), to estimate the welfare effects of non-marginal changes in climate and disaster-risk.

Throughout this dissertation, residential location decisions (Chapter 2 and 3) are modeled as a function of scientific hazard probabilities. In reality these decisions are based on household risk perceptions. Limited evidence suggests that hazard risk perceptions are in constant flux. Risk perception increases dramatically following a hazard event, particularly for those who are directly or indirectly affected, and then declines over time. Continued research on risk perceptions is essential to accurately valuing the non-market benefit of safer locations. Of particular importance is determining whether risk perceptions accurately reflect risk and whether hazards events have permanent effects on risk perception.

Despite research on mental health, displacement and interpersonal relationships, little is known about the long-term impacts of hazard events. Future research should focus on identifying the long-term welfare effects, either through measures of consumption or income, associated with natural hazards. Given the socioeconomic differences between intensive-risk (i.e. high-severity, low-probability) and extensive-risk disasters (i.e. low-severity, high-probability, concentrated in rural areas) it may be useful to conduct separate welfare analysis for the two disaster types. Lastly, future research evaluating how hazard events affect the distribution of household resources, between household members and between consumption categories, will aid in developing more effective DRM interventions.

Appendices

Appendix A: Migration Probabilities for MSA-Specific Wage Regressions

Table A.1 Migration Probabilities for Census Divisions by Education Attainment and Marital Status

	Birth Division	Division of Residence 1	2	3	4	5	6	7	8	9	Total
No High School Degree Married	1	60.42	2.88	3.14	0.42	14.83	0.76	5.17	4.75	7.63	100
	2	1.33	75.33	2.33	0.42	13.34	0.72	1.75	2.11	2.67	100
	3	0.21	1.11	74.51	1.47	9.66	2.61	3.51	3.51	3.41	100
	4	0.44	0.8	4.5	68.21	5.3	0.94	6.46	7.69	5.66	100
	5	0.56	3.02	4.35	0.41	84.95	2.93	1.63	0.86	1.31	100
	6	0.19	0.98	14.35	1.59	12.24	64.16	3.79	1.17	1.54	100
	7	0.03	0.26	2.53	1.56	3.61	1.71	80.18	3.73	6.4	100
	8	0.5	0.5	2	1.25	2.75	1	6.99	72.63	12.4	100
	9	0.47	0.74	2.29	1.85	2.69	0.98	5.49	11.01	74.49	100
	10	2.89	10.63	7.23	1.47	12.7	0.95	18.18	8.58	37.37	100
No High School Degree Not Married	1	67.17	3.39	2.3	0.5	13.37	0.7	2.3	3.89	6.39	100
	2	1.89	73.01	2.23	0.37	15.37	0.83	1.22	2.57	2.51	100
	3	0.3	0.81	75.61	1.59	9.9	1.97	2.64	3.45	3.72	100
	4	0.33	0.65	4.4	68.84	5.3	0.57	4.98	8.48	6.44	100
	5	0.89	4.5	2.81	0.33	85.97	1.76	1.45	0.82	1.47	100
	6	0.32	2.07	15.7	2.02	10.08	62.71	3.93	1.27	1.91	100
	7	0.06	0.42	3.72	2.18	3.59	1.57	78.92	3.43	6.1	100
	8	0.54	0.45	1.7	1.79	2.33	0.89	4.47	74.96	12.88	100
	9	0.38	0.66	2	1.76	3.08	0.73	4.56	10.51	76.32	100
	10	3.7	14.43	5.92	1.57	15.95	1.25	16.39	8.42	32.36	100
High School Degree Married	1	76.39	3.17	1.56	0.38	11.08	0.79	1.42	2.14	3.06	100
	2	1.34	76.65	2.13	0.39	13.13	0.71	1.48	1.97	2.2	100
	3	0.31	1.19	77.08	1.68	8.1	2.43	2.9	3.36	2.95	100
	4	0.28	0.69	5.46	67.76	4.71	1.28	5.74	7.51	6.57	100
	5	0.8	3.74	3.49	0.62	83.52	2.67	2.04	1.4	1.73	100
	6	0.47	1.07	9.86	1.19	12.07	68.13	4.16	1.25	1.79	100
	7	0.26	0.7	2.55	1.84	4.12	2.07	80.44	3.1	4.92	100
	8	0.41	0.97	2.2	2.4	3.64	1.02	5.69	70.15	13.51	100
	9	0.45	1.01	2.09	1.82	3.84	0.94	4.35	9.18	76.32	100
	10	4.85	18.5	8.75	1.69	19.28	1.34	10.61	6.74	28.24	100

Table A.1 (cont.) Migration Probabilities for Census Divisions by Education Attainment and Marital Status

	Birth Division	Division of Residence									
		1	2	3	4	5	6	7	8	9	Total
High School Degree Not Married	1	74.11	3.25	1.72	0.43	12.89	0.83	1.58	2.12	3.07	100
	2	1.64	74.25	1.89	0.37	14.5	0.78	1.32	2.49	2.75	100
	3	0.31	1.14	75.48	1.77	8.83	2.11	2.78	4.03	3.54	100
	4	0.23	0.87	4.83	66.64	5.5	1.22	5.5	7.81	7.39	100
	5	1.01	4.54	3	0.48	83.72	2.16	1.95	1.35	1.79	100
	6	0.52	1.55	11.48	1.53	11.73	65.15	3.94	1.47	2.63	100
	7	0.24	0.65	2.7	2.12	4.24	1.96	79.19	3.66	5.25	100
	8	0.37	0.91	2.24	1.96	3.84	0.8	5.17	72.29	12.42	100
	9	0.5	1.03	2.14	1.34	4.07	1.09	4.03	9.34	76.47	100
	10	6.26	19.71	7.41	1.54	22.1	1.53	9.41	7.18	24.86	100
Some College Education Married	1	67.63	4.11	1.99	0.62	14.22	0.84	2.33	3.33	4.92	100
	2	2.39	64.05	3.16	0.68	17.9	1.03	2.65	3.39	4.76	100
	3	0.49	1.43	68.82	2.3	9.86	2.49	4	4.97	5.64	100
	4	0.47	0.86	6.13	59.22	6.17	1.52	7.33	9.33	8.98	100
	5	0.83	3.39	3.49	0.91	78.35	3.47	3.6	2.39	3.58	100
	6	0.45	1.15	7.86	1.21	15.78	61.73	6.14	2.25	3.42	100
	7	0.37	0.71	2.29	1.87	5.59	2.35	76.38	4.2	6.24	100
	8	0.47	0.95	2.23	2.16	4.63	0.97	6.22	67.33	15.03	100
	9	0.5	0.77	1.91	1.36	4.28	0.98	4.27	8.97	76.95	100
	10	4.35	14.85	8.05	1.7	22.05	1.25	9.58	6.08	32.09	100
Some College Education Not Married	1	65.8	4.13	1.66	0.58	14.83	0.91	1.99	3.75	6.36	100
	2	2.43	61.95	2.65	0.52	19.76	0.88	2.45	3.84	5.51	100
	3	0.5	1.34	67.68	2.32	10.59	2.42	3.58	5.31	6.26	100
	4	0.48	0.93	5.49	58.56	6.57	1.55	6.68	9.54	10.2	100
	5	0.94	3.99	2.94	0.81	80.17	2.4	2.79	2.2	3.75	100
	6	0.45	1.65	9.7	1.48	14.64	60.4	5.35	2.08	4.27	100
	7	0.38	0.81	2.55	1.93	5.36	2.11	75.15	4.33	7.38	100
	8	0.5	1.04	1.98	2.17	4.13	0.85	5.81	67.38	16.14	100
	9	0.51	0.94	1.73	1.29	4.08	0.82	3.66	8.95	78.02	100
	10	5.09	16.59	6.74	1.75	24.34	1.36	8.78	6.96	28.39	100

Table A.1 (cont.) Migration Probabilities for Census Divisions by Education Attainment and Marital Status

	Birth Division	Division of Residence									
		1	2	3	4	5	6	7	8	9	Total
College Degree Married	1	63.38	6.88	3.02	0.84	13.82	1.01	2.57	2.9	5.57	100
	2	4.64	57.64	4.4	0.89	19.3	1.19	3.01	3.12	5.81	100
	3	1.33	2.49	61.24	3.78	11.99	2.74	4.8	5.13	6.5	100
	4	0.95	1.5	8.17	52.55	7.75	2.08	9.06	8.97	8.98	100
	5	1.49	4.24	3.8	1.06	73.73	4.15	4.63	2.51	4.39	100
	6	0.7	1.27	6.19	1.5	19.18	57.25	8.27	2.2	3.45	100
	7	0.47	1.19	2.25	2.27	7.8	2.97	73.19	4.32	5.53	100
	8	0.81	1.48	3.24	2.95	6.24	1.2	7.88	58.66	17.55	100
	9	0.91	1.46	2.6	1.61	6.09	1.09	4.73	8.66	72.85	100
	10	4.36	17.15	9.08	1.96	21.2	1.43	9.52	4.5	30.81	100
College Degree Not Married	1	59.49	7.27	2.09	0.67	15.55	0.9	2.32	3.46	8.25	100
	2	4.28	55.74	3.67	0.72	20.91	0.91	2.5	3.66	7.61	100
	3	1.41	3.05	58.18	3.46	12.72	2.46	4.19	5.7	8.83	100
	4	0.94	2.01	7.64	50.72	8.43	1.69	7.63	10.27	10.66	100
	5	1.6	4.85	3.34	0.91	75.15	2.91	3.26	2.63	5.36	100
	6	0.8	2.09	6.39	1.36	20.2	55.01	7.16	2.39	4.59	100
	7	0.69	1.61	2.36	1.99	8.05	2.44	70.97	4.53	7.36	100
	8	0.91	2.29	3.44	2.38	7.07	0.96	7.22	56.06	19.67	100
	9	0.96	2.23	2.31	1.15	5.64	0.86	3.71	7.58	75.55	100
	10	4.51	17.97	7.55	1.73	22.64	1.27	8.41	5.1	30.8	100
Graduate Degree Married	1	55.09	8.85	3.65	1.1	16.92	1.21	2.72	3.12	7.35	100
	2	6.83	51.74	5.38	1.14	19.97	1.34	2.92	3.35	7.33	100
	3	2.44	4.29	52.55	4.07	14.41	2.91	4.87	5.5	8.95	100
	4	1.93	3.28	10.41	41	11.41	2.38	8.67	9.54	11.38	100
	5	2.67	6.03	4.84	1.46	66.9	4.46	4.76	3.03	5.85	100
	6	1.04	2.39	7.32	1.62	23.42	49.54	6.96	2.99	4.73	100
	7	1.31	2.7	4.02	2.61	11.4	3.77	61.71	5.25	7.23	100
	8	1.9	2.99	4.63	3.11	8.77	1.47	8.12	50.55	18.47	100
	9	1.95	3.06	3.2	1.83	8.48	1.25	4.44	8.76	67.04	100
	10	6.23	17.63	11.04	2.34	21.26	2	9.12	4.01	26.37	100

Table A.1 (cont.) Migration Probabilities for Census Divisions by Education Attainment and Marital Status

	Birth Division	Division of Residence 1	2	3	4	5	6	7	8	9	Total
Graduate Degree Not Married	1	54.53	9.25	3.34	1.05	16.4	0.7	2.2	3.54	8.98	100
	2	6.48	52.41	4.03	0.76	20.68	1.04	2.26	3.6	8.74	100
	3	2.26	4.42	51.04	3.33	14.96	2.42	4.45	6.08	11.04	100
	4	1.91	4.02	9.43	38.36	11.76	2.05	7.87	10.39	14.21	100
	5	2.52	6.18	4	1.13	69.73	3.45	3.94	2.77	6.26	100
	6	1.48	2.83	8.35	1.2	21.7	48.35	7.09	2.73	6.27	100
	7	1.6	2.7	3.21	2.7	11.19	2.85	60.97	5.56	9.24	100
	8	1.59	2.72	4.69	2.34	10.17	1.1	6.96	48.43	21.99	100
	9	1.74	3.07	2.96	1.32	8.31	0.8	3.7	7.88	70.21	100
	10	6.18	18.47	8.92	2.08	23.15	1.78	8.09	4.54	26.8	100

1=New England 2=Middle Atlantic 3=East North Central 4=West North Central
5=South Atlantic 6=East South Central 7=West South Central 8=Mountain
9=Pacific 10=Outside the U.S

Appendix B: Cost of Living and Quality of Life Results

Table B.1 Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Abilene, TX	-0.6577***	0.0198	263	1.765***	0.3426	221
Akron, OH	-0.4238***	0.0163	143	2.5752***	0.3050	119
Albany, GA	-0.7153***	0.0205	281	2.135***	0.3431	182
Albany-Schenectady-Troy, NY	-0.2439***	0.0162	93	2.542***	0.3067	124
Albuquerque, NM	-0.3191***	0.0165	108	3.7912***	0.3026	41
Alexandria, LA	-0.6822***	0.0187	273	1.7507***	0.3414	224
Allentown-Bethlehem-Easton, PA-NJ	-0.1753***	0.0164	79	2.2583***	0.3080	161
Altoona, PA	-0.7029***	0.0189	275	1.1376***	0.3515	271
Amarillo, TX	-0.4982***	0.0180	181	2.2214***	0.3234	163
Anderson, IN	-0.6796***	0.0181	272	1.5552***	0.3488	243
Anderson, SC	-0.6558***	0.0184	262	2.0207***	0.3351	192
Ann Arbor, MI	-0.1011***	0.0176	68	2.1523***	0.3182	177
Anniston-Oxford, AL	-0.7505***	0.0203	288	1.6445***	0.3597	234
Asheville, NC	-0.2635***	0.0184	99	2.8131***	0.3189	94
Athens-Clarke County, GA	-0.4836***	0.0192	174	2.7385***	0.3249	106
Atlanta-Sandy Springs-Marietta, GA	-0.3579***	0.0159	120	5.5112***	0.2931	2
Atlantic City-Hammonton, NJ	0.0295	0.0181	45	1.0049**	0.3608	276
Auburn-Opelika, AL	-0.5073***	0.0190	188	2.8667***	0.3173	91
Augusta-Richmond County, GA-SC	-0.6232***	0.0170	245	2.7899***	0.3118	96
Austin-Round Rock-San Marcos, TX	-0.222***	0.0163	89	4.3614***	0.2977	22
Bakersfield-Delano, CA	-0.1054***	0.0169	70	2.149***	0.3080	178

* p<0.1 ** p<0.05 *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Baltimore-Towson, MD	0.0179	0.0160	47	4.3927***	0.2956	20
Barnstable Town, MA	0.3509***	0.0176	15	0.9219**	0.3963	283
Baton Rouge, LA	-0.5032***	0.0166	185	3.0776***	0.3033	77
Battle Creek, MI	-0.5557***	0.0186	213	1.3084***	0.3320	257
Beaumont-Port Arthur, TX	-0.6311***	0.0176	250	2.0752***	0.3217	187
Bellingham, WA	0.0169	0.0192	48	2.5173***	0.3270	128
Bend, OR	0.0701***	0.0210	33	2.2619***	0.3382	160
Billings, MT	-0.439***	0.0194	152	2.159***	0.3379	176
Binghamton, NY	-0.6299***	0.0173	248	1.2633***	0.3378	262
Birmingham-Hoover, AL	-0.5272***	0.0165	199	3.7166***	0.3001	44
Bloomington, IN	-0.5275***	0.0200	200	2.278***	0.3290	159
Bloomington-Normal, IL	-0.4862***	0.0173	175	1.96***	0.3172	198
Boise City-Nampa, ID	-0.3825***	0.0169	131	3.6831***	0.3042	45
Boston-Cambridge-Quincy, MA-NH	0.3759***	0.0159	14	4.5937***	0.2953	16
Boulder, CO	0.0858***	0.0186	32	2.9223***	0.3205	87
Bremerton-Silverdale, WA	0.009	0.0182	49	2.2974***	0.3312	155
Bridgeport-Stamford-Norwalk, CT	0.4878***	0.0164	13	2.9243***	0.3089	86
Brownsville-Harlingen, TX	-0.7351***	0.0190	284	1.6474***	0.3406	232
Buffalo-Niagara Falls, NY	-0.4997***	0.0162	183	2.7101***	0.3060	110
Burlington, NC	-0.5493***	0.0187	209	1.7623***	0.3438	222
Burlington-South Burlington, VT	-0.1075***	0.0182	71	2.4333***	0.3255	138

* p<0.1 ** p<0.05 *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Canton-Massillon, OH	-0.564***	0.0167	218	2.1337***	0.3150	183
Cape Coral-Fort Myers, FL	0.0269	0.0169	46	3.4393***	0.3058	59
Cedar Rapids, IA	-0.5057***	0.0175	187	2.0435***	0.3211	189
Champaign-Urbana, IL	-0.4825***	0.0179	173	2.202***	0.3217	165
Charleston, WV	-0.7059***	0.0178	276	2.4402***	0.3249	137
Charleston-North Charleston-Summerville, SC	-0.2099***	0.0180	87	3.4602***	0.3057	58
Charlotte-Gastonia-Rock Hill, NC-SC	-0.3963***	0.0162	137	4.3941***	0.2969	19
Charlottesville, VA	-0.0956***	0.0198	65	2.3515***	0.3289	148
Chattanooga, TN-GA	-0.5591***	0.0170	215	2.8908***	0.3094	89
Chicago-Joliet-Naperville, IL-IN-WI	0.0571***	0.0158	37	4.8041***	0.2949	11
Chico, CA	0.0662***	0.0184	34	1.8284***	0.3226	213
Cincinnati-Middletown, OH-KY-IN	-0.4003***	0.0160	138	3.5825***	0.2968	53
Clarksville, TN-KY	-0.7101***	0.0184	277	2.3759***	0.3380	146
Cleveland-Elyria-Mentor, OH	-0.3618***	0.0159	122	3.5899***	0.2972	52
College Station-Bryan, TX	-0.4581***	0.0198	160	1.9623***	0.3341	196
Colorado Springs, CO	-0.3408***	0.0165	111	3.8526***	0.3056	36
Columbia, MO	-0.5903***	0.0195	228	2.6359***	0.3179	116
Columbia, SC	-0.5147***	0.0167	192	3.7965***	0.3016	40
Columbus, GA-AL	-0.6113***	0.0190	236	2.0431***	0.3368	190
Columbus, OH	-0.3943***	0.0160	135	3.8438***	0.2982	37
Corpus Christi, TX	-0.4749***	0.0181	168	2.416***	0.3193	142

* p<0.1   ** p<0.05   *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Dallas-Fort Worth-Arlington, TX	-0.3582***	0.0158	121	5.4783***	0.2930	3
Danville, VA	-0.7404***	0.0193	287	1.4867***	0.3682	246
Davenport-Moline-Rock Island, IA-IL	-0.5548***	0.0168	212	1.8745***	0.3159	208
Dayton, OH	-0.4991***	0.0162	182	2.9921***	0.3030	82
Decatur, AL	-0.74***	0.0187	286	1.9222***	0.3456	201
Decatur, IL	-0.6595***	0.0188	264	1.1699***	0.3423	268
Deltona-Daytona Beach-Ormond Beach, FL	-0.2279***	0.0172	91	2.6928***	0.3273	111
Denver-Aurora-Broomfield, CO	-0.1206***	0.0160	73	5.1956***	0.2947	5
Des Moines-West Des Moines, IA	-0.421***	0.0169	142	2.7771***	0.3107	101
Detroit-Warren-Livonia, MI	-0.2335***	0.0159	92	3.779***	0.2964	42
Dover, DE	-0.356***	0.0183	118	2.1948***	0.3275	167
Duluth, MN-WI	-0.4675***	0.0186	166	1.6797***	0.3370	230
Durham-Chapel Hill, NC	-0.308***	0.0172	107	3.1341***	0.3107	73
Eau Claire, WI	-0.5049***	0.0185	186	1.4311***	0.3512	248
El Centro, CA	-0.1911***	0.0214	84	0.1096	0.3927	294
Elkhart-Goshen, IN	-0.5903***	0.0180	229	1.9733***	0.3285	195
El Paso, TX	-0.6907***	0.0167	274	2.622***	0.3153	117
Erie, PA	-0.6165***	0.0173	241	1.8434***	0.3289	212
Eugene-Springfield, OR	-0.0928***	0.0173	64	2.8686***	0.3171	90
Evansville, IN-KY	-0.5957***	0.0172	232	2.4005***	0.3145	144
Fargo, ND-MN	-0.5281***	0.0199	201	2.4162***	0.3193	141

* p<0.1   ** p<0.05   *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Farmington, NM	-0.3559***	0.0234	116	1.3474***	0.3839	255
Fayetteville, NC	-0.6374***	0.0177	253	2.278***	0.3293	158
Fayetteville-Springdale-Rogers, AR-MO	-0.5021***	0.0174	184	3.4989***	0.3057	55
Flagstaff, AZ	-0.0076	0.0233	53	1.7978***	0.3643	218
Flint, MI	-0.6402***	0.0190	254	0.6412*	0.3799	292
Florence-Muscle Shoals, AL	-0.7626***	0.0193	293	1.8828***	0.3419	207
Fort Collins-Loveland, CO	-0.2231***	0.0171	90	2.9487***	0.3185	85
Fort Smith, AR-OK	-0.7577***	0.0188	291	2.1451***	0.3398	179
Crestview-Fort Walton Beach-Destin, FL	-0.1911***	0.0192	85	2.5477***	0.3290	121
Fort Wayne, IN	-0.7109***	0.0168	278	2.7897***	0.3094	97
Fresno, CA	-0.0095	0.0168	54	2.3125***	0.3113	152
Gadsden, AL	-0.754***	0.0201	289	1.8615***	0.3490	209
Gainesville, FL	-0.2801***	0.0182	102	2.6642***	0.3234	112
Gainesville, GA	-0.4536***	0.0194	157	2.5093***	0.3228	130
Glens Falls, NY	-0.3889***	0.0188	133	0.6613*	0.3544	291
Goldsboro, NC	-0.632***	0.0183	251	1.8501***	0.3494	211
Grand Junction, CO	-0.2558***	0.0192	96	1.9846***	0.3439	194
Grand Rapids-Wyoming, MI	-0.4377***	0.0166	150	2.4202***	0.3088	139
Greeley, CO	-0.3535***	0.0177	115	2.7813***	0.3173	99
Green Bay, WI	-0.3818***	0.0172	129	1.801***	0.3274	215
Greensboro-High Point, NC	-0.5249***	0.0165	196	3.3999***	0.3027	62

* p<0.1 ** p<0.05 *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Greenville, NC	-0.537***	0.0189	204	2.2951***	0.3234	156
Greenville-Mauldin-Easley, SC	-0.5432***	0.0170	205	3.3168***	0.3070	65
Gulfport-Biloxi, MS	-0.4908***	0.0197	176	2.4845***	0.3311	133
Hagerstown-Martinsburg, MD-WV	-0.2214***	0.0183	88	1.9607***	0.3321	197
Hanford-Corcoran, CA	-0.1896***	0.0214	83	0.7496**	0.3638	288
Harrisburg-Carlisle, PA	-0.444***	0.0165	154	2.66***	0.3071	113
Hartford-West Hartford-East Hartford, CT	0.065***	0.0160	35	3.4033***	0.3015	61
Hattiesburg, MS	-0.6557***	0.0210	261	2.3048***	0.3305	153
Hickory-Lenoir-Morganton, NC	-0.6123***	0.0171	237	2.742***	0.3128	105
Holland-Grand Haven, MI	-0.4248***	0.0176	145	1.472***	0.3271	247
Houma-Bayou Cane-Thibodaux, LA	-0.5796***	0.0216	224	0.9418**	0.3788	281
Houston-Sugar Land-Baytown, TX	-0.3821***	0.0159	130	5.0168***	0.2938	9
Huntsville, AL	-0.6278***	0.0172	247	3.052***	0.3111	79
Indianapolis-Carmel, IN	-0.5115***	0.0161	191	4.1808***	0.2965	26
Iowa City, IA	-0.3714***	0.0194	126	1.8975***	0.3358	204
Jackson, MI	-0.5091***	0.0191	190	0.8475**	0.3603	285
Jackson, MS	-0.576***	0.0171	223	3.1111***	0.3055	74
Jackson, TN	-0.7344***	0.0193	283	1.528***	0.3788	245
Jacksonville, FL	-0.2676***	0.0164	100	4.3416***	0.2979	23
Jacksonville, NC	-0.5635***	0.0206	217	2.1872***	0.3486	172
Janesville, WI	-0.4239***	0.0183	144	1.2449***	0.3570	263

* p<0.1 ** p<0.05 *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Johnson City, TN	-0.5967***	0.0204	233	2.1445***	0.3412	180
Johnstown, PA	-0.8121***	0.0185	296	1.3584***	0.3336	254
Joplin, MO	-0.8006***	0.0185	295	2.0599***	0.3361	188
Kalamazoo-Portage, MI	-0.4934***	0.0173	178	2.0264***	0.3184	191
Kankakee-Bradley, IL	-0.3842***	0.0183	132	1.1888***	0.3399	266
Kansas City, MO-KS	-0.4284***	0.0160	148	4.1491***	0.2978	28
Kennewick-Pasco-Richland, WA	-0.4686***	0.0177	167	2.1639***	0.3294	174
Killeen-Temple-Fort Hood, TX	-0.6201***	0.0175	243	2.4577***	0.3222	136
Kingsport-Bristol-Bristol, TN-VA	-0.7141***	0.0176	280	2.1928***	0.3350	169
Kingston, NY	0.0073	0.0182	50	0.5528	0.3511	293
Knoxville, TN	-0.5432***	0.0166	206	3.655***	0.3040	48
Kokomo, IN	-0.6635***	0.0187	266	0.976**	0.3691	279
La Crosse, WI-MN	-0.4396***	0.0191	153	1.6221***	0.3415	237
Lafayette, LA	-0.4121***	0.0216	140	1.8181***	0.3513	214
Lake Charles, LA	-0.6472***	0.0189	258	1.6352***	0.3423	235
Lakeland-Winter Haven, FL	-0.3382***	0.0166	110	3.4229***	0.3089	60
Lancaster, PA	-0.3248***	0.0167	109	2.2828***	0.3112	157
Lansing-East Lansing, MI	-0.4378***	0.0166	151	2.2235***	0.3110	162
Laredo, TX	-0.6208***	0.0205	244	1.176**	0.3860	267
Las Cruces, NM	-0.4792***	0.0200	170	2.3963***	0.3475	145
Las Vegas-Paradise, NV	-0.0445**	0.0161	58	5.0425***	0.2951	8

* p<0.1 ** p<0.05 *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Lebanon, PA	-0.4266***	0.0191	147	0.9342**	0.3475	282
Lewiston-Auburn, ME	-0.3473***	0.0200	113	1.8855***	0.3509	206
Lexington-Fayette, KY	-0.4669***	0.0176	164	3.0985***	0.3116	75
Lincoln, NE	-0.48***	0.0174	171	2.6464***	0.3145	114
Little Rock-North Little Rock-Conway, AR	-0.5717***	0.0167	220	3.4835***	0.3027	57
Los Angeles-Long Beach-Santa Ana, CA	0.6356***	0.0158	6	5.0548***	0.2945	7
Louisville/Jefferson County, KY-IN	-0.4445***	0.0161	155	3.6383***	0.2976	51
Lubbock, TX	-0.5583***	0.0183	214	2.5373***	0.3181	126
Lynchburg, VA	-0.5919***	0.0184	231	2.5007***	0.3191	131
Macon, GA	-0.6538***	0.0199	260	1.4296***	0.3606	249
Madera-Chowchilla, CA	0.000***	NA	51	0.000***	NA	295
Madison, WI	-0.1011***	0.0168	69	3.0686***	0.3078	78
Manchester-Nashua, NH	0.0429**	0.0167	40	2.8527***	0.3115	93
Mansfield, OH	-0.6407***	0.0186	256	0.8535**	0.3651	284
McAllen-Edinburg-Mission, TX	-0.7645***	0.0175	294	2.3047***	0.3237	154
Medford, OR	0.0365*	0.0191	43	2.4151***	0.3308	143
Memphis, TN-MS-AR	-0.497***	0.0164	180	3.6427***	0.3025	50
Merced, CA	-0.0179	0.0211	55	0.9866**	0.3512	278
Miami-Fort Lauderdale-Pompano Beach, FL	0.1694***	0.0159	24	4.5201***	0.2957	17
Michigan City-La Porte, IN	-0.5081***	0.0201	189	1.2291***	0.3547	264
Midland, TX	-0.462***	0.0209	162	1.0263**	0.3612	275

* p<0.1   ** p<0.05   *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Milwaukee-Waukesha-West Allis, WI	-0.1142***	0.0162	72	3.2629***	0.2987	68
Minneapolis-St. Paul-Bloomington, MN-WI	-0.1002***	0.0159	67	4.2646***	0.2961	25
Mobile, AL	-0.6123***	0.0175	238	2.7887***	0.3153	98
Modesto, CA	0.1111***	0.0175	30	1.893***	0.3147	205
Monroe, LA	-0.6745***	0.0207	270	1.8526***	0.3436	210
Monroe, MI	-0.3512***	0.0178	114	0.7779**	0.3596	287
Montgomery, AL	-0.6164***	0.0176	240	2.7615***	0.3167	103
Morgantown, WV	-0.6479***	0.0219	259	2.5439***	0.3314	123
Mount Vernon-Anacortes, WA	0.1585***	0.0179	27	2.5527***	0.3284	120
Muncie, IN	-0.7123***	0.0188	279	1.2802***	0.3671	261
Muskegon-Norton Shores, MI	-0.5991***	0.0184	234	1.4276***	0.3377	251
Myrtle Beach-North Myrtle Beach-Conway, SC	-0.3414***	0.0181	112	3.0828***	0.3161	76
Napa, CA	0.6405***	0.0209	5	0.7453**	0.3563	289
Naples-Marco Island, FL	0.3085***	0.0180	18	2.188***	0.3275	171
Nashville-Davidson--Murfreesboro--Franklin, TN	-0.4195***	0.0162	141	4.666***	0.2970	14
New Haven-Milford, CT	0.1521***	0.0162	28	3.1493***	0.3061	72
New Orleans-Metairie-Kenner, LA	-0.2588***	0.0169	98	3.3274***	0.3024	64
New York-Northern New Jersey-Long Island, NY-NJ-PA	0.5014***	0.0158	12	4.7454***	0.2959	12
Niles-Benton Harbor, MI	-0.4967***	0.0189	179	1.7342***	0.3292	226
Norwich-New London, CT	0.0417**	0.0172	41	2.3573***	0.3202	147
Ocala, FL	-0.3633***	0.0172	123	2.5327***	0.3262	127

* p<0.1 ** p<0.05 *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Ocean City, NJ	0.2998***	0.0207	19	-0.0133	0.4667	296
Odessa, TX	-0.663***	0.0220	265	0.6896*	0.4072	290
Ogden-Clearfield, UT	-0.4637***	0.0166	163	3.2771***	0.3054	67
Oklahoma City, OK	-0.5807***	0.0163	225	4.0018***	0.2993	34
Olympia, WA	-0.0769***	0.0181	62	2.7779***	0.3180	100
Omaha-Council Bluffs, NE-IA	-0.4669***	0.0165	165	3.4864***	0.3029	56
Orlando-Kissimmee-Sanford, FL	-0.1459***	0.0160	76	4.7111***	0.2957	13
Owensboro, KY	-0.7319***	0.0197	282	1.681***	0.3534	229
Oxnard-Thousand Oaks-Ventura, CA	0.6079***	0.0165	8	2.1948***	0.3108	168
Palm Bay-Melbourne-Titusville, FL	-0.2686***	0.0167	101	3.0216***	0.3111	80
Panama City-Lynn Haven-Panama City Beach, FL	-0.2467***	0.0200	94	2.1651***	0.3374	173
Pascagoula, MS	-0.5475***	0.0206	208	1.7251***	0.3471	228
Pensacola-Ferry Pass-Brent, FL	-0.4263***	0.0180	146	2.7279***	0.3187	107
Peoria, IL	-0.5271***	0.0166	198	2.4808***	0.3094	134
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	-0.028*	0.0158	56	4.0359***	0.2953	31
Phoenix-Mesa-Glendale, AZ	-0.0959***	0.0159	66	5.5746***	0.2934	1
Pittsburgh, PA	-0.5177***	0.0160	194	3.8286***	0.2989	38
Pittsfield, MA	-0.1592***	0.0192	78	1.3347***	0.3578	256
Portland-South Portland-Biddeford, ME	-0.055***	0.0173	59	2.9565***	0.3146	84
Portland-Vancouver-Hillsboro, OR-WA	0.0385**	0.0160	42	4.6461***	0.2957	15
Port St. Lucie, FL	-0.0382**	0.0174	57	2.854***	0.3148	92

* p<0.1 ** p<0.05 *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Poughkeepsie-Newburgh-Middletown, NY	0.1464***	0.0163	29	1.6197***	0.3155	238
Prescott, AZ	-0.0576**	0.0192	60	2.7692***	0.3311	102
Providence-New Bedford-Fall River, RI-MA	0.168***	0.0160	25	3.2307***	0.3015	69
Provo-Orem, UT	-0.3943***	0.0169	136	3.3425***	0.3066	63
Pueblo, CO	-0.572***	0.0181	221	2.1965***	0.3453	166
Punta Gorda, FL	-0.0664***	0.0190	61	1.593***	0.3682	239
Racine, WI	-0.1952***	0.0182	86	1.1553***	0.3501	269
Raleigh-Cary, NC	-0.3672***	0.0163	124	4.3782***	0.2972	21
Rapid City, SD	-0.4562***	0.0206	158	2.0802***	0.3556	186
Reading, PA	-0.3805***	0.0169	128	2.0952***	0.3155	185
Redding, CA	0.0488**	0.0186	39	1.0951**	0.3488	274
Reno-Sparks, NV	0.1004***	0.0174	31	3.7464***	0.3041	43
Richmond, VA	-0.2933***	0.0162	105	4.0826***	0.2977	29
Riverside-San Bernardino-Ontario, CA	0.1642***	0.0160	26	3.972***	0.2965	35
Roanoke, VA	-0.4766***	0.0174	169	2.6089***	0.3189	118
Rochester, MN	-0.3904***	0.0186	134	1.7857***	0.3346	220
Rochester, NY	-0.4604***	0.0162	161	2.7249***	0.3066	108
Rockford, IL	-0.4564***	0.0168	159	1.9543***	0.3207	199
Rocky Mount, NC	-0.6353***	0.0193	252	1.5599***	0.3487	242
Sacramento--Arden-Arcade--Roseville, CA	0.2473***	0.0160	20	3.6556***	0.2981	47
Saginaw-Saginaw Township North, MI	-0.582***	0.0178	226	1.1133***	0.3380	272

* p<0.1 ** p<0.05 *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
St. Cloud, MN	-0.4036***	0.0180	139	2.1892***	0.3200	170
St. Joseph, MO-KS	-0.6761***	0.0197	271	1.7318***	0.3392	227
St. Louis, MO-IL	-0.3801***	0.0159	127	3.8134***	0.2973	39
Salem, OR	-0.1888***	0.0173	81	2.6455***	0.3165	115
Salinas, CA	0.705***	0.0206	4	1.2904***	0.3250	259
Salt Lake City, UT	-0.2576***	0.0164	97	4.0064***	0.2989	33
San Antonio-New Braunfels, TX	-0.5289***	0.0162	202	4.0806***	0.2987	30
San Diego-Carlsbad-San Marcos, CA	0.5176***	0.0160	11	4.0271***	0.2989	32
San Francisco-Oakland-Fremont, CA	0.7611***	0.0160	3	4.4045***	0.2970	18
San Jose-Sunnyvale-Santa Clara, CA	0.806***	0.0163	2	2.9023***	0.3040	88
San Luis Obispo-Paso Robles, CA	0.5603***	0.0184	10	1.4057***	0.3299	252
Santa Barbara-Santa Maria-Goleta, CA	0.6298***	0.0191	7	1.5754***	0.3284	240
Santa Cruz-Watsonville, CA	0.8227***	0.0185	1	1.0037**	0.3377	277
Santa Fe, NM	0.0342*	0.0205	44	2.5392***	0.3331	125
Santa Rosa-Petaluma, CA	0.6042***	0.0170	9	2.0111***	0.3137	193
North Port-Bradenton-Sarasota, FL	0.0578***	0.0165	36	3.6744***	0.3040	46
Savannah, GA	-0.2838***	0.0194	103	2.7132***	0.3173	109
Scranton--Wilkes-Barre, PA	-0.4928***	0.0167	177	2.4176***	0.3119	140
Seattle-Tacoma-Bellevue, WA	0.1981***	0.0159	22	5.1253***	0.2936	6
Sheboygan, WI	-0.356***	0.0187	119	1.2836***	0.3494	260
Shreveport-Bossier City, LA	-0.6179***	0.0190	242	2.5105***	0.3151	129

* p<0.1 ** p<0.05 *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Sioux City, IA-NE-SD	-0.7384***	0.0204	285	1.7983***	0.3600	216
Sioux Falls, SD	-0.5253***	0.0200	197	2.3166***	0.3320	150
South Bend-Mishawaka, IN-MI	-0.626***	0.0171	246	2.3138***	0.3264	151
Spartanburg, SC	-0.6681***	0.0176	268	2.8089***	0.3163	95
Spokane, WA	-0.3559***	0.0168	117	3.5145***	0.3069	54
Springfield, IL	-0.5537***	0.0196	211	1.304***	0.3428	258
Springfield, MA	-0.0858***	0.0163	63	2.4916***	0.3167	132
Springfield, MO	-0.67***	0.0169	269	3.181***	0.3112	71
Springfield, OH	-0.591***	0.0176	230	0.9639**	0.3441	280
State College, PA	-0.4502***	0.0191	156	1.5466***	0.3418	244
Stockton, CA	0.1762***	0.0172	23	1.798***	0.3192	217
Sumter, SC	-0.7575***	0.0204	290	1.2072**	0.4130	265
Syracuse, NY	-0.5465***	0.0164	207	2.4685***	0.3066	135
Tallahassee, FL	-0.2993***	0.0178	106	3.0015***	0.3149	81
Tampa-St. Petersburg-Clearwater, FL	-0.147***	0.0160	77	4.9712***	0.2951	10
Terre Haute, IN	-0.7624***	0.0180	292	1.7599***	0.3313	223
Toledo, OH	-0.4819***	0.0165	172	2.5439***	0.3037	122
Topeka, KS	-0.6305***	0.0184	249	2.1442***	0.3250	181
Trenton-Ewing, NJ	0.2242***	0.0174	21	1.5652***	0.3292	241
Tucson, AZ	-0.1894***	0.0164	82	4.1682***	0.3016	27
Tulsa, OK	-0.5615***	0.0165	216	3.6433***	0.3027	49
Tuscaloosa, AL	-0.5701***	0.0188	219	2.1599***	0.3350	175

* p<0.1 ** p<0.05 *** p<0.01

Table B.1 (cont.): Cost of Living and Quality of Life Results for 296 MSAs

MSA	Cost of Living			Quality of Life		
	Coefficient	Robust Std. Err.	Rank	Coefficient	Robust Std. Err.	Rank
Tyler, TX	-0.5162***	0.0187	193	1.7948***	0.3286	219
Utica-Rome, NY	-0.6149***	0.0171	239	1.6744***	0.3189	231
Valdosta, GA	-0.6662***	0.0209	267	1.9122***	0.3568	202
Vallejo-Fairfield, CA	0.3191***	0.0174	17	1.6341***	0.3208	236
Vineland-Millville-Bridgeton, NJ	-0.2469***	0.0183	95	0.8452**	0.3511	286
Virginia Beach-Norfolk-Newport News, VA-NC	-0.1368***	0.0162	74	4.3138***	0.2981	24
Visalia-Porterville, CA	-0.139***	0.0183	75	1.1421***	0.3380	270
Waco, TX	-0.5864***	0.0180	227	2.2042***	0.3225	164
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.3212***	0.0159	16	5.2072***	0.2944	4
Waterloo-Cedar Falls, IA	-0.5509***	0.0191	210	1.6453***	0.3453	233
Wausau, WI	-0.5207***	0.0181	195	1.8997***	0.3256	203
Wichita, KS	-0.647***	0.0166	257	3.2143***	0.3051	70
Wichita Falls, TX	-0.6032***	0.0203	235	1.4276***	0.3651	250
Williamsport, PA	-0.5723***	0.0180	222	1.3863***	0.3398	253
Wilmington, NC	-0.1815***	0.0183	80	2.9703***	0.3166	83
Winston-Salem, NC	-0.5315***	0.0167	203	3.3152***	0.3071	66
Worcester, MA	0.0496**	0.0162	38	2.7574***	0.3053	104
Yakima, WA	-0.4315***	0.0187	149	1.9538***	0.3374	200
York-Hanover, PA	-0.37***	0.0166	125	2.1254***	0.3183	184
Youngstown-Warren-Boardman, OH-PA	-0.6405***	0.0164	255	1.7375***	0.3122	225
Yuba City, CA	-0.0017	0.0203	52	1.1028***	0.3436	273
Yuma, AZ	-0.2919***	0.0199	104	2.3234***	0.3386	149

* p<0.1 ** p<0.05 *** p<0.01

Appendix C: Alternate ASC Decompositions

Hazard risk (HRISK) is highly skewed. This raises concerns that outliers unduly influence parameter estimates. Table C.1 presents three alternate model specifications that address this issue. Model 5 implements Stata's *rreg* command, which is a form of robust regression that eliminates gross outliers and produces case weights for remaining observations. The Cook's D method is used to identify and eliminate outliers. Huber weights and biweights are used to calculate case weights. A detailed description of this methodology is available in Berk (1990). Models 6 and 7 are 2SLS estimations that, as in Model 4, instrument for MSA population. Model 6 uses a restricted dataset whereby the top 5% of HRISK values are excluded. Model 7 transforms the HRISK variable into a vector of dummy variables: *LowRisk*, *MediumRisk*, and *HighRisk*. The variable *LowRisk*=1 for MSAs with fewer than 2 expected hazard events (0 otherwise), *MediumRisk*=1 for MSAs with between 2 and 10 expected hazard events (0 otherwise), and *HighRisk*=1 for MSAs with more than 10 expected hazard events (0 otherwise). In all three alternate regression models there is a negative and significant relationship between quality of life and hazard risk.

Table C.1 Alternate ASC Decompositions

Variable	Model 5: Robust Regression		Model 6: 2SLS		Model 7: 2SLS	
	Coefficient	Robust Std. Err.	Coefficient	Robust Std. Err.	Coefficient	Robust Std. Err.
Constant	-9.551***	(0.810)	-10.484***	(1.188)	-10.577***	(1.151)
lnPOP	1.004***	(0.039)	0.960***	(0.075)	0.974***	(0.073)
UNEMP	-16.167***	(1.953)	-10.475***	(3.746)	-10.885***	(3.439)
PCTAX	-0.204**	(0.101)	-0.206*	(0.113)	-0.197*	(0.110)
VCRIME	-0.237*	(0.136)	-0.412***	(0.141)	-0.426***	(0.148)
PHYSICIANS	0.013	(0.015)	0.026*	(0.015)	0.024	(0.016)
ARTINDEX	0.070***	(0.025)	0.063**	(0.029)	0.060**	(0.029)
SUBWAY	0.029	(0.113)	0.110	(0.132)	0.117	(0.131)
TEACHERS	-0.032*	(0.018)	-0.043**	(0.019)	-0.052***	(0.019)
DROPOUT	-0.192	(1.332)	-0.023	(1.591)	-0.687	(1.641)
VOTERS	-0.136	(0.440)	0.325	(0.490)	0.684	(0.476)
TEMP	0.005	(0.008)	0.021*	(0.011)	0.018*	(0.010)
PRECIP	0.012***	(0.003)	0.011***	(0.003)	0.012***	(0.003)
OCEAN	-0.117*	(0.065)	-0.074	(0.059)	-0.134**	(0.061)
EMISSIONS	-0.002	(0.002)	-0.003**	(0.001)	-0.002*	(0.001)
NPLSITES	-0.007*	(0.004)	-0.005**	(0.002)	-0.006**	(0.003)
PARKS	-4.041	(4.538)	-5.673	(5.473)	-7.375*	(4.150)
HRISK	-0.008***	(0.001)	-0.011***	(0.003)		
MediumRisk					-0.014	(0.056)
HighRisk					-0.307***	(0.094)
REG1	-0.531***	(0.101)	-0.442***	(0.102)	-0.434***	(0.103)
REG3	0.184*	(0.099)	0.209**	(0.093)	0.247**	(0.096)
REG4	0.872***	(0.113)	0.968***	(0.127)	0.949***	(0.125)
Adjusted R^2	0.871		0.866		0.862	
N	296		281		296	

* p<0.1 ** p<0.05 *** p<0.01

Appendix D: Relative Contribution to Quality of Life Index

ASC decompositions can be supplemented with an analysis of the relative contribution of location-specific attributes. Fields (2004) develops a method, presented in Equations D1 and D2, for calculating relative importance.

$$p-weight(X^k) = \left(\frac{s(X^k)}{R^2}\right) * 100 \tag{D1}$$

$$\sum_{k=1}^{K} s-weight(X^k) = \frac{\sum_{k=1}^{K} \text{cov}[X^{ik}\beta^k, Y]}{\text{var}(Y)} \tag{D2}$$

Here X^k is the vector of explanatory variables, β^k the vector of estimated coefficients, and Y the dependent variable (i.e. quality-of-life index). The *s-weight* values indicate the share of total variation explained by each dependent variable. The *p-weight* values, which are scaled *s-weights*, indicate the percent contribution of each dependent variable to explained variation. As discussed by Fields (2004), this method meets three essential criteria: 1) explanatory power is assigned such that other factors are held constant, 2) the sum of *p-weights* equals the total explanatory power of the model, and 3) variation in the dependent variable is gauged by an index other than variance.

This method is applied to Model 2 from Table 2.11. *P-weights* are determined for each location-specific attribute. These weights are then summed across several categories: population (lnPOP), socioeconomic characteristics (UNEMP, PCTAX, VCRIME, PHYSICIANS, ARTINDEX, SUBWAY, TEACHERS, DROPOUT, and VOTERS), climate amenities (TEMP, PRECIP, and OCEAN), and non-climate amenities (EMISSIONS, NPLSITES, PARKS, and HRISK). Results are presented in Table D.1. Unsurprisingly, population is the dominant contributor to the housing-cost-adjusted quality-of-life index. This is a consequence of the endogenous relationship between population and quality of life discussed in Section 2.6.4. The remaining categories, in order of their contribution to quality of life are socioeconomic characteristics (7.00%), non-climate amenities (2.82%) and climate amenities (1.42%).

Table D.1 Relative Contribution to Quality of Life by Variable Category

Category	Percent Contribution to Housing-Cost Adjusted ‘Quality of Life’
Population	88.76%
Socioeconomic Characteristics	7.00%
Climate Amenities	1.42%
Non-Climate Amenities	2.82%
Total	100.00%

Appendix E: Distribution of Environmental Amenities

Figure E.1 Spatial Distribution of Environmental Amenities: Quartiles

[Note: Lightest color denotes first quartile and darkest denotes forth quartile.]

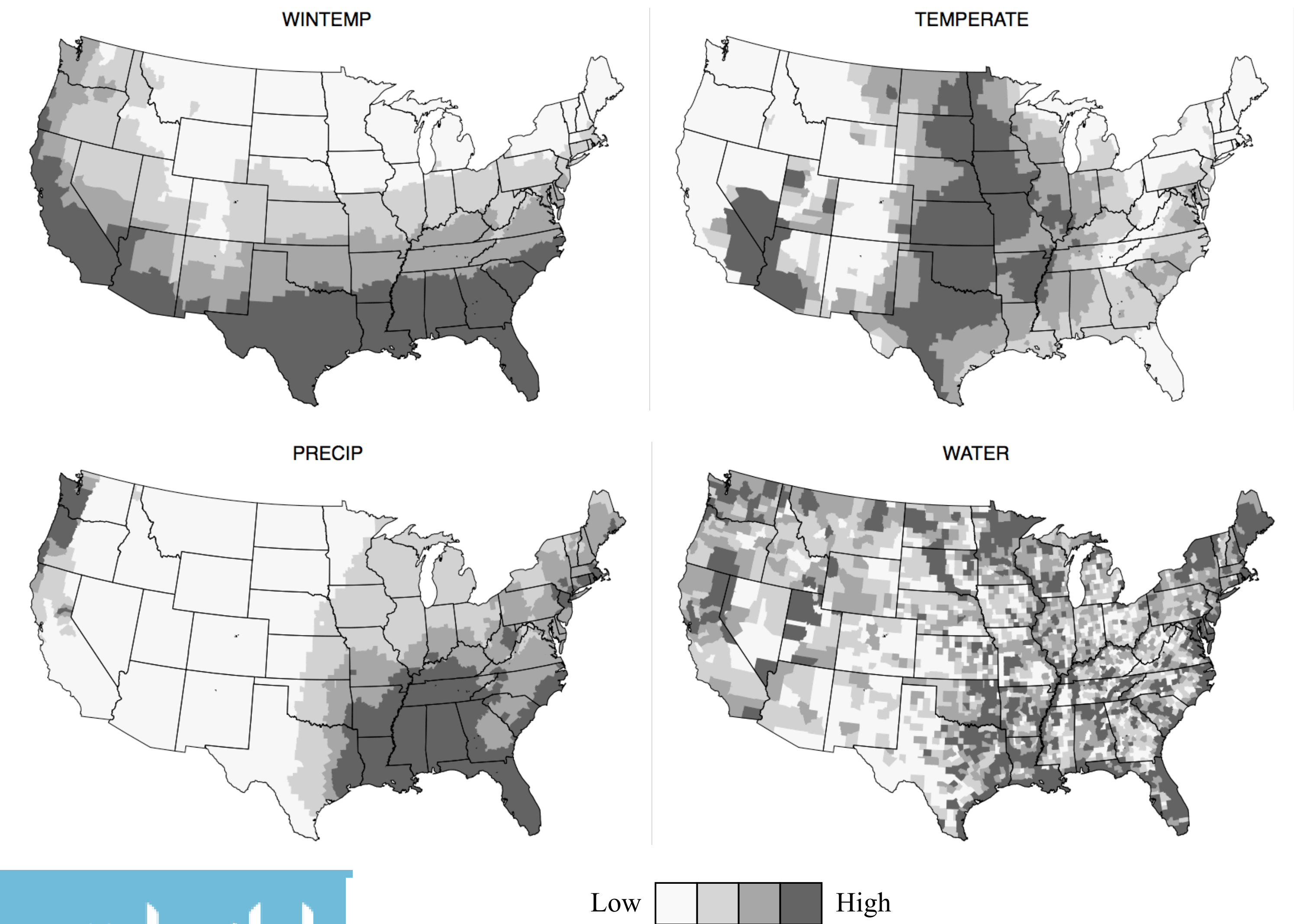
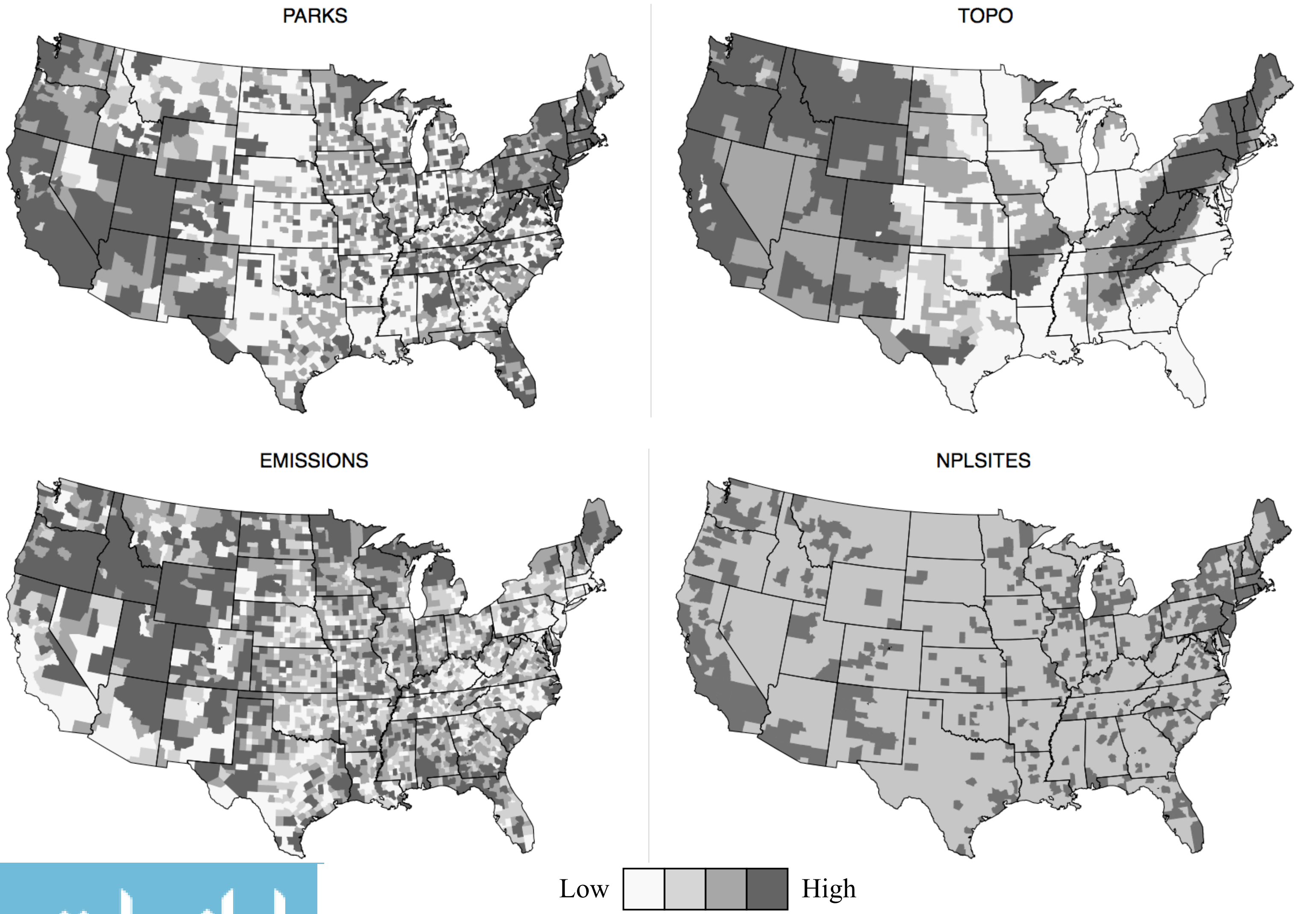


Figure E.1 (cont.). Spatial Distribution of Environmental Amenities: Quartiles
[Note: Lightest color denotes first quartile and darkest denotes forth quartile.]



Appendix F: Distribution of GWR Coefficients for Select Amenities

Figure F.1 Distribution of GWR Coefficients for Environmental Amenities: Standard Deviation from Mean

[Note: Lightest color denotes values <1 s.d. below mean and darkest color denotes values >1 s.d. above mean.]

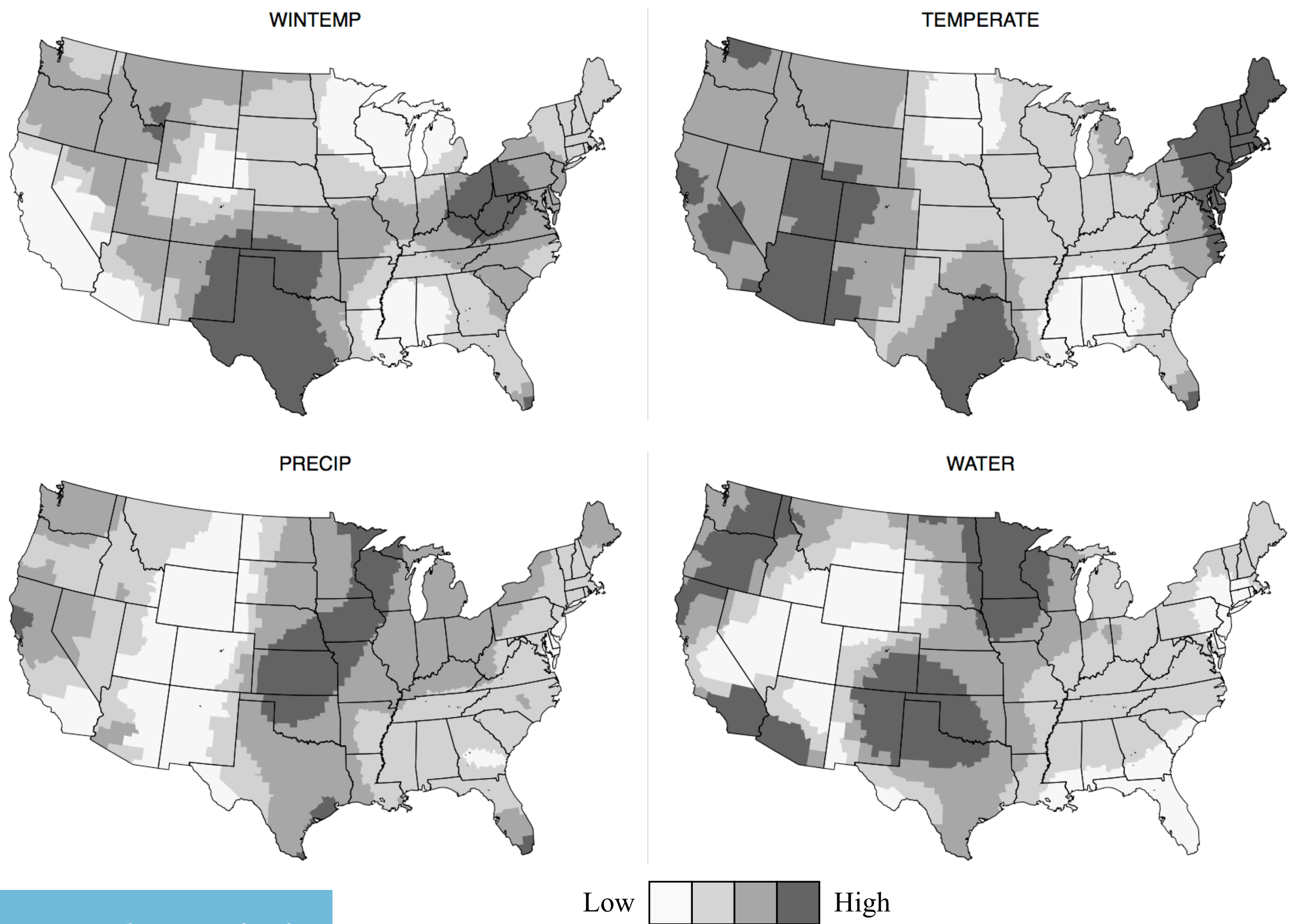
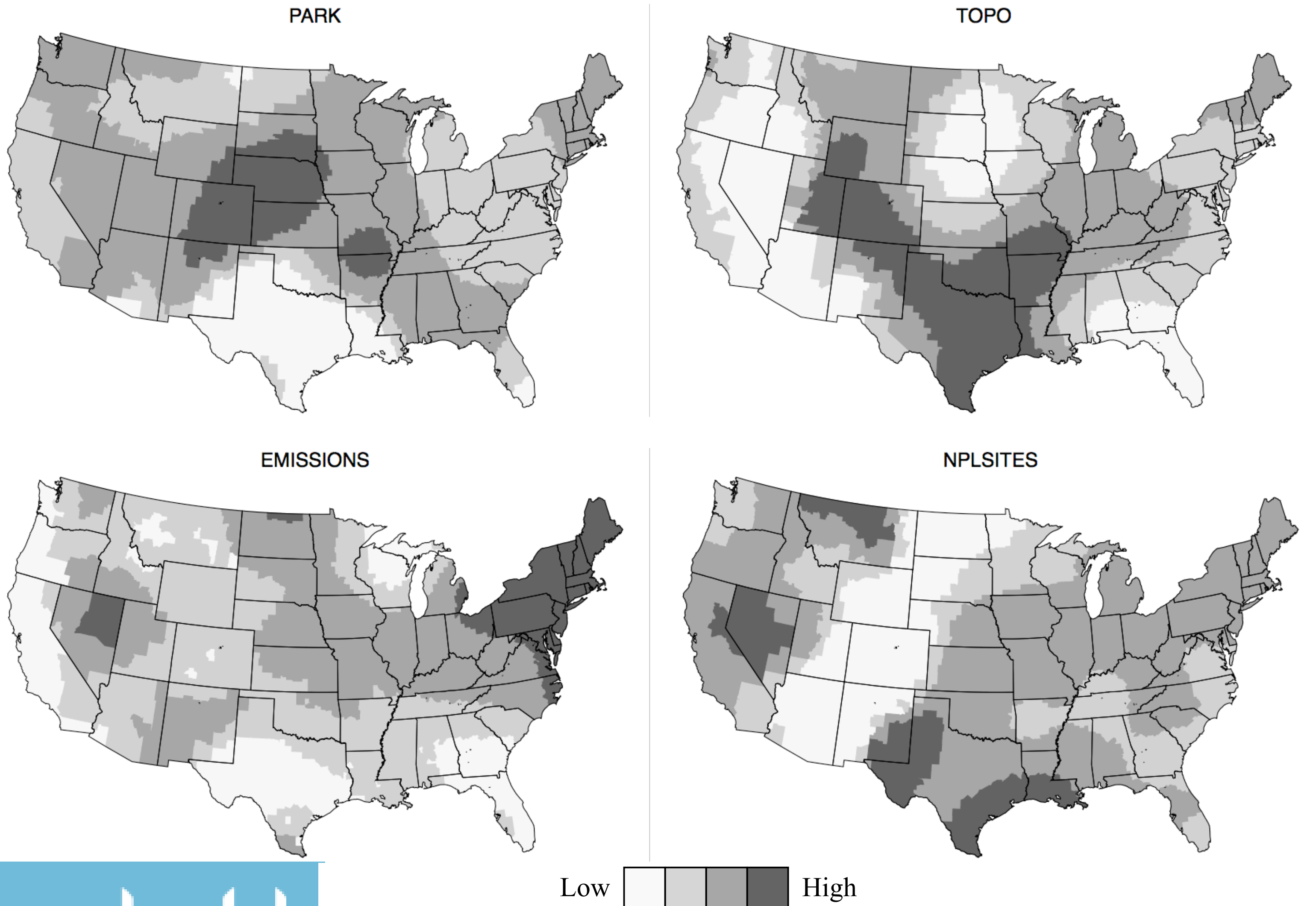


Figure F.1 (cont.). Distribution of GWR Coefficients for Environmental Amenities: Standard Deviation from Mean
[Note: Lightest color denotes values <1 s.d. below mean and darkest color denotes values >1 s.d. above mean.]



Appendix G: Survey Questions Used to Construct the PDVI

The PDVI is constructed using a principle components analysis and six questions from the PSID supplementary questionnaire. These questions are reproduced below.

A. How long were you displaced from your home?[51]

    1. Not Displaced
    2. Displaced for less than a month
    3. Displaced for more than a month but less than six months
    4. Displaced for more than six months

B. Altogether, how much damage to your property or possessions did you experience as a result of Katrina or Rita? Would you say no damage, some damage, a moderate amount of damage, or a lot of damage?

    1. No damage
    2. Some damage
    3. A moderate amount of damage
    4. A lot of damage

C. In the first month after Katrina or Rita, to what extent did you experience a shortage of food? Would you say not at all, a little, some, or a lot?

    1. Not at all
    2. A little
    3. Some
    4. A lot

D. In the first month after Katrina or Rita, to what extent did you experience a shortage of water? Would you say not at all, a little, some, or a lot?

    1. Not at all
    2. A little
    3. Some
    4. A lot

[51] In order to maintain consistency among components of the PDVI, answer categories for this question are constructed from a continuous variable. Answer categories for all other questions are taken directly from the 2007 PSID.

E. In the first month after Katrina or Rita, to what extent did you experience unsanitary conditions, such as inadequate toilets? Would you say not at all, a little, some, or a lot?

1. Not at all
2. A little
3. Some
4. A lot

F. In the first month after Katrina or Rita, to what extent did you experience a loss of electricity? Would you say not at all, a little, some, or a lot?

1. Not at all
2. A little
3. Some
4. A lot

Appendix H: ML Derivation for the Hurdle-Weibull Model

As previously discussed, hurdle models are modified survival (or count) models that assume different processes govern zero outcomes and positive outcomes. Under certain conditions, namely that parameter vectors for the two processes are separable, the log likelihood function is simply the sum of log likelihoods for the binary and survival component (McDowell, 2003). To see this, assume a binary process determines whether the dependent variable is zero (i.e. the household is not displaced) or positive (i.e. the household is displaced). The probability function is given in Equation H1, where π denotes the probability that a household is not displaced.

$$\Pr(Y_1 = y_1) = \begin{cases} \pi, & y_1 = 0 \\ 1-\pi, & y_1 = 1,2,3,\ldots \end{cases} \tag{H1}$$

Likewise assume a survival process, specifically a zero-truncated Weibull process, determines positive outcomes. This density function is presented in Equation H2.

$$\Pr(Y_1 = y_1 | Y_1 \neq 0) = \begin{cases} \left[\frac{\kappa}{\lambda}\left(\frac{y_1}{\lambda}\right)^{\kappa-1}\right]^{y_3} \exp\left(\frac{-y_1}{\lambda}\right)^{\kappa}, & y_1 = 1,2,3,\ldots \\ 0, & \text{otherwise} \end{cases} \tag{H2}$$

As before, κ and λ are parameters of the Weibull distribution and y_3 is a right censor variable. The unconditional probability function is given in Equation H3.

$$\Pr(Y_1 = y_1) = \begin{cases} \pi, & y_1 = 0 \\ (1-\pi)\left[\frac{\kappa}{\lambda}\left(\frac{y_1}{\lambda}\right)^{\kappa-1}\right]^{y_3} \exp\left(\frac{-y_1}{\lambda}\right)^{\kappa}, & y_1 = 1,2,3,\ldots \end{cases} \tag{H3}$$

Assuming that observations are independently and identically distributed, the log likelihood can be written as in Equation H4.

$$\ln L\left(\pi_j,\lambda_j,y_{1j}\right)=\begin{cases}\ln\left\{\pi_j\right\}, & y_{1j}=0\\ \ln\left\{\left(1-\pi_j\right)\left[\dfrac{\kappa}{\lambda_j}\left(\dfrac{y_{1j}}{\lambda_j}\right)^{\kappa-1}\right]^{y_{3j}}\exp\left(\dfrac{-y_{1j}}{\lambda_j}\right)^{\kappa}\right\}, & y_{1j}=1,2,3,\ldots\end{cases}\tag{H4}$$

Modeling the binary process with a probit model and using the link functions described in Equations 4.9 and 4.10 yields the final log-likelihood function.

$$\begin{aligned}\ln L = y_{1j}\ln\left[\Phi(-x_{1j}\beta_1)\right]+(1-y_{1j})\ln\left[\Phi(1-x_{1j}\beta_1)\right]\\ -\left(\frac{y_{2j}}{\lambda}\right)^{\kappa}+y_{3j}\left(\ln(\kappa)-\ln(\lambda)+(\kappa-1)\left[\ln(y_{2j})-\ln(\lambda)\right]\right)\end{aligned}$$

$$\begin{aligned}&\kappa=\exp(\gamma)\\ &\lambda=\exp(\eta)=\exp(x_{2j}\beta_2)\\ &x_{1j}=f(H_j,DS_j,E_j^h,B_j^h,SS_j)\\ &x_{2j}=f(H_j,DS_j,E_j^h,B_j^h,SS_j)\end{aligned}\tag{H5}$$

Here Φ represents the cumulative normal distribution function, x_1 is the vector of explanatory variables that determine the binary process, x_2 is the vector of explanatory variables that determine the survival process. The corresponding vectors of estimated parameters are represented by β_1 and β_2. For the purpose of illustration the dependent variable is decomposed into y_1 (binary process) and y_2 (survival process). Because parameter vectors (i.e. x_{1j} and x_{2j}) are separable, the covariance between β_1 and β_2 is zero. As a result, the joint log-likelihood function is the sum of the log-likelihood functions for a probit model and a truncated-at-zero Weibull model.

Appendix I: Stata and R Codes

Hedonic Housing Regression (Stata): Section 2.6.1

```
use "/PUMS_HOUSING.dta", clear
keep usercost2 owner acre1_10 unit1-unit5 noheat nokit noplm rms bds ybl2-ybl9
        puma_id cbsa_id
sort cbsa_id puma_id
tab cbsa_id, gen(MSA)
reg usercost2 owner acre1_10 unit1-unit5 noheat nokit noplm rms bds ybl2-ybl9 MSA1-
        MSA157 MSA159-MSA296 if owner==1, r
```

Hedonic Wage Regression (Stata): Section 2.6.2

```
use "/PUMS_WAGES.dta", clear
set matsize 800
drop fagep-pwgtp80

//Due to memory limitation, hedonic wage estimates are exported into two files

forvalues i=1/150{
eststo: reg wagerate agep age_sq gender married race_white hsdedu scoledu coledu
        gradedu ser_occ2-prd_occ2 prob1 prob1_sq if cbsa_grp==`i', r
        }
esttab using "/Hedonic Wage Regressions1", b(4) not/*
                */ title("Hedonic Wage Regression by MSA")/*
                */ varwidth(14) modelwidth(7)/*
                */ ar2 r csv nostar
eststo clear

forvalues i=151/296{
eststo: reg wagerate agep age_sq gender married race_white hsdedu scoledu coledu
        gradedu ser_occ2-prd_occ2 prob1 prob1_sq if cbsa_grp==`i', r
        }
esttab using "/Hedonic Wage Regressions2", b(4) not/*
                */ title("Hedonic Wage Regression by MSA")/*
                */ varwidth(14) modelwidth(7)/*
                */ ar2 r csv nostar
eststo clear
```

Conditional Logit Analysis (Stata): Section 2.6.3

```
//The conditional logit model is conducted using 50,000 randomly selected households

use "/CLogit.dta", clear
keep serialno cbsa_grp
```

```
sort serialno
expand 296
sort serialno
egen cbsa=seq(), by(serialno)
gen choice=1 if cbsa_grp==cbsa
replace choice=0 if choice==.
set seed 12093320
sample 19 if choice==0, count by(serialno)

merge m:1 serialno using "CLogit .dta", nogen
do "/CLogit_xbWage.do"      //Predict wages using hedonic wage estimates

gen totinc=exp(xb_wage)
replace totinc=ln(totinc*hrwork)
compress
sort serialno cbsa

forvalues i=1/296{
        gen MSA`i'=0
        replace MSA`i'=1 if cbsa==`i'
        }

gen mig1=0 if pobp==state1|pobp==state2|pobp==state3
gen mig2=0 if b_div==div1|b_div==div2|b_div==div3
gen mig3=0 if b_rgn==rgn1|b_rgn==rgn2|b_rgn==rgn3
mvencode mig1 mig2 mig3, mv(1)
keep serialno choice totinc mig1 mig2 mig3 MSA* cbsa

clogit choice totinc mig1 mig2 mig3 MSA1-MSA157 MSA159-MSA296,
        group(serialno) r
```

<u>ASC Decomposition (Stata): Section 2.6.4</u>

```
//ASC Decompositions are conducted after merging the quality-of-life (QOL) index,
//housing services index, location specific amenities and adjusting the QOL index for
//the price of housing services

use "Decomposition.dta"

reg depvar unemploy pctax vcrime physrate artindex subwaylightrail ptratio
        droprate voters sumtemp update_precip ocean pcemis nplsites locpark2
        MultiFreq, r
reg depvar lnpop unemploy pctax vcrime physrate artindex subwaylightrail ptratio
        droprate voters sumtemp update_precip ocean pcemis nplsites locpark2
        MultiFreq, r
reg depvar lnpop unemploy pctax vcrime physrate artindex subwaylightrail ptratio
```

```
        droprate voters sumtemp update_precip ocean pcemis nplsites locpark2 MultiFreq
        reg1 reg3 reg4, r
ivregress 2sls depvar unemploy pctax vcrime physrate artindex subwaylightrail
        ptratio droprate voters sumtemp update_precip ocean pcemis nplsites locpark2
        MultiFreq reg1 reg3 reg4 (lnpop= in_popden2 lnpop1910), r
rreg depvar lnpop unemploy pctax vcrime physrate artindex subwaylightrail ptratio
        droprate voters sumtemp update_precip ocean pcemis nplsites locpark2 MultiFreq
        reg1 reg3 reg4
ivregress 2sls depvar unemploy pctax vcrime physrate artindex subwaylightrail
        ptratio droprate voters sumtemp update_precip ocean pcemis nplsites locpark2
        MultiFreq reg1 reg3 reg4 (lnpop= in_popden2 lnpop1910) if MultiFreq<58, r
ivregress 2sls depvar unemploy pctax vcrime physrate artindex subwaylightrail
        ptratio droprate voters sumtemp update_precip ocean pcemis nplsites locpark2
        MFD2 MFD3 reg1 reg3 reg4 (lnpop= in_popden2 lnpop1910), r
```

<u>Krinsky-Robb Simulation (R): Section 2.6.4</u>

```
library(foreign)
library(MSBVAR)
library(boot)
library(psych)

betaX=#Coefficient for Amenity x
varX=#Variance for Amenity x
betaI=#Coefficient for Income
varI=#Variance for Income
Ibar=#Median Income

src<-c(1,2)
theta<-c(betaX, betaI)
sbeta<-matrix(theta, ncol=1)
varcov<-c(varX, 0, 0, varI)
scov_b<-matrix(varcov, ncol=2)
sbeta_sim<- rmultnorm(5000, mu=sbeta, vmat=scov_b, tol = 1e-10)
swtp<-function(sbeta_sim){
      b1<-sbeta_sim[,1]
      b2<-sbeta_sim[,2]
      fb=(b1/b2)*Ibar
      return(fb)
      }
swtpvalues5<-eval(swtp(sbeta_sim[,src]))
mean(swtpvalues)
quantiles<-quantile(swtpvalues, c(.025, .05, .5, .95, .975))
quantiles
```

SAR Analysis (Stata): Section 3.6.1

```
use  "\NetMrg.dta"

spmat import wm1 using  "\WMatrix_Q1.gal",  geoda

tab stmg, gen(st)
replace st8=1 if st8==.
mvencode st1-st48, mv(0) override

eststo: spreg ml netmrg exptinc pctax02 IND2-IND6 popden popden_sq metro1 metro2
        hsedu medage vcrime artrec wintemp temperate precip ocean lnwater lnpark
        toposcale emis emis_sq nplsites disarea EFreq HFreq FFreq, id(id) dlmat(wm1)
        elmat(wm1)
eststo: spreg ml netmrg exptinc pctax02 IND2-IND6 popden popden_sq metro1 metro2
        hsedu medage vcrime artrec wintemp temperate precip ocean lnwater lnpark
        toposcale emis emis_sq nplsites disarea EFreq HFreq FFreq div2-div9, id(id)
        dlmat(wm1) elmat(wm1)
eststo: spreg ml netmrg exptinc pctax02 IND2-IND6 popden popden_sq metro1 metro2
        hsedu medage vcrime artrec wintemp temperate precip ocean lnwater lnpark
        toposcale emis emis_sq nplsites disarea EFreq HFreq FFreq st2-st48, id(id)
        dlmat(wm1) elmat(wm1)
eststo: spreg gs2sls netmrg exptinc pctax02 IND2-IND6 popden popden_sq metro1
        metro2 hsedu medage vcrime artrec wintemp temperate precip ocean lnwater
        lnpark toposcale emis emis_sq nplsites disarea EFreq HFreq FFreq st2-st48, id(id)
        dlmat(wm1) elmat(wm1)
```

GWR Analysis (Stata): Section 3.6.2

```
use "/NetMrg.dta"

gwr netmrg exptinc pctax02 IND2-IND6 popden popden_sq hsedu medage vcrime artrec
        wintemp temperate precip lnwater lnpark toposcale emis emis_sq nplsites
        MultiFreq, east(xcoord) north(ycoord) bandwidth (378838.58) test dots reps(1)
        saving("/GWR Results") replace
```

Mental Health Analysis (Stata): Section 4.2.4

```
use "/PSID.dta"

ivprobit PTSD female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 (PDVI=DSI near_miles), first r
ivprobit PTSD female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 smoke07 regdrink07 inact07 chronic07 prior07 (PDVI=DSI
        near_miles), first r
```

```
ivprobit PTSD female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 smoke07 regdrink07 inact07 chronic07 prior07 SSI (PDVI=DSI
        near_miles), first r
ivprobit PTSD2 female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 smoke07 regdrink07 inact07 chronic07 prior07 SSI (PDVI=DSI
        near_miles), first r
ivprobit PHQ female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 (PDVI=DSI near_miles), first r
ivprobit PHQ female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 smoke07 regdrink07 inact07 chronic07 prior07 (PDVI=DSI
        near_miles), first r
ivprobit PHQ female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 smoke07 regdrink07 inact07 chronic07 prior07 SSI (PDVI=DSI
        near_miles), first r
ivprobit PHQ2 female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 smoke07 regdrink07 inact07 chronic07 prior07 SSI (PDVI=DSI
        near_miles), first r
ivprobit GAD female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 (PDVI=DSI near_miles), first r
ivprobit GAD female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 smoke07 regdrink07 inact07 chronic07 prior07 (PDVI=DSI
        near_miles), first r
ivprobit GAD female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 smoke07 regdrink07 inact07 chronic07 prior07 SSI (PDVI=DSI
        near_miles), first r
ivprobit GAD2 female07 black07 age07 married07 child07 edu07 unemploy07 lninc07
        ownhome07 smoke07 regdrink07 inact07 chronic07 prior07 SSI (PDVI=DSI
        near_miles), first r
```

Displacement Analysis (Stata): Section 4.3.4

```
use "/PSID.dta"

stset duration, failure(censor)

capture program drop probit_weib
program define probit_weib
        args lnf xa leta lgam
        tempvar pr p M R
        quietly gen double `pr'=1-(normprob(-`xa'))
        quietly gen double `p'=exp(`lgam')
        quietly gen double `M'=($ML_y2*exp(-`leta'))^`p'
        quietly gen double `R'=ln($ML_y2)-`leta'
        quietly replace `lnf'=cond($ML_y1==0, ln(1-`pr'), ln(`pr')-`M'+$ML_y3*(`lgam'-
`leta'+(`p'-1)*`R'))
end
```

```
ml model lf probit_weib (probit: displace=dam2 dam3 dam4 DSI near_miles) (weibull:
        duration censor=dam2 dam3 dam4 DSI near_miles) /lnp, vce(r)
ml maximize
estat ic

ml model lf probit_weib (probit: displace=dam2 dam3 dam4 DSI near_miles female05
        black05 age05 married05 edu05 unhealthy05 unemploy05 child05 lninc05
        ownhome05 mobhome05) (weibull: duration censor=dam2 dam3 dam4 DSI
        near_miles female05 black05 age05 married05 edu05 unhealthy05 unemploy05
        child05 lninc05 ownhome05 mobhome05) /lnp, vce(r)
ml maximize
estat ic

ml model lf probit_weib (probit: displace=dam2 dam3 dam4 DSI near_miles female05
        black05 age05 married05 edu05 unhealthy05 unemploy05 child05 lninc05
        ownhome05 mobhome05 SSI) (weibull: duration censor=dam2 dam3 dam4 DSI
        near_miles female05 black05 age05 married05 edu05 unhealthy05 unemploy05
        child05 lninc05 ownhome05 mobhome05 SSI stayfam insuramnt05 remit05) /lnp,
        vce(r)
ml maximize
estat ic

ml model lf probit_weib (probit: displace=dam2 dam3 dam4 DSI near_miles female05
        black05 age05 married05 edu05 unhealthy05 unemploy05 child05 lninc05
        ownhome05 mobhome05 SSI st1 st3) (weibull: duration censor=dam2 dam3
        dam4 DSI near_miles female05 black05 age05 married05 edu05 unhealthy05
        unemploy05 child05 lninc05 ownhome05 mobhome05 SSI stayfam insuramnt05
        remit05 st1 st3) /lnp, vce(r)
ml maximize
estat ic
```

References

Adeola, F. 2009. Mental Health and Psychosocial Distress Sequelae of Katrina: An Empirical Study of Survivors. Research in Human Ecology 16, 195-210.

Adger, N. 1999. Social Vulnerability to Climate Change and Extremes in Coastal Vietnam. World Development 27(2), 249-269.

Allen Klaiber, H., Phaneuf, D. 2010. Valuing Open Space in a Residential Sorting Model of the Twin Cities. Journal of Environmental Economics and Management 60 (2), 57-77.

Anselin, L. 2001. Spatial Econometrics. In: Baltagi, B. (Ed.), A Companion to Theoretical Econometrics, Blackwell, Oxford, UK, pp. 310-330.

Arraiz, I., Druker, D., Kelejian, H., Prucha, I. 2010. A Spatial Cliff-Ord-Type Model with Heteroskedastic Innovations: Small and Large Sample Results. Journal of Regional Science 50 (2), 592-614.

Averina, M., Nilssen, O., Brenn, T., Brox, J., Arkhipovsky, V., Kalinin, A. 2005. Social and Lifestyle Determinants of Depression, Anxiety, Sleeping Disorder and Self-Evaluated Quality of Like in Russia: A Population-Based Study in Arkhangelsk. Social Psychiatry and Psychiatric Epidemiology 40, 511-518.

Baker, J., Shaw, W., Bell, D., Brody, S., Riddel, M., Woodward, R., Neilson, W. 2009. Explaining Subjective Risks of Hurricanes and the Role of Risks in Intended Moving and Location Choice Models. Natural Hazards Review 10 (3), 102-112.

Bal, S., Crombez, G., Bourdeaudhuij, I., Van Oost, P. 2009. Symptomatology in Adolescents Following Initial Disclosure of Sexual Abuse: The Roles of Crisis Support, Appraisals and Coping. Child Abuse and Neglect 33(10), 717-727.

Bauer, T., Epstein, G., Gang, Ira. 2005. Enclaves, Language, and Location Choice of Migrants. Journal of Population Economics 18, 649-662.

Bayer, P., Ferreira, F., McMillan, R. 2004. Tiebout Sorting, Social Multipliers and the Demand for School Quality. NBER Working Paper 10871.

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

Bayer, P., McMillan, R. 2005. Racial Sorting and Neighborhood Quality. NBER Working Paper 11813.

Bayer, P., McMillan, R., Rueben, K. 2005. Residential Segregation in General Equilibrium. NBER Working Paper 11095.

Beaudoin, C. 2007. News, Social Capital and Health in the Context of Katrina. Journal of Health Care for the Poor and Underserved 18, 418-430.

Berger, M., Blomquist, G., Kenkel, D., Tolley, G. 1987. Valuing Changes in Health Risks: a Comparison of Alternative Measures. Southern Economic Journal 53 (4), 967-984.

Berk, R. 1990. A Primer on Robust Regression. In: Fox, J., Long, S. (Eds.), Modern Methods of Data Analysis, Sage, Newbury Park, CA.

Bernknopf, R., Brookshire, D., Thayer, M. 1990. Earthquake and Volcano Hazard Notices: An Economic Evaluation of Changes in Risk Perceptions. Journal of Environmental Economics and Management 18 (1), 35-49.

Beron, K., Murdoch, J., Thayer, M., Vijverberg, W. 1997. An Analysis of the Housing Market Before and After the 1989 Loma Prieta Earthquake. Land Economics 73 (1), 101-113.

Berry, S., Levinsohn, J., Pakes, A. 1995. Automobile Prices in Market Equilibrium Econometrica: Journal of the Econometric Society 63 (4), 841-890.

Berry, S., Linton, O., Pakes, A. 2004. Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems. Review of Economic Studies 71 (3), 613-654.

Bethell, J., Rhodes, A., Bondy, S., Low, W., Guttmann, A. 2010. Repeat Self-Harm: Application of Hurdle Models. The British Journal of Psychiatry 196, 243-244.

Bilgic, A., Florkowski, W. 2007. Application of a Hurdle Negative Binomial Count Data Model to Demand for Bass Fishing in the Southeastern United States. Journal of Environmental Management 83(4), 478-490.

Bin, O., Kruse, J., Landry, C. 2008. Flood Hazards, Insurance Rates, and Amenities: Evidence from the Coastal Housing Market. Journal of Risk and Insurance 75 (1), 63-82.

Bin, O., Polasky, S. 2004. Effects of Flood Hazards on Property Values: Evidence Before and After Hurricane Floyd. Land Economics 80 (4), 490-500.

Bödvarsdóttir, I., Elklit, A. 2004. Psychological Reactions in Icelandic Earthquake Survivors. Scandinavian Journal of Psychology 45, 3-13.

Bohara, A., Kreig, R. 1996. A Poisson Hurdle Model of Migration Frequency. The Journal of Regional Analysis and Policy 26(1), 37-45.

Brookshire, D., Schulze, W., Tschirhart, J., Thayer, M., Hageman, R., Pazand, R., Ben-David, S. 1980. Methods Development for Valuing Hazards Information. Technical Report. Washington: U.S. Geological Survey.

Brookshire, D., Thayer, M., Tschirhart, J., Schulze, W. 1985. A Test of the Expected Utility Model: Evidence from Earthquake Risks. The Journal of Political Economy 93 (2), 369-389.

Bruch, E., Mare, R. 2011. Methodological Issues in the Analysis of Residential Preferences and Residential Mobility. University of Michigan, Population Studies Center Research Report 11-725.

Cao, H., McFarlane, A., Klimidis, S. 2003. Prevalence of Psychiatric Disorder Following the 1998 Yun Nan (China) Earthquake-the First 5-Month Period. Social Psychiatry and Psychiatric Epidemiology 38, 204-212.

Cardona, O., Ordaz, M., Yamin, L., Marulanda, M., Barbat, A. Earthquake Loss Assessment for Integrated Disaster Risk Management. Journal of Earthquake Engineering 12(S2), 48-59.

Carling, G., Mills, E. 1987 The Determinants of County Growth. Journal of Regional Science 27 (1), 39-54.

Cleves, M., Gould, W., Gutierrez, R., Marchenko, Y. 2008. An Introduction to Survival Analysis Using Stata. Stata Press, College Station, TX, pp. 221-233.

Curtis, A., Mills, J., Leitner, M. 2007. Katrina and Vulnerability: The Geography of Stress. Journal of Health Care for the Poor and Underserved 18(2), 315-330.

Dahl, G. 2002. Mobility and the Return to Education: Testing a Roy Model with Multiple Markets. Econometrica 70 (6), 2367-2420.

Davies, P., Greenwood, M., Li, H. 2001. A Conditional Logit Approach to U.S. State-to-State Migration. Journal of Regional Science 41 (2), 337-360.

Davis, T., Grills-Taquechel, A., Ollendick, T. 2010. The Psychological Impact From Hurricane Katrina: Effect of Displacement and Trauma Exposure on University Students. Behavior Therapy 41, 340-349.

Day, K., Winer, S. 2006. Policy-Induced Internal Migration: An Empirical Investigation of the Canadian Case. International Tax and Public Finance 13 (5), 535-564.

Dilley, M., Chen, R., Deichmann, U., Lerner-Lam, A., Arnold, M. 2005. Natural Disaster Hotspots: A Global Risk Analysis. World Bank, Washington DC.

Donovan, G., Champ, P., Butry, D. 2007. Wildfire Risk and Housing Prices: A Case Study from Colorado Springs. Land Economics 83 (2), 217-233.

Drukker, D., Prucha, I., Raciborski, R. 2011. Maximum-Likelihood and Generalized Spatial Two-Stage Least-Squares Estimators for a Spatial-Autoregressive Model with Spatial-Autoregressive Disturbances. Technical report, Stata.

Dun, O. 2011. Migration and Displacement Triggered by Floods in the Mekong Delta. International Migration. 49, 201-223

Ehrlich, I., Becker, G. 1972. Market Insurance, Self-Insurance, and Self-Protection. The Journal of Political Economy 80 (4), 623-648.

Elklit, A., Hartvig, T., Christiansen, M. 2007. Psychological Sequelae in Parents of Extreme Low and Very Low Birth Weight Infants. Journal of Clinical Psychology in Medical Settings 14, 238-247.

Fields, G. 2004. Regression-Based Decompositions: A New Tool for Managerial Decision-Making. Unpublished paper, Cornell University, Ithaca, NY.

Filmer, D., Pritchett, L. 2001. Estimating Wealth Effects Without Expenditure Data—Or Tears: An Application To Educational Enrollments In States Of India. Demography 38 (1), 115-132.

Finney, M., Goetzke, F., Yoon, M. 2011. Income Sorting and the Demand for Clean Air: Evidence from Southern California. Land Economics 87 (1), 19-27

Freeman, P., Martin, L., Linnerooth-Bayer, J., Mechler, R., Pflug, G., Warner, K. 2003. Disaster Risk Management: National Systems for the Comprehensive Management of Disaster Risk and Financial Strategies for Natural Disaster Reduction. Inter-American Development Bank, Washington, DC.

Fronstin, P., Holtmann, A. 1994. The Determinants of Residential Property Damage Caused by Hurricane Andrew. Southern Economic Journal 61 (2), 387-397.

Fussell, E., Sastry, N., VanLandingham, M. 2010. Race, Socioeconomic Status, and Return Migration to New Orleans after Hurricane Katrina. Population and Environment 31, 20-42.

Galea, S., Ahern, J., Nandi, A., Tracy, M., Beard, J., Vlahov, D. 2007. Urban Neighborhood Poverty and the Incidence of Depression in a Population-Based Cohort Study. Annals of Epidemiology 17(3), 171-179.

Galea, S., Melissa, T., Norris, F., Coffey, S. 2008. Financial and Social Circumstances and the Incidence and Course of PTSD in Mississippi During the First Two Years After Hurricane Katrina. Journal of Traumatic Stress 21(4), 357-368.

Galea, S., Vlahov, D., Resnick, H., Ahern, J., Susser, E., Gold, J., Bucuvalas, M., Kilpatrick, D. 2003. Trends of Probable Post-Traumatic Stress Disorder in New York City After the September 11 Terrorist Attacks. American Journal of Epidemiology 158(6), 514-524.

Gawande, K., Bandyopadhyay, U. 2000. Is Protection for Sale? Evidence on the Grossman Helpman Theory of Endogenous Protection. Review of Economics and Statistics 82 (1), 139-152.

Gawande, K., Bohara, A., Berrens, R., Wang, P. 2000. Internal Migration and the Environmental Kuznets Curve for US Hazardous Waste Sites. Ecological Economics 33, 151-166.

Glaeser, E., Gottlieb, J. 2008. The Economics of Place-Making Policies. NBER Working Paper 14373.

Greene, W. 2003. Econometric Analysis 5th Edition. Prentice Hall, Upper Saddle River, New Jersey, pp. 790-801.

Greenough, G., McGeehin, M., Bernard, S., Trtanj, J., Riad, J., Engelberg, D. 2001. The Potential Impacts of Climate Variability and Change on Health Impacts of Extreme Weather Events in the United States. Environmental Health Perspectives 109 (Suppl. 2), 191-198.

Groen, J., Polivka, A. 2010. Going Home After Hurricane Katrina: Determinants of Return Migration and Changes in Affected Areas. Demography 47(4), 821-844.

Hallstrom, D., Smith, V. 2005. Market Responses to Hurricanes. Journal of Environmental Economics and Management 50 (3), 541-561.

Harrison, D., Smersh, G., Schwartz, A. 2001. Environmental Determinants of Housing Prices: the Impact of Flood Zone Status. Journal of Real Estate Research 21 (1), 3-20.

Hoehn, J., Berger, M., Blomquist, G. 1987. A Hedonic Model of Interregional Wages, Rents, and Amenity Values. Journal of Regional Science, 27 (4), 605-620.

Holladay, J., Schwartz, J. 2010. Flooding the Market: The Distributional Consequences of the NFIP. Institute for Policy Integrity, Policy Brief No. 7. New York University School of Law New York, NY.

Hori, M., Schafer, M. 2010. Social Costs of Displacement in Louisiana After Hurricanes Katrina and Rita. Population and Environment 31, 64-86.

Horney, J., MacDonald, P., van Willigen, M., Berke, P., Kaufman, J. 2010. Individual Actual or Perceived Property Flood Risk: Did it Predict Evacuation from Hurricane Isabel in North Carolina, 2003? Risk Analysis 30 (3), 501-511.

IPCC. 2011. Summary for Policymakers. In: Field, C., Barros, V., Stocker, T., Qin, D., Dokken, D., Ebi, K., Mastrandrea, M., Mach, K., Plattner, G., Allen, S., Tignor, M., Midgley, P. (Eds.), Intergovernmental Panel on Climate Change Special Repost on Managing the Risks of Extreme Events and Disaster to Advance Climate Change Adaptation, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Jaspers, S., Shoham, J. 1999. Targeting the Vulnerable: A Review of the Necessity and Feasibility of Targeting Vulnerable Households. Disasters 23(4), 359-372.

Joseph, S., Williams, R., Yule, W. 1992. Crisis Support, Attributional Style, Coping Style, and Post-Traumatic Symptoms. Personality and Individual Differences 13(11), 1249-1251.

Kaplan, E., Meier, P. 1958. Nonparametric Estimation From Incomplete Observations. Journal of the American Statistical Association 53, 457-481.

Kario, K., McEwen, B., Pickering, T. 2003. Disasters and the Heart: A Review of the Effects of Earthquake-Induced Stress on Cardiovascular Disease. Hypertension Research 26, 355-367.

Kelejian, H., Prucha, I. 1988. A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. Journal of Real Estate Finance and Economies 17 (1), 99-121.

Knabb, R., Rhome, R., Brown, D. 2005. Tropical Cyclone Report: Hurricane Katrina. National Hurricane Center, National Weather Service, National Oceanic and Atmospheric Administration. [online document] Accessed 27 December 2011. <http://www.nhc.noaa.gov/pdf/TCR-AL122005_Katrina.pdf>.

Krinsky, I., Robb, L. 1986. On Approximating the Statistical Properties of Elasticities. The Review of Economics and Statistics, 68 (4), 715-719.

Kroenke, K., Spitzer R., Williams, J. 2001. The PHQ-9. Journal of General Internal Medicine 16, 606-613.

Krugman, P. 1991. Increasing Returns and Economic Geography. Journal of Political Economy 99 (3), 483-499.

Kunreuther, H. 2006. Disaster Mitigation and Insurance: Learning from Katrina. The Annals of the American Academy of Political and Social Science 604 (1), 208-227.

Landry, C., Bin, O., Hindsley, P., Whitehead, J., Wilson, K. 2007. Going Home: Evacuation-Migration Decisions of Hurricane Katrina Survivors. Southern Economic Journal 74(2), 326-343.

Loomis, J. 2004. Do Nearby Forest Fires Cause a Reduction in Residential Property Values? Journal of forest economics 10 (3), 149-157.

Lu, J., Zhang, L. 2010. Evaluation of Parameter Estimation Methods for Fitting Spatial Regression Models. Forest Science 56 (5), 505-514.

MacDonald, D., Murdoch, J., White, H. 1987. Uncertain Hazards, Insurance, and Consumer Choice: Evidence from Housing Markets. Land Economics 63 (4), 361-371.

Madrid, P., Grant, R., Reilly, M., Redlener, N. 2006. Challenges in Meeting Immediate Emotional Needs: Short-term Impact of a Major Disaster on Children's Mental Health: Building Resiliency in the Aftermath of Hurricane Katrina. Pediatrics 117(5), S448-S453.

McDowell, A. 2003. From the Help Desk: Hurdle Models. The Stata Journal 3(2), 178-184.

McFadden, D. 1978. Modeling the Choice of Residential Location. In: Karlquist, A. (Ed.), Spatial Interaction Theory and Residential Location. North Holland, Amsterdam, pp. 75-96.

McGranahan, D. 1999. Natural Amenities Drive Rural Population Change. Agricultural Economic Report No. 781, Food and Rural Economics Division Report, U.S. Department of Agriculture, Economic Research Service, Washington, D.C. [online document] Accessed 22 November 2011 <http://www.ers.usda.gov/publications/aer781/>

McGranahan, D. 2008. Landscape Influence on Recent Rural Migration in the U.S. Landscape and Urban Planning 85, 228-240.

McKenzie, R., Levendis, J. 2010. Flood Hazards and Urban Housing Markets: The Effects of Katrina on New Orleans. The Journal of Real Estate Finance and Economics 40 (1), 62-76.

Miguel-Tobal, J., Cano-Vindel, A., Gonzalez-Ordi, H., Iruarrizaga, I., Rudenstine, S., Vlahov, D., Galea, S. 2006. PTSD and Depression After the Madrid March 11 Train Bombings. Journal of Traumatic Stress 19(1), 69-80.

Morrow, B. 1999. Identifying and Mapping Community Vulnerability. Disasters 23(1), 1-18.

Mueser, P., Graves, P. 1995. Examining the Role of Economic Opportunity and Amenities in Explaining Population Redistribution. Journal of Urban Economics 37, 176-200.

National Hurricane Center. 2007. Tropical Weather Summary—2005. National Hurricane Center, National Weather Service, National Oceanic and Atmospheric Administration. [online document] Accessed 27 December 2011. <http://www.nhc.noaa.gov/archive /2005/tws/ MIATWSAT_nov_final.shtml>

NOAA. 2011. NOAA's 1981-2010 Climate Normals. [online document] Accessed 20 October 2011. < http://lwf.ncdc.noaa.gov/oa/climate/normals/usnormals.html>

Norris, F., Friedman, M., Watson, P., Byrne, C., Diaz, E., Kaniasty, K. 2002. 60,000 Disaster Victims Speak: Part 1. An Empirical Review of the Empirical Literature, 1981-2001. Psychiatry 65(3), 207-239.

Norris, F., Murphy, A., Baker, C., Perilla. J. 2004. Postdisaster PTSD Over Four Waves of a Panel Study of Mexico's 1999 Flood. Journal of Traumatic Stress 17, 283-292.

Norris, F., Uhl, G. 1993. Chronic Stress as a Mediator of Acute Stress: The Case of Hurricane Hugo. Journal of Applied Social Psychology 23, 1263-1284.

North, C., Nixon, S., Shariat, S., Mallonee, S., McMillen, J., Spitznagel, E. 1999. Psychiatric Disorders Among Survivors of the Oklahoma City Bombing. Journal of American Medical Associations 282, 755-762.

Önder, Z., Dokmeci, V., Keskin, B. 2004. The Impact of Public Perception of Earthquake Risk on Istanbul's Housing Market. Journal of Real Estate Literature 12 (2), 181-196.

Ord, J. 1975. Estimation Methods for Models of Spatial Interaction. Journal of the American Statistical Association 70, 120-126.

Partridge, M. 2010. The Dueling Models: NEG vs Amenity Migration in US Engines of Growth. Papers in Regional Science 89 (3), 513-537.

Partridge, M., Rickman, D., Ali, K., Olfert, M. 2008. The Geographic Diversity of U.S. Nonmetropolitan Growth Dynamics: A Geographically Weighted Regression Approach. Land Economics 84 (2), 241-266.

Perry, M. 2006. Domestic Net Migration in the United States: 2000 to 2004. Current Population Reports P25-1135, U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau, Washington, D.C. [online document] Accessed 22 November 2011. <http://www.census.gov/prod/2006pubs/p25-1135.pdf>

Poston, D., Zhang, L., Gotcher, D., Gu, Y. 2009. The Effect of Climate on Migration: United States, 1995-2000. Social Science Research 38 (3), 743-753.

Qouta, S., Punamaki, R., Sarraj, E. 2003. Prevalence and Determinants of PTSD Among Palestinian Children Exposed to Military Violence. European Child and Adolescent Psychiatry 12, 265-272.

Quiggin, J. 1992. Risk, Self-Protection and Ex Ante Economic Value-Some Positive Results. Journal of Environmental Economics and Management 23 (1), 40-53.

Roback, J. 1982. Wages, Rents, and the Quality of Life. Journal of Political Economy 90 (6), 1257-1278.

Rupasingha, A., Goetz, S. 2004. County Amenities and Net Migration. Agricultural and Resource Economics Review 33 (2), 245-254.

Saldaña-Zorrialla, S., Sandberg, K. 2009. Spatial Econometric Model of Natural Disaster Impacts on Human Migration in Vulnerable Regions of Mexico. Disasters 33 (4), 591-607.

Savageau, D. 2007. Places Rated Almanac: The Classic Guide for Finding Your Best Places to Live in America. Places Rated Books, Washington DC.

Shogren, J., Crocker, T. 1999. Risk and Its Consequences. Journal of Environmental Economics and Management 37 (1), 44-51.

Simmons, K., Kruse, J., Smith, D. 2002. Valuing Mitigation: Real Estate Market Response to Hurricane Loss Reduction Measures. Southern Economic Journal 68 (3), 660-671.

Smith, M., Redd, W., DuHamel, K., Vickberg, S., Ricketts, P. 1999. Validation of the PTSD Checklist-Civilian Version in Survivors of Bone Marrow Transplantation. Journal of Traumatic Stress 12(3), 485-499.

Smith, V. 1986. Benefit Analysis for Natural Hazards. Risk Analysis 6 (3), 325-334.

Sinha, Paramita. 2008. The Value of Climate Amenities: A Disequilibrium Approach. Unpublished Doctoral Dissertation, University of Maryland, College Park, MD.

Speyrer, J., Ragas, W. 1991. Housing Prices and Flood Risk: An Examination Using Spline Regression. The Journal of Real Estate Finance and Economics 4 (4), 395-407.

Spitzer, R., Kroenke, K., Williams, J., Lowe, B. 2006. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. Archives of Internal Medicine 166, 1092-1097.

StataCorp LP. 2009. Stata Base Reference Manual Release 11. Stata Press, College Station, TX, pp. 724-736.

Steinberg, J., Aysha, A., Kowalski, M., Kukar, A., Suma, V., Volka, M., Ehlert, F., Herweg, B., Donnelly, J., Philip, J., Reed, G., Rozanski, A. 2004. Increased Incidence of Life-threatening Ventricular Arrhythmias in Implantable Defibrillator Patients After the World Trade Center Attack. Journal of the American College of Cardiology 44, 1261-1264.

Taylor, L. 2003. The Hedonic Method. In: Champ, P., Boyle, K., Brown, T. (Eds.), A Primer on Nonmarket Valuation. Kluwer Academic Publishers, Boston, pp. 331-394.

Tiebout, C. 1956. A Pure Theory of Local Expenditures. Journal of Political Economy 64 (5), 416-424.

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

Tra, C. 2010. A Discrete Choice Equilibrium Approach to Valuing Large Environmental Changes. Journal of Public Economics 94 (1-2), 183-196.

Trevisan, M., Celentano, E., Meucci, C., Farinaro, E., Jossa, F., Krogh, V., Giumetti, D., Panico, S., Scottoni, A., Mancini, M. 1986. Short-Term Effect of Natural Disasters on Coronary Heart Disease Risk Factors. Arteriosclerosis 6, 491-494.

UNISDR. 2009. 2009 UNISDR Terminology on Disaster Risk Reduction. United Nations International Strategy for Disaster Reduction, Geneva, Switzerland.

UNISDR. 2011. Global Assessment Report on Disaster Risk Reduction. United Nations International Strategy for Disaster Reduction, Geneva, Switzerland.

U.S. Census Bureau. 2008. A Compass for Understanding and Using American Community Survey Data: What General Data Users Need to Know. U.S Government Printing Office, Washington DC.

Uscher-Pines, L. 2008. Health Effects of Relocation Following Disaster: A Systematic Review of the Literature. Disasters 33 (1), 1-22.

van Aalst, M. 2006. The Impacts of Climate Change on the Risk of Natural Disasters. Disasters 30 (1), 5-18.

Varano, S., Schafer, J., Cancion, J., Decker, S., Greene, J. 2010. A Tale of Three Cities: Crime and Displacement after Hurricane Katrina. Journal of Criminal Justice 38, 42-50.

Vigdor, J. 2007. The Katrina Effect: Was There a Bright Side to the Evacuation of Greater New Orleans? The B.E. Journal of Economic Analysis and Policy 7 (1), Article 64. [online document] Accessed 22 November 2011 <http://www.bepress.com/bejeap/vol7/iss1/art64>

Walker, E., Newman, E., Dobie, D., Ciechanowski, P., Katon, W. 2002. Validation of the PTSD Checklist in an HMO Sample of Women. General Hospital Psychiatry 24, 375-380.