

CHOICE DIFFICULTY AND RISK PERCEPTIONS IN ENVIRONMENTAL  
ECONOMICS

by

ERIC NIGEL DUQUETTE

A DISSERTATION

Presented to the Department of Economics  
and the Graduate School of University of Oregon  
in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy

September 2010

**University of Oregon Graduate School**

**Confirmation of Approval and Acceptance of Dissertation prepared by:**

Eric Duquette

Title:

"Choice Difficulty and Risk Perceptions in Environmental Economics"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Economics by:

Trudy Cameron, Chairperson, Economics  
William Harbaugh, Member, Economics  
Jason Lindo, Member, Economics  
Ulrich Mayr, Outside Member, Psychology

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

September 4, 2010

Original approval signatures are on file with the Graduate School and the University of Oregon Libraries.

© 2010 Eric N. Duquette

## An Abstract of the Dissertation of

Eric N. Duquette

for the degree of

Doctor of Philosophy

in the Department of Economics

to be taken

September 2010

Title: CHOICE DIFFICULTY AND RISK PERCEPTIONS IN ENVIRONMENTAL  
ECONOMICS

Approved: \_\_\_\_\_

Trudy Ann Cameron, Ph.D.

Economists typically assume that individuals behave in accordance with rational choice theory. In practice, however, individual behavior can deviate from the predictions of models founded upon basic economic theory. The extent to which these deviations are important to individual decision-making in environmental economics, and thus to the development of sound environmental policies, is not fully understood. The objective in this dissertation research is to investigate potential deviations from rational choice behavior in some environmental economics contexts and to identify their relevance to environmental policy.

Chapter I uses a stated-preference survey for the valuation of environmental health-risk reductions in which respondents rate the subjective difficulty of each key choice they are asked to consider. Existing literature identifies many potential categories

of biases in the empirically estimated valuation of non-market goods in stated-preference research. One potential source of bias stems from the “objective complexity” of the choice scenario. I find that existing objective measures of choice set complexity do not fully explain subjective choice difficulty ratings in this valuation survey. Instead, subjective difficulty appears to result from the interplay among objective complexity, preferences, and cognitive resource constraints.

In Chapter II, I consider the possible consequences of choice difficulty from the standpoint of neuroeconomics. Within the scope of neuroeconomics, one can identify some neurobiological correlates of economic decision-making activity. I study the apparent effects of choice difficulty on the neurobiological encoding of individuals’ value assessments. Information from this study provides a neurological basis for deviations from simple economic theory based on conventional models of rational choice.

Chapter III examines risk perceptions that may influence individuals’ decisions to migrate within the U.S. to reduce potential health and economic risks related to climate change. My analysis treats historical patterns of migration among counties as a function of varying spatial and temporal patterns in tornado activity, along with other spatially and temporally delineated variables intended to capture the evolution of subjective perceptions of these tornado risks. Results suggest that the perception of risk from extreme weather events can have a small but statistically discernible effect on migration behavior across sociodemographic groups for both out-migrants and in-migrants.

## CURRICULUM VITAE

NAME OF AUTHOR: Eric N. Duquette



### GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED

University of Oregon, Eugene  
Florida State University, Tallahassee  
University of Arkansas, Fayetteville

### DEGREES AWARDED:

Doctor of Philosophy, Economics, 2010, University of Oregon  
Master of Science, Economics, 2007, University of Oregon  
Bachelor of Science, Mathematical Sciences, 2002, University of Arkansas  
Bachelor of Science, Electrical Engineering, 2002, University of Arkansas

### AREAS OF SPECIAL INTEREST:

Environmental and Resource Economics  
Behavioral Economics

### PROFESSIONAL EXPERIENCE:

Teaching and Research Assistant, Department of Economics, University of Oregon, Eugene, 2005-2010

### GRANTS, AWARDS, AND HONORS:

Mikesell Award for Best Environmental Economics Paper (Co-winner, 2009)  
Kleinsorge Summer Research Fellowship (2008)

## ACKNOWLEDGEMENTS

I thank Professor Trudy Ann Cameron for a great five years. Trudy's guidance during the research phase of the dissertation was invaluable. Much of my initial professional experiences would not have occurred without her wonderful encouragement and support from the Raymond F. Mikesell Laboratory in Environmental and Resource Economic. I thank Professor William T. Harbaugh for donating the resources of his "Best of Show" brain. His guidance helped identify fundamental questions to the research and he brought humor at all the right times. I thank Professor Ulrich Mayr for his excellent research guidance and patience. I thank Professor Jason Lindo for his remarkable assistance and high marginal value suggestions at the late phases of the dissertation. His insights greatly improved the overall clarity of the research.

Additional assistance, constructive comments, and shared experiences that were helpful towards the completion of the dissertation came from many people. I thank Dan Burghart, J.R. DeShazo, Silke Friedrich, Eric Gauss, Erica Johnson, Yohei Matani, Nino Sitchinava, Peter Stiffler, Brian Vander Naald, and Glen Waddell. I thank participants at the Oregon Resource and Environmental Economics Workshop, the Colorado University Environmental and Resource Economics Workshop, the Western International Economics Association conference (Portland and Vancouver), and the Micro Group Seminar participants at the University of Oregon.

This research has been supported in part by a grant from the National Science Foundation (SES-0551009) to the University of Oregon (PI: Trudy Ann Cameron). It employs original survey data from an earlier project supported by the U.S. Environmental Protection Agency (R829485) and Health Canada (Contract H5431-010041/001/SS) at UCLA (PI: J.R. DeShazo). Additional support has been provided by the Raymond F. Mikesell Foundation at the University of Oregon. Office of Human Subjects Compliance approval filed as protocol #C4-380-07F at the University of Oregon. This work has not been formally reviewed by any of the sponsoring agencies.



For Amanda and Monroe, "We did it!"

## TABLE OF CONTENTS

Chapter	Page
<b>I. SUBJECTIVE CHOICE DIFFICULTY IN STATED-PREFERENCE SURVEYS</b> .....	1
Introduction .....	1
Dimensions of Choice Difficulty .....	5
Objective Measures of Choice Complexity in Utility Space .....	5
Objective Measures of Choice Complexity in Attribute Space .....	9
Observable Individual Characteristics .....	11
Observable Measures of Cognitive Resource Constraints .....	11
The Stated Preference Survey Data .....	13
The Survey Design .....	13
Data Description .....	15
Empirical Models .....	20
Preliminary Estimation of Utility Function Parameters .....	21
Models for Subjective Choice Difficulty .....	21
Results .....	24
Discussion .....	32
Conclusions .....	37
 <b>II. THE NEUROBIOLOGICAL ROLE OF DECISION CONFLICT IN VALUATION</b> .....	 39
Introduction .....	39
Materials and Methods .....	43
Participants .....	43
Experiment Stimuli and Tasks .....	44
fMRI Image Acquisition .....	46

Chapter	Page
Behavioral Model .....	47
fMRI Data Analysis .....	49
Results .....	51
Behavioral Factors That Affect a Decision to Give .....	51
Changes in Neural Activation Due to Difficulty and Charity Importance .....	53
Discussion .....	55
III. EXTREME WEATHER RISKS AND MIGRATION .....	57
Introduction .....	57
Extreme Weather Events and Perceived Risks .....	60
The Perceived Risks of Tornado Activity .....	62
Spatial and Temporal Variability .....	62
In-flow and Out-flow Migration Asymmetry .....	64
Demographic Differences .....	65
Data .....	66
Modeling Migration in Response to Changes in Perceived Tornado Risks .....	71
Count of Migrants .....	72
Historic Risk and Temporal Dynamics .....	74
Spatial Dynamics .....	75
Conditional Distribution of Migration Distances .....	76
Results .....	77
Frequency and Intensity Effects on the Size of Migrant Flows .....	78
Effects by Household Types .....	81
Changes in Distances Traveled and Variation by Subflows .....	84
Spatially Displaced Tornado Activity .....	88
Discussion .....	89

Chapter	Page
APPENDICES .....	98
A. FIGURES AND TABLES .....	98
B. CALCULATION DETAILS OF THE OBJECTIVE COMPLEXITY MEASURES .....	139
C. REVIEW OF THE BASIC “STRUCTURAL” SPECIFICATION FOR UTILITY .....	143
D. THE STANDARD DEVIATION OF FITTED UTILITY AS A DETERMINANT OF SUBJECTIVE CHOICE DIFFICULTY .....	147
E. ADDITIONAL BEHAVIORAL MODEL SPECIFICATIONS .....	152
F. MIGRATION DATA CONSTRUCTION .....	154
G. ADDITIONAL (ABBREVIATED) TABLES .....	162
H. UNABBREVIATED SOCIODEMOGRAPHIC TABLES .....	164
REFERENCES .....	167

## LIST OF FIGURES

Figure	Page
1. Attribute- versus Utility-Space Complexity .....	98
2. Example of a Choice Scenario .....	99
3. Wording of the Follow-up Question Concerning Choice Difficulty .....	100
4. Subjective Choice Difficulty Response Frequencies by Choice Occasion ...	101
5. Pattern of Std. dev. of fitted U by Choice Occasion .....	102
6. Relationship between Entropy and Std. dev. of fitted U .....	103
7. Example Charity Rating Screen .....	104
8. Sequence of Screen Presentations in Mandatory and Voluntary Runs .....	105
9. Average Rate of Acceptance for Payout Conditions .....	106
10. Frequency Distribution of Fitted Choice Difficulty Trials .....	107
11. Neural Activation for Increasing Difficulty .....	108
12. Neural Activation for Increasing Charity Importance .....	109
13. Tornadoes - Conterminous U.S., Single State; Single County .....	110
14. Within- and Adjacent-to-county Tornado Risks .....	111
15. Spatial Variation in Peak Hazard .....	112

## LIST OF TABLES

Table	Page
1. Summary Statistics (n=22176) .....	113
2. Invariant Difficulty Ratings.....	115
3. Simple Preliminary Conditional Logit Models .....	116
4. Utility-Space Determinants of Choice Difficulty.....	117
5. Utility-Space and Attribute-Space Determinants of Choice Difficulty.....	118
6. Additional Determinants of Subjective Choice Difficulty .....	119
7. Comparison of Ad hoc and Structural Specifications .....	122
8. The Set of Charitable Organizations .....	123
9. Estimated Marginal Effects on the Propensity to Give .....	124
10. The Marginal Effects on the Probability of Giving.....	125
11. Peak Voxels for the D+A+ Contrasted Image.....	126
12. Peak Voxels for the C+ Contrasted Image .....	126
13. Dependent Variables: Migration Response Variables, 1992-2005 .....	127
14. Tornado Activity Variables, Presidential Disaster Indicators, and Socioeconomic Characteristics; Summarized for all U.S. Counties, 1992-2005.....	128
15. County-Level Presidential Disaster Declarations, by Incident Type and General Category, 1964-2008 .....	130
16. The Effects of Tornado Occurrences on County-level Migrant Household Tax Return Flows .....	131
17. The Effects of Tornado Intensity on County-level Migrant Household Tax Return Flows .....	132
18. Socioeconomic Response Heterogeneity: Baseline Effects with Socioeconomic Interactions.....	133
19. The Effect of an Additional Tornado by Migration Response Variables and Tornado Prone Regions .....	135

Table	Page
20. The Effect of an F5 Tornado by Migration Response Variables and Tornado Prone Regions.....	136
21. Recency Effects on Contemporaneous Migration Response.....	137
22. The Effects of Spatially Displaced Tornado Occurrences on County-level Migration .....	138

## CHAPTER I

### SUBJECTIVE CHOICE DIFFICULTY IN STATED-PREFERENCE SURVEYS

#### Introduction

The stated preference literature abounds with examples of consumer choice data from valuation surveys that probe individuals' preferences for non-market or pre-test-market goods. To properly estimate preferences and derive willingness to pay (*WTP*) measures, individual-level choice data require empirical models that can handle heterogeneity in preferences, in decision heuristics, and in the context variables that constitute the "ancillary conditions" of the choice environment and contribute to an individual's overall preferred bundle of goods (as in Bernheim and Rangel (2009)). For empirical choice models based in random utility theory, it is unclear exactly how these models should include other aspects of choice that are apart from the formal minimalist representation of a utility function that individuals are typically assumed to maximize.

In this paper, we emphasize the notion of "choice difficulty" as an aspect of the behavioral context for a choice task. If ignored by the researcher, choice difficulty can lead to apparent inconsistencies in the outcomes of utility maximization, even after conditioning on observable determinants of an individual's preferences and the salient features of a good. We elicit subjective choice difficulties and explore their determinants. Our goal is to evaluate the potential role of explicit subjective choice difficulty measures as important adjuncts to choice modeling, especially since they have the potential to



index more comprehensively the variety of choice set features that are more typically employed, in reduced form specifications, as controls for unobserved subjective choice difficulty.

The typical strategy used to investigate context effects in stated preference research has been to assess the influence of different types of survey design elements (such as the number of attributes, alternatives, or choice tasks) on individuals' decision-making behavior. Existing studies commonly refer to these types of objectively measured external influences stemming from the survey design as "choice complexity" and have shown that various dimensions of choice complexity can significantly impact estimates of the marginal utility parameters and thus the resulting calculations of WTP (see Louviere, et al. (2002), Louviere, et al. (2005), and Adamowicz and DeShazo (2006)).<sup>1</sup>

We extend the usual definition of choice complexity to build a broader notion of "choice difficulty." Choice difficulty encompasses interactions between choice set complexity and respondent characteristics, such as sociodemographic traits, idiosyncratic subjective experiences, cognitive capacity, interest in the task, and current attention budgets. Our direct measure of choice difficulty, unique to this survey, comes from a follow-up question that elicits each individual's subjective assessment of the difficulty of the conjoint choice task just completed.<sup>2</sup> We show that our respondents' subjective

---

<sup>1</sup> In a revealed-preference setting, Beshears, et al. (2008) discuss factors that can potentially contribute to decision-making errors and, thus, a disparity between revealed preferences and "normative preferences"—preferences that represent an individual's true interests. They identify five important factors that contribute to the disparity. These include passive choice, complexity, limited personal experience, third-party marketing, and intertemporal choices.

<sup>2</sup> The elicitation of respondents' subjective impressions about their earlier survey responses has also been used in the literature on preference uncertainty e.g. Evans, et al. (2003), Li and Mattsson (1995), Vossler, et al. (2003), and Welsh and Poe (1998). In this literature, researchers incorporate subjective measures of preference *uncertainty* into the estimation process to improve *WTP* estimates that might otherwise be

choice difficulty ratings are correlated with (1) some common objective measures of choice set complexity, (2) observed individual characteristics that may proxy for better abilities (or opportunities) to make consistent choices, and (3) variables based upon evidence from elsewhere in the same survey, for that respondent, that may capture other factors expected to influence the perceived difficulty level for the choice task in question.

Our research also explores an additional candidate measure for choice difficulty that quantifies the distance between alternatives in *utility space*, rather than *attribute space*. With the exception of the entropy measure employed by Swait and Adamowicz (2001a, 2001b), all other common empirical measures of choice complexity are, conceptually, distances between choice-set alternatives in *attribute space*. These attribute-space measures can be problematic in that they can fail to capture the type of choice complexity that arises when alternatives are far apart in attribute-space but nevertheless close in terms of a particular individual's utility function (i.e. when two alternatives are far apart in attribute-space but still lie close to the same indifference curve).<sup>3</sup> In comparison to entropy, our utility-space measure of the similarity of alternatives can reflect the same types of heterogeneity, but may be easier to interpret. We find that, like entropy, our measure is strongly correlated with subjective choice difficulty and its effects on respondents' ratings of choice difficulty are consistent with

---

biased. In our analysis, we recognize that the effects of preference uncertainty and choice difficulty on choice outcomes are likely to be correlated. As our results indicate, cognitive capacity can play a large role in choice difficulty, and this is also likely true for preference uncertainty. However, choice difficulty can arise even when respondents are certain about their preferences.

<sup>3</sup> Utility-space measures also have the potential to be unique for each individual when preferences are allowed to be heterogeneous.

our priors. However, neither our alternative measure, nor entropy, is the *only* systematic determinant of subjective choice difficulty.

In contrast to previous studies, we do not simply embed proxies for choice difficulty directly into our conjoint choice model, using these proxies to shift either the estimated utility parameters or the scale factor (error dispersion) for the choice model. Instead, the attributes of each choice set, in some cases along with the estimated utility parameters from a preliminary choice model, are used to build an array of variables which we use to analyze the determinants of perceived choice difficulty. Among the proxies normally used to control for choice difficulty, we explore which candidates seem to do the best job. Specifically, we investigate the factors which may contribute additional explanatory power. Furthermore, if the subjective difficulty variable adequately captures the various proxy variables which have been used elsewhere in the literature then it might be a good practice to attempt to elicit subjective choice difficulties directly in all stated preference surveys. Our results suggest that this is a strong possibility—the fitted values from an estimated choice difficulty model could function as a single-valued index variable. This index could be used to purge a fitted choice model of systematic variation in marginal utility parameters (or implied *WTP*) due to choice sets which are outliers in terms of choice difficulty, for at least some respondents, without vastly increasing the size of the parameter space of the model.

## **Dimensions of Choice Difficulty**

We entertain a variety of different dimensions of choice complexity which may contribute to a respondent's perception that a choice task is more or less difficult. We consider measures which are calculated in utility space and therefore require preliminary estimation of a preference function before complexity can be quantified. We also consider an array of complexity measures calculated simply in attribute space, which are independent of the preferences of the individual who is making the choice. Finally we will assess the impact on subjective choice difficulty of some available variables that may reflect cognitive capacity or cognitive constraints, plus a variety of individual respondent characteristics and even some variables that describe observed patterns of respondent behavior across the five choice tasks that most respondents completed.

### *Objective Measures of Choice Complexity in Utility Space*

The most typical measures of choice set complexity are based upon calculations in attribute space. The information contained within these measures may encompass the nature of the different alternatives (goods) in the choice set, their number, the level of detail used (or needed) to describe them, and the number of choice sets presented to the individual. However, we suspect that important aspects of choice complexity can also originate from individual *preferences* over the different attributes describing each alternative, in combination with the *levels* of the attributes themselves.

In the quantification of objective choice set complexity, Swait and Adamowicz (2001a, b) propose a measure known as entropy. This measure was first introduced in the

field of *information theory* by Shannon (1948) and can be defined over any set of probabilistic events. Probabilistic events with relatively large degrees of uncertainty will have outcomes that reveal relatively greater amounts of information, or entropy. Swait and Adamowicz (2001a, b) define the entropy of a choice set to be a function of the choice probabilities associated with each alternative in a choice set. Random-utility theory assigns a choice probability to each alternative,  $\pi_{ij} = \pi(x_{ij}, \hat{\beta})$ , that is a function of the latent utility associated with each alternative as measured by the estimated value of the utility index,  $x_{ij}, \hat{\beta}$ , for a conditional logit choice model. Thus, the entropy of a choice set,  $H_i = -\sum_{j=1}^J \pi_{ij} \log(\pi_{ij})$ , can potentially capture choice complexity stemming both from the levels of attributes,  $x_{ij}$ , and preferences over those attributes,  $\beta$ .

Entropy is minimized when there is one dominant alternative in the choice set, and maximized if each alternative is equally likely. The authors model the “scale factor” in a random utility model (the inverse of the error variance) as a quadratic function of entropy and are able to identify systematic effects on choice consistency related to both the linear and the squared terms in entropy. Swait and Adamowicz (2001a) find that the estimated variance (noise) for the utility function is first increasing, and then decreasing, in the level of entropy for a choice set.

The most important thing about the entropy measure is that it incorporates utility-space information, whereas most other existing measures of complexity are restricted simply to attribute space and are assumed to be independent of individual preferences. In the two-alternative case, Figure 1 depicts the relationship between proximity of

alternatives in attribute space and proximity in utility space. These two types of proximity can potentially have independent effects on the difficulty of a choice task. In Figure 1, after Mas-Colell, et al. (1995), the utility of an alternative is represented by the intersection of the utility curve in attribute space with the 45° line. The standard deviation of utility is a measure of the “distance” among alternatives in a one-dimensional utility space, as opposed to a multi-dimensional attribute space. As illustrated, the utility-space distance of a set of alternatives can vary independently from attribute-space distance, for example, as measured by the variability of any single attribute’s levels across alternatives. The steeper (lighter) utility curves represent the possible preferences of a different individual and suggest that preferences can significantly affect the amount of complexity measured in utility space, even for the same set of alternatives. Thus, any complexity that arises in utility-space is strongly dependent on an individual’s preferences over the attributes of the alternatives. For more than two alternatives, of course, a simple distance measure alone is inadequate. To measure the extent to which there is a clear winner in utility space, entropy is a potentially more useful summary.

Entropy is thus one way to very succinctly encapsulate choice set complexity in a way that also reflects individual preferences. Swait and Adamowicz (2001a) explicitly invoke the contribution of this complexity to *choice difficulty* in formulating their hypotheses regarding complexity and variance (p. 158):

“Because complexity is hypothesized to demand additional outlays of effort on the part of consumers to find the utility-maximizing choice, we expect that variance (scale) will be increasing (decreasing) in complexity.”

Since they have no direct measure of choice difficulty, Swait and Adamowicz must simply *assume* that choice difficulty is the unobserved behavioral link between entropy as a convenient one-dimensional summary of choice complexity and the resulting observed heteroscedasticity in choice models. In this paper, we use our direct measure of choice difficulty to assess the extent to which subjective choice difficulty is related to choice set entropy, other candidate measures of complexity, measures of cognitive capacity and/or constraints, sociodemographic/attitudinal variables, and even some systematic patterns in observed choice behaviors.

Along with the entropy measure proposed by Swait and Adamowicz (2001a, b), we consider an alternative and somewhat simpler utility-space measure: the standard deviation of the systematic component of the estimated utility across alternatives in the choice set. In contrast to entropy, this alternative measure is simply a summary of the extent to which estimated utility-levels differ across alternatives for each individual. When the utility-differences across alternatives in the choice set are smaller, an individual's most-preferred alternative will be more difficult to discern. Entropy does a better job of identifying large positive outliers in terms of utility, but it unavoidably subsumes an additional dimension—the number of alternatives in the choice set. This is moot when all choices involve the same number of alternatives, but may be relevant in cases where the sizes of choice sets vary.<sup>4</sup>

---

<sup>4</sup> Yet another measure of choice difficulty in utility space might be the size of the “lead” held by the highest-utility alternative.

*Objective Measures of Choice Complexity in Attribute Space*

Aside from the entropy measure, most researchers have adopted the term “choice complexity” to describe the systematic influence of survey design elements and/or the design of the individual choice sets on response patterns, independent of the characteristics of any particular respondent. DeShazo and Fermo (2002) demonstrate the effects of choice complexity on “choice consistency” via the scale of the error term in a random utility model. They find choice consistency to be systematically affected by the number of alternatives, the number of attributes per alternative, and the number of attributes which are constant across alternatives. In addition, they calculate a measure of the standard deviation of attribute levels within each alternative, and then compute the across-alternative mean of these standard deviations, as well as the across-alternative standard deviation of these within-alternative standard deviations.<sup>5</sup> These additional choice set properties, collectively described as the “information structure” for each choice set, are also shown to have systematic effects on the consistency of responses by individuals. Hensher and his co-authors (see Hensher (2004, 2006a, b)) likewise include a three-level attribute-space measure (wide, narrow, and base) for the range of each attribute level as an objective measure of complexity and find varying evidence for the effects of these measures on choice consistency.

---

<sup>5</sup> Each attribute in the DeShazo and Fermo study is offered at one of just three equally spaced levels (which can be denoted as -1, 0, or +1). This finesses the problem of different units of measurement for each alternative, but it may confound interpretation of the standard deviation across attribute levels because of the inherent scale differences across attributes. This may compromise the estimated effects of the complexity measures (see Lancsar, et al. (2007)). In contrast, but analogously, the attributes we use are cardinal measures so that they have well-defined scales of measurement.



For more general types of attributes, our analysis attempts to improve upon the purely objective measures of information structure used in DeShazo and Fermo (2002). Our attribute levels are cardinal variables. Prior to calculating the standard deviation of the levels of different types of attributes within an alternative, we first standardize the scales of measurement for each attribute (see Appendix B). Standardization prevents the different scales of measurement of the different attributes from acting like weights on their influence. If each alternative is good on all attributes, or undesirable on all attributes, the choice task can be expected to be easier. If, instead, each alternative is good on some dimensions and undesirable on others, respondents will have to make more types of tradeoffs among attributes to identify the most-preferred alternative. To measure these tendencies within a given choice set, we use the choice-set-level variables *Mean SD<sub>i,k</sub>* and *Disp. SD<sub>i,k</sub>* (as developed in Appendix B).

As do Hensher and his co-authors (see Hensher (2004, 2006a, b)) and Johnson (2006), we also employ continuous measures of the standard deviation, across alternatives in the choice set, for the levels of each attribute. We refer to these separate measures as descriptions of the “across-alternative attribute variability.” The impacts of these components of complexity on choice difficulty are theoretically indeterminate. A large standard deviation of an attribute’s level may place cognitive stress on a respondent by forcing him or her to actively consider regions of attribute space that may not be contained in the respondent’s everyday choice set. Likewise, a small range could also be a source of stress if a respondent lacks the cognitive ability to discriminate between small differences in attribute levels. In the limit however, a small enough range could render

the alternatives nearly indistinguishable along that dimension of the attribute in question, reducing the number of attributes remaining to be compared and thus making the choice task easier. Thus quadratic forms or other nonlinear relationships need to be explored.

### *Observable Individual Characteristics*

In this paper, we consider the effects of observable respondent characteristics directly upon subjective choice difficulty. The existing choice literature also considers the effects of these types of variables on the decision making process, but only in a reduced-form sense—by using them as direct shifters of either the marginal utility parameters or the scale of the error term. For example, Hensher, et al. (2005), Hensher (2006a), and Hensher, et al. (2007) study the effects of choice set complexity using an elegant design-of-designs (DoD) approach.<sup>6</sup> Even though these studies, and others, consider income and age, there are likely to be many other individual characteristics that could indirectly affect the choice outcomes of respondents via their effects on subjective choice difficulty. We entertain a wide array of sociodemographic characteristics and other respondent-specific factors as potential covariates for choice difficulty.

### *Observable Measures of Cognitive Resource Constraints*

We consider both educational attainment, and response times for other choice tasks by the same respondent, as objective proxies for cognitive resource constraints that

---

<sup>6</sup> Choice sets with different complexity characteristics are assigned randomly across split samples of respondents. In their mixed logit models, Hensher and his collaborators constrain to zero the marginal utilities associated with the attributes which each individual self-reports to have ignored in making their choices.

potentially co-vary with choice difficulty. Educational attainment, in part, may reflect an individual's ability to make decisions under increasingly difficult choice scenarios. If so, greater educational attainment may lead to lower average subjective ratings of choice difficulty. Further, an individual's capacity to process a difficult choice may affect the time required to make each choice. Thus, longer response times may also be associated with more difficult choices (although they may also belie looser time constraints). If a tighter time constraint results in shorter response times, then choices may be judged to be more difficult.

Several existing studies provide evidence to support the potential usefulness of these "cognitive" measures to explain choice behavior. Haaijer, et al. (2000) and Rose and Black (2006) allow the scale factor (error variance), or the variances of slope coefficients in a random parameters choice model, to depend upon the response times (response latencies) of individuals. Both studies find large improvements in explanatory power over models which do not incorporate information on response times. In a non-economic social science choice context, Fischer, et al. (2000) show that the within-alternative "attribute conflict" that arises due to variation in the attribute levels of an alternative contributes to longer response times and noisier responses in choices among of alternatives.<sup>7</sup>

---

<sup>7</sup> In Fischer, et al. (2000), each respondent in an experimental setting provides a preference rating for a set of twenty alternatives that are presented sequentially and then ratings for an identical set of alternatives, but with a different randomized ordering, after a period of "filler." Therefore, respondents rate each of the assigned alternatives twice. The authors use the difference in rating for an alternative as a measure of response error.

## **The Stated Preference Survey Data**

### *The Survey Design*

Our analysis uses an existing large sample of stated preference survey data concerning preferences with respect to privately supplied programs to reduce health risks. We also take advantage of the random-utility-based theoretical model developed in Cameron and DeShazo (2009) as a basic framework for our analysis.<sup>8</sup> This nationally representative survey includes adults aged 25 years and older in the United States.<sup>9</sup> In brief, the stated preference survey consists of five modules. The first module asks respondents about their subjective risks of contracting the major illnesses or injuries which are the focus of the survey, the extent to which lifestyle changes might reduce their risks of these illnesses, and how taxing it might be to implement these lifestyle changes.

The second module is a tutorial that explains the concept of an “illness profile,” which is a sequence of prospective future health states. An illness profile has attributes that include the number of years before the individual becomes sick (also referred to as the latency of the illness), illness-years while the individual is sick, recovered/remission years after the individual recovers from the illness, and lost life-years if the individual dies earlier than he would have without the disease or injury. Then the tutorial informs

---

<sup>8</sup> For more information on the survey instrument and the data, see the appendices which accompany Cameron and DeShazo (2009): Appendix A – Survey Design & Development, Appendix B – Stated Preference Quality Assurance and Quality Control Checks, Appendix C – Details of the Choice Set Design, Appendix D – The Knowledge Networks Panel and Sample Selection Corrections, Appendix E – Model, Estimation and Alternative Analyses, and Appendix F – Estimating Sample Codebook.

<sup>9</sup> Knowledge Networks, Inc administered an internet survey to a sample of 2,439 of their panelists with a response rate of 79 percent.

the individual that he might be able to purchase a new diagnostic testing program, at a monthly cost, that would reduce his risk of experiencing each illness profile.<sup>10</sup>

The third, and key, module of each survey consists of five different three-alternative conjoint choice experiments where the individual is asked to choose between two possible health-risk reduction programs and a status quo alternative. The survey design is essentially orthogonal, in that each illness program attribute—monthly program cost, risk reduction, the latency of the illness, its duration, and the lost-life years—is randomized across alternatives, choice occasions, and individuals. In addition, a “label” for each illness profile is randomly selected from five specific types of cancer, heart attack, heart disease, stroke, respiratory illness, diabetes, traffic accident or Alzheimer’s disease (with occasional exclusions based on plausibility). Illness profiles need to be unique to each age/gender combination, so simple randomization proved more viable than any attempt at some type of fractional factorial design. One single example of a randomized choice scenario from the survey is presented in Figure 2.

Each choice exercise is immediately followed by a set of debriefing questions designed to help the researcher understand the individual’s reasons for their particular choice. Some debriefing questions depend on the alternative chosen by the respondent—in particular, those who choose the status quo alternative (“Neither Program”) are asked why it is their preferred alternative. Other debriefing questions, including the key “choice difficulty” question for this paper, are asked regardless of which alternative the individual selects. The crucial question for this paper, shown in Figure 3, is “How difficult was your

---

<sup>10</sup> Each illness-related risk-reduction program consists of diagnostic blood tests, drug therapies, and life-style changes, the costs of which would need to be paid annually, since they would not be covered by health insurance.

choice on the previous screen?” Subjects were invited to respond on a Likert-type scale from 1=“easy” to 7=“very difficult.”

The fourth module of the survey contains additional debriefing questions that permit us to explore other potential determinants of the individual’s responses. A final module is collected separately from the same consumer panel and contains the respondent’s socio-demographic characteristics and a detailed medical history, including which major diseases the individual has already faced.

#### *Data Description*

Our data set contains information on 1789 individuals who collectively made choices from a total of 8807 choice sets.<sup>11</sup> With these data, we are not able to study the effects of the number of alternatives and the number of attributes on subjective choice difficulty because these dimensions of the choice scenarios are held constant across all five choice sets posed to each respondent (i.e. all choice sets have three alternatives, and every alternative is described in terms of the same list of attributes). However, this restriction still allows us to study many of the other potential determinants of choice difficulty. Table 1 summarizes the key variables in our analysis.

*How difficult*, our dependent variable, consists of the 1-to-7 subjective difficulty rating by respondents for each choice, with higher-numbered categories conveying greater difficulty. Figure 4 displays the distribution of difficulty ratings by choice occasion. The average difficulty rating for individuals is 2.88. The remaining variables in

---

<sup>11</sup> Of the 8817 choice sets with otherwise sufficiently complete data for analysis, 10 are dropped because each of these choice sets is the sole usable choice set for a respondent.

Table 1 are potential determinants of choice difficulty. We divide these determinants into three broad categories: objective measures of choice set complexity, observable sociodemographic characteristics of the respondent, and proxies for the likely cognitive resources or constraints for each individual respondent.

As in previous studies which consider objective measures of choice set complexity, we employ a number of constructed variables. We consider Swait and Adamowicz's entropy value as one measure of complexity in utility space, but we also consider the simpler standard deviation of the fitted utility indices across alternatives, *Std. dev. of fitted U*, as an alternative. Figure 5 shows the distribution of *Std. dev. of fitted U* across choice occasions, based on the estimated utility parameters from a preliminary choice model that we discuss in the following section. Choice set attributes were random by construction, of course. Thus, the distribution of *Std. dev. of fitted U* should be unchanged across the five different choice occasions even though this measure is preference-dependent.

In addition to the two possibilities for utility-space measures of choice set complexity, we also examine some of the other customary measures within attribute space. Following DeShazo and Fermo (2002), these measures might include the mean and dispersion across alternatives of the standard deviation in attribute levels (e.g. *Mean SD<sub>i,k</sub>* and *Disp. SD<sub>i,k</sub>*), and the standard deviation across alternatives of each program attribute (e.g., *Std. dev. of latency*). Furthermore, we include two sets of measures of across-alternative attribute-level correlations, corresponding to two different

representations for how utility depends on program attributes—one ad hoc, and one more structural.

The “ad hoc” program attributes consist of the unprocessed attribute levels taken directly from the survey’s choice sets. These attributes include the monthly program cost, the size of the risk reduction, the latency of the illness, its duration, and the lost-life years. In our ad hoc specification, these raw attributes enter as linear and additively separable determinants of indirect utility. For our more-“structural” model, the program attributes are first processed to permit a structural random utility model within a formal discounted expected utility framework. Appendix B provides specific details on the construction of the typical DeShazo/Fermo-type objective complexity measures.

We employ a number of observable sociodemographic variables to explain differences in the subjective difficulty of choice tasks. These include age, gender (*Female*), marital status (*Single, Divorced*), race (*Black, Hispanic, and Other ethnicity*, relative to the omitted category *White*), number of household members, number of children in the household, an indicator for single parenthood, income, and an indicator for a dual-income household.

We also have a variety of health history variables for each respondent, as well as other subjectively reported variables. We make use of individuals’ subjective reports about their prior experiences with each class of illness, their subjective risk of suffering a future episode of each class of illness, and their perceptions about the subjective controllability of each type of illness. To accommodate occasional instances of missing health data, we construct an indicator variable,  $1(\textit{Missing health})$ , which has a value of



one (zero otherwise) to identify these individuals.<sup>12</sup> To contain the dimensionality of the parameter space, we use only the individual's mean ratings across the list of illnesses for both the subjective risk measure and the subjective controllability measure.<sup>13</sup> To quantify each individual's personal experience with the major illnesses addressed in the survey, we introduce an individual-level variable which provides a simple count of the number of major illnesses the respondent indicates he or she has already experienced.

Our proxies for cognitive capacity include indicators for the highest level of education attained by the individual (i.e. 1(*Less than h.s.*) and 1(*High school*) (i.e. earned diploma), relative to the omitted category, at least *Some College*). We also include a measure of average time-on-task. To minimize endogeneity, this average is calculated for the respondent's other choice tasks, not including the choice in question (*Avg. duration on other choice occasions*).<sup>14</sup> This is consequently a choice-set-specific variable, since the nature of the "other" choice occasions will vary from choice to choice for an individual. Finally, there are some choice sets in the data for which the response time of individuals is exceedingly long. In many cases, this is probably because the individual took a break in the process of completing the survey. We handle these occurrences with an indicator variable, 1(*Valid duration*), which takes on a value of zero for exceedingly long response times and one otherwise. We interact this variable with the data on choice

---

<sup>12</sup> Information on some or all of the health variables is missing for 166 individuals (or 812 choice sets) because these individuals chose not to respond to some health questions in Module 1 of the survey.

<sup>13</sup> Models where we use the disaggregated subjective responses of each illness type, instead of the mean value for these variables, provide qualitatively similar results.

<sup>14</sup> In empirical results not reported, we find that the use of current choice set response duration as a time-on-task measure to be positive and highly significantly correlated with choice difficulty, suggesting a strong endogeneity between the two measures. Also, we recognize that *Avg. duration on other choice occasions* may not completely mitigate the concerns of endogeneity bias because of the potential for joint dependence across choice occasions.

durations and use only duration data judged most likely to be valid in calculating the average time-on-task variable.

Several features of the raw distribution of the *How difficult* variable, as displayed in Figure 4, merit discussion. This figure highlights some of our concerns about the stability of preferences in a multiple choice-occasion stated-preference environment and suggests the likely need for additional control variables. First, the distribution of subjective ratings appears to be approximately normal, except for a mass of observations associated with the easiest difficulty rating (*How difficult=1*) on each of the five choice occasions. We believe that this heaping at “1” suggests that some proportion of our respondents may devote little attention to the question about their subjective difficulty rating. While they may engage sufficiently with the substantive program choice question, they may also recognize that the difficulty rating question is not as important and automatically choose the left-most option so that they can proceed more quickly through the survey.<sup>15</sup>

Another prominent feature of the distribution in Figure 4 is that respondents, on average, tend to rate their choices as being easier on each subsequent choice occasion. We thus introduce indicators for choice occasions two through five and treat the first choice for each respondent as the baseline throughout our empirical analysis.

To further explore the question of inattentive behavior, Table 2 displays the distribution of subjective difficulty ratings across all 26,451 choice occasions. For each difficulty rating, the table also displays the number and proportion of individuals who use

---

<sup>15</sup> Here, it would have been helpful to randomize the left-to-right order of the difficulty rating, sometimes putting “easy” on the left, and sometimes putting “very difficult” on the left, to check for primacy effects.

the *identical* difficulty rating for all of their choices. A disproportionate share of responses for the “easy” rating—roughly forty percent—are from respondents who maintain the same rating across all five choice occasions. Of course, these individuals cannot express increasing ease of choices because they began at the “easy” end of the bounded scale. The other sixty percent of responses, however, come from respondents who alter their difficulty rating at least once across the sequence of choice occasions.

If inattentive behavior is a consequence of choice difficulty, then the estimated marginal effects of each of the determinants of choice difficulty (such as objective choice set complexity) may suffer from a type of “attention” bias. We control for possible respondent inattention with an indicator variable, *All status quo*, for those individuals who always choose “Neither Program” for their conjoint choices. We also use an indicator, *No change in difficulty rating*, for those respondents who report the identical difficulty rating for all choice occasions. However, Malhotra (2009) finds evidence that inattention is more likely in the case of simple tasks (“survey satisficing”), and that people are “more motivated to persist in completing tasks [which are] intricate, challenging, and enriching.” Thus we cannot automatically assume that increased difficulty leads to less attention to a choice problem.

### **Empirical Models**

Before we can explore our models to explain subjective choice difficulty, we need to estimate some approximate utility parameters from a preliminary conditional logit choice model. These are needed so that we can build the “fitted” utilities required to

construct the utility-space choice complexity measures— *Entropy* and *Std. dev. in fitted U*. These key measures, along with our other potential determinants of difficulty, are then used in the main model to explain respondents’ subjective difficulty ratings for each choice task.<sup>16</sup>

#### *Preliminary Estimation of Utility Function Parameters*

We consider two different specifications for the preliminary utility model from which we construct our utility-space measures of complexity. The ad hoc model has been outlined above. It merely uses the main raw attributes of each choice scenario to build a linear and additively separable utility “index” for construction of the utility-space measures. Our structural model, which borrows heavily from previous research with this same survey sample, allows us to construct utility-space measures based on a discounted expected utility specification. In Appendix B, we review in some detail the construction of the structural program attributes.

#### *Models for Subjective Choice Difficulty*

For individual  $i$  on choice occasion  $t$ , we model the subjective choice difficulty (*How difficult*,  $d_{it}$ ) using a seven-interval ordered probit specification. Our goal is to assess the extent to which respondents’ subjective choice difficulties are affected by objective measures of choice set complexity, by observable individual characteristics, and

---

<sup>16</sup> Work in progress includes the rather daunting task of developing a joint model that simultaneously uses respondents’ reported choice difficulty ratings to shift the estimated preference parameters (and/or the scale factor) in our choice models. Here, we concentrate specifically on the determinants of perceived choice difficulty.

by apparent cognitive constraints,. We allow the latent continuous subjective choice difficulty,  $d_{it}^*$ , to be a linear-in-parameters function of several types of determinants:

$$d_{it}^* = \delta' w_{it} + \beta' x_i + \varepsilon_{it}, \quad (1.1)$$

where  $i = 1, \dots, N$  respondents and  $t = 1, \dots, T$  choice occasions per respondent. The vector  $w_{it}$  contains several objective measures of choice set complexity. The vector  $x_i$  captures a number of observable sociodemographic characteristics and proxies for cognitive capacity, which are assumed to be invariant over choice occasions for the same respondent. The error term,  $\varepsilon_{it}$ , is both individual- and choice-occasion specific. To identify the parameter vectors  $\delta$  and  $\beta$  for the observable determinants of subjective choice difficulty, we assume that  $\varepsilon_{it}$  is distributed  $N(0, \sigma_\varepsilon^2)$  and that  $\varepsilon_{it}$  is uncorrelated with  $w_{it}$  and  $x_i$  for all individuals and all choice occasions.

The relationship between the observable ordered categorical response (represented by the individual's subjective difficulty rating,  $d_{it}$ ) and the continuous latent difficulty variable is:

$$d_{it} = j \text{ if } \mu_{j-1} < d_{it}^* < \mu_j \quad (1.2)$$

where  $j = 1, \dots, 7$ ,  $\mu_0 = -\infty$ ,  $\mu_7 = \infty$  and the other cut points  $\mu_1, \dots, \mu_6$  are estimated thresholds from the ordered probit regression analysis. Under the assumption that the error term is normally distributed, the probability of observing response  $d_{it} = j$ , conditional on  $w_{it}$  and  $x_i$ , is:

$$P_{ij} = P(d_{it} = j | w_{it}, x_i) = \Phi(\mu_j - \delta' w_{it} - \beta' x_i) - \Phi(\mu_{j-1} - \delta' w_{it} - \beta' x_i), \quad (1.3)$$

where  $\Phi(\cdot)$  is the standard normal cumulative density function.<sup>17</sup>

As usual for discrete outcome models, neither the location nor the scale of the latent difficulty variable is known. Parameter identification requires the normalization to zero of either the intercept term in the latent regression specification, or one of the cut points (i.e. the thresholds between ratings). We make the assumption that the average intercept across individuals and choice sets is zero. Furthermore, to accommodate the arbitrariness of scale, the estimated coefficients, the cut points, and the error term are all normalized on the error standard deviation, e.g.  $\delta = (\delta^* / \sigma_\varepsilon)$ , and the error variance for the normalized model is thus equal to one.

For completeness, we also consider a fixed-effects specification. In this context, we will assume that subjective choice difficulty, as measured by the difficulty ratings, is a cardinal variable. In a least-squares framework, we introduce non-zero individual fixed effects,  $\alpha_i$ , to parameterize the model for unobserved heterogeneity that is specific to each individual.

$$d_{it}^* = \delta' w_{it} + \alpha_i + \varepsilon_{it} \quad (1.4)$$

The fixed effects,  $\alpha_i$ , will subsume all factors which are invariant over choice occasions for the same individual, so the  $\beta$  coefficients on the  $x_i$  variables in equation (1.1) cannot be separately identified. While the cardinality assumption is a compromise, this is one way to determine whether any unobserved heterogeneity might be biasing the estimated

---

<sup>17</sup> Note that the ordered probit model is readily able to accommodate the heaping of responses that we observe at the easiest difficulty rating.

slopes on the remaining  $w_{it}$  variables (specifically, those variables which are not randomly assigned).

## Results

Table 3 shows results for our two different specifications of the preliminary conditional logit models. These models produce estimates of the preference parameters for each attribute in a random-utility framework. Model 1 is our ad hoc specification of the latent utility index and includes only the raw attributes. In contrast, the structural model uses the constructed attributes described in Appendix B. This specification is reported as Model 2. Since the maximized log-likelihood for ad hoc Model 1 is larger than that for structural Model 2, we focus on choice set complexity measures in utility space and attribute space for the ad hoc model in our main illustration. However, we discuss a sensitivity analysis across the two possible models later in the paper.

Table 4 begins our estimation results for progression of variations on the ordered probit specifications in equation (2). Models 1 and 2, our most parsimonious specifications, provide striking evidence of the possible effects of choice complexity in utility space on subjective choice difficulty. Model 1 treats the *Std. dev. of fitted U* as the only determinant of respondents' subjective choice difficulty. Model 2 employs the calculated *Entropy* of the choice set as the only determinant. The results show that, independently, each utility-space measure is a strongly statistically significant

determinant of the difficulty ratings selected by respondents ( $p < 1\%$ ). The maximized value of the log-likelihood is also essentially identical across the two specifications.<sup>18</sup>

When we include *both* the *Std. dev. of fitted U* and *Entropy*, as in Model 3, their individual effects are insignificant ( $p > 10\%$ ). To illustrate why this occurs, Figure 6 reveals that there exists a very close, although somewhat non-linear, relationship between these two utility-space measures. Thus, we are unable to distinguish between their separate effects when both variables are included in one model, and specifications like Model 3 are unhelpful.<sup>19</sup>

Table 5 preserves the utility-space *Entropy* variable from Table 4, but introduces two other attribute-space measures of objective choice set complexity as further explanatory variables for subjective choice difficulty. In section (b)(1) of this table, Model 2a reveals that the mean across alternatives of the within-alternative standard deviation of (standardized) attribute levels has a statistically significant negative effect on perceived difficulty.<sup>20</sup> A low mean value for these measures means that alternatives tend to have levels of attributes that are either all good, all bad, or all neutral, rather than mixes of attributes with some good and some bad, necessitating more tradeoffs during the decision process. We probably expect choices to be easier when the *Mean std. dev.* is small, and harder when more tradeoffs must be considered, which appears not to be the

---

<sup>18</sup> In results not shown, we extend the linear specification of these variables to a quadratic form. The linear component for *Std. dev. of fitted U* is unchanged in regards to magnitude, sign, and significance, but the additional quadratic term is insignificant. Both linear and quadratic terms are insignificant when the specification of subjective difficulty is quadratic in *Entropy*.

<sup>19</sup> We perform nested likelihood-ratio tests of the restrictions present in Models 1 and 2 against the unrestricted model of Model 3. Confirming the Wald-type test embodied in the individual asymptotic t-test statistics on each parameter, these tests fail to reject the hypothesis ( $p > 10\%$ ) of a zero incremental contribution for either variable when the other is already present in the model.

<sup>20</sup> See Appendix B.2 for a detailed exposition of how this objective complexity variable is calculated.



case. Model 2b, on the other hand, suggests that *Disp. of std. dev.*, the dispersion, across alternatives, of these same standard deviations (if added on its own) has a positive and statistically significant effect on perceived difficulty. In this case, some alternatives would have all good, all bad, or all neutral attribute levels, while others would have mixes of good and bad attribute levels. Choices appear to be judged more difficult when this is the case.

In Model 2c in Table 5, adding both the mean and dispersion of these standard deviations to this same model leaves only the mean term statistically significantly different from zero, so some of the information in these two measures appears to be duplicative. Model 2d, however, reveals that perceived difficulty is not linear in the *Mean std. dev.* characteristic of a choice set. This variable enters quadratically with a negative coefficient on the squared term. As the *Mean std. dev.* of within-alternative attribute levels increases from its minimum of 0.82 to its maximum of 1.31, perceived choice difficulty first increases, is maximized at a value of 1.06 for this variable, then decreases. The negative effect thus dominates if only a linear term is used, as revealed in Model 2a.

However, the models in Table 5 neglect other factors which may help to explain the variation in subjective difficulty ratings across individuals and across choices. If these other determinants are correlated with the *Entropy* variable (or with the *Std. dev. of fitted U* variable), then its coefficient may be biased. We check for this possibility, using just the *Entropy* variable as a utility-space measure, in the additional models presented in Table 6.

Section (b)(1) of Table 6 includes controls for both the *Mean std. dev.* and the square of *Mean std. dev.* as suggested by the results in Table 5. Section (b)(2) of Table 6 then summarizes the effects on perceived difficulty of standard deviations in attribute levels across alternatives on an attribute-by-attribute basis. Attribute levels are randomly assigned, except for occasional implausibility exclusions, so we expect no multicollinearity in these standard deviation measures. Only the standard deviation across alternatives in the number of sick-years appears to have a positive and significant effect on subjective choice difficulty. These results conflict with our basic intuition that greater dissimilarity in an attribute should make alternatives easier, rather than harder, to compare. All other individual attribute-space standard deviation measures in the ad hoc specification bear coefficients which are statistically insignificant. These results, however, are for the case where we already control for the factors listed in Section (b)(1) of the table.

We note that the estimated effect of the *Entropy* variable changes only slightly between Models 1 and 4 despite the addition of the seven additional attribute-space measures. Overall, these results suggest that respondents' perceptions of choice difficulty are sensitive to the proximity of the alternatives in utility space as well as to the mix of objective attributes within a choice set (in ways that may be independent of preferences over these attributes).

Due to the essentially randomized design of the illness profiles, our measures of objective choice set complexity are orthogonal to all of the sociodemographic variables. Thus our models can in principle be estimated without controls for sociodemographic

characteristics, without concerns about omitted variables bias in the coefficients of any of the purely attribute-space variables. However, we extend the specification in Model 5—to include indicators for choice occasions 2 through 5 as well as a range of sociodemographic variables—to see whether these variables further increase our ability to predict subjective difficulty ratings. Model 5 also includes our observable proxies for cognitive capacity, health history and subjective health variables, and some other controls that may capture inattention to the choice task.

In Model 5, we find that the coefficient on the *Entropy* becomes about one-third larger when we control for choice occasions and a wide range of respondent-specific characteristics. In section (b)(1) of the table, the coefficients on the quadratic form of *Mean std. dev.* change only slightly and maintain their significance. Section (b)(2) of the table reveals that respondents seem to view choices as *more* difficult if costs are more different across alternatives but with the standard deviation in years sick now becomes statistically insignificant. This asymmetry between the cost variable and the other illness profile attributes (such as years sick) may not be surprising, however.

Model 5 also allows us to identify the effects of important observable proxies for cognitive capacity and some observable sociodemographic characteristics of individuals. Among the cognitive capacity measures, we find a very clear gradient in the effect of education on subjective difficulty ratings. There is no significant difference in subjective difficulty between those individuals who completed college and those with only some college experience so we combine these as the omitted category. However, there exists a significant increase in rated choice difficulty for respondents who have only a high

school degree compared to the baseline individuals who have at least some college education. This effect is even larger when the comparison involves individuals with less than a high school education.

Our proxy for other aspects of cognitive capacity (or constraints on its utilization), *Avg. duration on other choice occasions*, indicates that the net effects of these unobserved determinants, collectively, have a positive effect on rated choice difficulty. People who spent more time on other choice tasks tend to rate the current choice as more difficult.

Sociodemographic characteristics also influence subjective choice difficulty. Perceived choice difficulty seems to increase with income (which may actually measure the opportunity costs of time spent on these choice tasks) and to decline with the respondent's age (which may reflect either greater confidence about decision-making ability or more familiarity with health-risk related choices after controlling for educational attainment).<sup>21</sup> Perceived difficulty is also higher for females, lower for blacks but higher for Hispanics (relative to whites), and lower for households with more children. Membership in a dual-income household may correspond to lower perceived difficulty, although the mechanism for such an outcome is not clear.<sup>22</sup> Furthermore, subjective choice difficulty decreases with the number of illnesses with which a

---

<sup>21</sup> We thought we might possibly identify an increase in choice difficulty for some of the oldest seniors, but perhaps selection bias among these oldest seniors means there are too few seniors in our sample who are old enough to be cognitively compromised to a statistically detectible extent. The point estimate of the coefficient on the square of age is positive, suggesting a U-shaped profile for perceived difficulty as a function of age, but the coefficient is not quite statistically significant at the 10% level.

<sup>22</sup> To the extent that choices are easy if the respondents simply checks "Neither Program" in every case, we have been careful to control for cases with this universal rejection of the offered programs. However, individuals who selected "Neither Program" in most cases, but not all, are not captured by this variable.

respondent has had prior personal experience. In contrast, larger values for the average subjective risk of future major illness (which tends to increase with age) and the average subjective controllability of illness correspond to greater subjective difficulty for choice tasks. These controls may offset some of the tendency of greater age to affect subjective choice difficulty.

Respondents are also more likely to rate a choice as being easier if they choose the status quo (“Neither Program”) option for all the conjoint choices or report a constant subjective difficulty rating for all of the choice sets they considered. This last result supports with our conjecture that these particular respondents may have been relatively inattentive to the various choice tasks.

An obvious potential concern about Model 5 in Table 6, given that we have panel data in the form five choices for most individuals, is the possibility of bias from remaining unobserved heterogeneity. While the design of the offered attributes is randomized, so that we expect minimal correlation between these choice set design variables and any unobservable respondent characteristics, it is possible that some of the observed respondent heterogeneity is correlated with unobserved heterogeneity. So we next explore a fixed-effects specification.

Model 6 in Table 6 involves a least-squares-based fixed-effects model to condition on the full set of choice-invariant characteristics, whether observed or unobserved, to better identify the effects of the choice-set varying factors.<sup>23</sup> Importantly,

---

<sup>23</sup> We have also carried out the fixed-effects analysis with an unconditional fixed-effects ordered probit model that is available in the LIMDEP 9.0 software. The unconditional fixed-effects model is subject to the incidental parameters problem when the number of time periods (or choice sets) is small (see Greene (2008)). We did not find any qualitative changes in the relative sizes of coefficient estimates when moving

we find that the effect on subjective difficulty ratings of a change in the *Entropy* variable remains significant and negative. This particular measure of choice complexity in utility-space plays an important role in the conjoint choice responses of individuals—but not the *only* role.

We find in Section (b)(1) of Table 6 that the coefficients for the quadratic effect of the within-alternative attribute-space choice complexity measure, *Mean std. dev.*, retain their relative magnitudes. In section (b)(2), the point estimates of the effects of the individual standard deviations of annualized costs and the illness profile attributes are relatively robust across Models 5 and 6 (with the exception of a sign change for the statistically insignificant effects of variability in the risk difference attribute).

Concerning the choice occasion indicators in Section (b)(3) of Table 6, Model 5 suggests that respondents rate choices as becoming successively easier, on average, with each additional choice. A test of equality of the estimated coefficients for the choice occasion indicator variables reveals significant differences across choice occasions. In Model 6, however, relative to the first choice, we fail to reject equality among the ratings differentials for choice tasks 2 through 4. However, we can reject equality of the ratings differentials across choice tasks 2 through 5.<sup>24</sup> Thus, respondents (on average) report

---

to this model, although we do find an upward shift (in absolute terms) in *all* the coefficients of the model, which is to be expected given our relatively small number of choice sets (i.e. five per respondent). Instead, we choose to report results for the simpler linear fixed-effects models under the assumption the individuals' apparent subjective difficulty ratings are approximately cardinal.

<sup>24</sup> Respondents were informed prior to their last choice occasion that the following choice would be their final one. This may explain the difference in the average difficulty rating of the last choice from the average ratings of the intervening choice occasions. Also, in an alternative specification, we incorporate choice occasion effects by using a linear index for the number of choice occasions. Under this alternative treatment, we include a linear and quadratic term for the index and find only the linear term (negatively) significant.

lower difficulty ratings across choice tasks but rated difficulty does not appear to decline in a linear fashion.

Finally, the sole striking difference in sign across Models 5 and 6 is for the coefficient on the average duration on other choice tasks for a respondent. In Model 5, choices were perceived as more difficult when a respondent tended to spend more time on other choices. This suggests that longer choice durations may reflect *lesser cognitive capacity*. Net of any unobserved heterogeneity, however, Model 6 suggests that longer choice durations correspond to judgments of *lesser choice difficulty*. This would be consistent with longer durations reflecting *lesser time constraints* on the decision-making process.

## Discussion

Our results suggest that there are a number of important factors which influence the subjective difficulty of a choice task. We find that utility-distances between alternatives in the choice set clearly have a significant effect on the choice difficulty that respondents perceive. The *Entropy* variable and our alternative utility-space measure of these distances, *Std. dev. of fitted U*, appear to perform about equally well in capturing the choice difficulty that stems from the closeness of alternatives in utility-space.<sup>25</sup>

Our *Entropy* measure could, of course, be rendered more sensitive to differences in preferences across individuals if the preliminary conditional logit choice model involved greater parameter heterogeneity. We have stayed with the simplest possible

---

<sup>25</sup> In general, for data in which the number of alternatives is constant throughout the survey, any effects on choice difficulty from entropy can safely be attributed to the distances between alternatives in utility-space.

specifications in this model because some of our attribute-space variables (notably those under the heading of *Across-alt. attrib. variability*) are calculated on an attribute-by-attribute basis. However, there is no requirement that the list of attributes, or the functional form, in the preliminary choice model used to create the *Entropy* variable must be the same as the list of attributes introduced as regressors in the choice difficulty model. *Entropy* may, in fact, explain choice difficulty even better if the preliminary choice model is richer.<sup>26</sup>

Overall, our analysis of the effects of choice context suggest that the determinants of perceived choice difficulty likely extend well beyond just the simple proxies based on measures of objective choice complexity which have typically been explored in the existing literature. In addition to our alternative measures of proximity in utility space, subjective choice difficulty appears to vary systematically with a variety of dimensions of individual heterogeneity (e.g. income, age, ethnicity, and the number of children present in the household).

Explicit information about subjective choice difficulty could be incorporated into a richer (and much more complicated) joint empirical model. Importantly, our results suggest that the empirical estimation of demand or *WTP* may be affected by a rather wide variety of factors that have not typically been accounted for in the choice complexity literature. In this paper, we have used a very crude preliminary discrete choice model merely to produce the initial estimates of the utility parameters needed to build any measure of alternative similarity in utility space. In principle, this sub-model could be

---

<sup>26</sup> Models which assess this possibility are currently being explored. Appendix D provides alternative tables for similar models where the *Std. dev. of fitted U.* is used instead of *Entropy* as the key utility-space measure.



estimated simultaneously with another sub-model to explain subjective choice difficulty. Actual or fitted choice difficulty could be used simultaneously to shift the utility parameters and/or the error variances in the choice model. This would allow for a much broader analysis of the direct effects of choice difficulty on *WTP*. Having focused here on the details of the sub-model for the subjective difficulty measure, however, we leave this more comprehensive analysis for subsequent work.<sup>27</sup>

It is also possible that the results based on the ad hoc specification for the choice model may not carry over to specifications. To assess this possibility, we use the structural attributes of the theoretical model that we outline in Appendix C, which is a simplification of the model employed in Cameron and DeShazo (2009). Table 7 reproduces the main coefficient estimates for the ad hoc specification (Model 6 in Table 6) along with the estimates for the corresponding structural specification (Model 7). Both models are estimated using the linear fixed-effects estimator as an approximation. The comparable attribute-space measures bear very similar coefficients, but the use of the standard deviations of the structural variables, instead of the ad hoc variables, causes the coefficient on the key *Entropy* variable to fall by half.

Another consequence of controlling for individual fixed effects in Model 6 is that the effect of *Avg. duration on other choice tasks* changes from significant and positive, to significant and negative. This generally implies that across individuals, subjective

---

<sup>27</sup> We have research already in progress concerning this joint model. It is straightforward to specify such a model. However, because the estimated utility parameters show up in more than one place, convergence is difficult to achieve in a full information maximum likelihood context. We have had success with a model that alternates between (1) a logit model involving parameters and/or the error variance expressed as functions of the fitted values from the previous iteration of the difficulty model, and (2) an ordered logit subjective difficulty model conditional on fitted logit parameters from the previous iteration of the choice model. Iterating between the two conditional models permits convergence.

perceptions of difficulty are positively correlated with average duration on the choice tasks. This generally implies that unobserved heterogeneity across individuals is likely correlated with average processing times for these types of choices. A binding time constraint may be one such omitted variable. We cannot control directly for how binding respondents' time constraints might have been. Longer observed response times might correspond to an opportunity for a more leisurely consideration of the alternatives, which might result in a perception of less difficult choices.

Within each subject's choice tasks, however, the subjective rating of difficulty and the duration on the choice tasks are inversely related. Fischer, et al. (2000) note that while choice set complexity is likely to lead to longer response times, the observed pattern will be confounded if individuals endogenously adopt decision strategies in response to the level of complexity in a fashion similar to those types of behaviors modeled in the effort-accuracy literature (e.g., Payne (1993)). If decision strategies are flexible, then response times could decrease for a given level of difficulty, leaving the general relationship between choice difficulty and response times ambiguous. Every subject in our study was presented with only five choice occasions, which prevents us from effectively exploring, in any depth, some of the potentially confounding effects of evolving endogenous decision strategies on the relationship between response times and choice difficulty. However, we control crudely for these possible changes in the average individuals' response strategy with the choice occasion indicators in our empirical analysis. In general, this ambiguity suggests that future conjoint choice survey research

may benefit if respondents were also asked about the extent to which they had to rush to make their choices.

We also hypothesize in our study that stated subjective choice difficulty potentially captures all of the different things that can conspire to make a particular stated preference choice situation “difficult” from the perspective of the individual respondent. Furthermore, subjective choice difficulty may differ across respondents even for identical choice tasks.<sup>28</sup> However, a potential concern with the use of subjective assessment of choice difficulty is that respondents may lack the experience necessary to properly locate the difficulty of the initial choice on any absolute scale, which may distort coefficient estimates for any of the factors of choice difficulty associated with the context and/or design of the survey.

Given that respondents have insufficient knowledge of the likely distribution of subjective choice difficulty for the *first* choice occasion, each respondent may select a difficulty rating for the first choice in a relatively arbitrary fashion. As respondents proceed through additional choice occasions, however, they begin to update their beliefs about the distribution of difficulty levels across choice occasions. In a survey containing a large enough number of choice occasions, the influence of the initial rating—affected by the respondents’ prior belief about the distribution of possible choice difficulties and his or her guess about where the first choice may lie on the overall difficulty spectrum—

---

<sup>28</sup> In a similar quest to our analysis, Luce, et al. (2003) extends the efforts by Fischer, et al. (2000) and allows half of the subjects the opportunity to put 90% confidence bounds on their initial ratings. The authors use these confidence bounds as subjective measures of the level of conflict that individuals consciously or unconsciously perceive. The response errors and confidence bounds are both shown to be affected by variation in attribute conflict, attribute extremity, and choice context.

might eventually disappear as respondents gain experience. One could omit the first choice tasks and their ratings from the analysis, treating them as part of a “burn-in” phase. However, our survey involves only five choice occasions per respondent and this limits our ability to fully address the possibility of these initial reference level effects. It may be possible to address this concern by normalizing on respondents’ initial choice ratings, although the boundedness of the seven-point scale is a limitation.

## **Conclusions**

Previous studies have not enjoyed the advantage of a directly elicited subjective difficulty rating for each one of a large set of stated choice tasks, with multiple choices per respondent, such as the atypical variable we exploit in this paper. As a result, existing studies have typically relied upon on only some of the many possible proxies for choice difficulty. “Choice difficulty” is often invoked as the latent factor which explains why some of these proxies have the systematic effects on marginal utilities or scale factors that they are observed to produce. However, it has only been possible to speculate that “choice difficulty” is the relevant missing link (i.e. unobserved mediating variable).

Without a specific choice difficulty variable, researchers who wish to control for choice difficulty need to be satisfied with controlling for it *indirectly* instead, using one or more raw or constructed quantities based upon observable variables. Each of these variables may be able to explain *some* of the variation in the missing choice difficulty variable, but none does it all. We have demonstrated this handicap by showing that

several different classes of variables seem to be predictive of individuals' reported subjective choice difficulty ratings.

Our findings suggest that directly elicited subjective choice difficulty ratings may have the potential to serve as a sound univariate summary of these numerous determinants of choice difficulty. Thus, future stated and revealed-preference research may be able to circumvent the adoption of some of the more sophisticated empirical choice models (e.g., Louviere (2001), Swait and Adamowicz (2001b), Hensher (2004), Greene and Hensher (2007)) to account for factors in the choice environment that can bias parameter estimates. In particular, our relatively non-intrusive follow-up question about the difficulty of the preceding choice may reduce the need for a highly parameterized empirical choice model with many kinds of objective proxies for contextual determinants of choice difficulty. In future analyses, a direct measure of subjective choice difficulty may be a viable way to control for some or all of the potential effects on respondent behavior originating from the challenges of the choice environment, in general, or for different types of individuals.

## CHAPTER II

### THE NEUROBIOLOGICAL ROLE OF DECISION CONFLICT IN VALUATION

#### Introduction

Economic policies for the provision of public goods are necessary because market mechanisms can fail to provide adequate levels of the goods. The optimal design of policies requires that economists be able to measure the value that individuals have for these goods. To do this, economists typically infer the values of goods based upon the observed choices by individuals. The standard model of rational choice has individuals equate marginal utilities when making choices between alternatives, but this may not be how people make decisions in all situations. For instance, individuals potentially adopt some type of choice heuristic under difficult decisions. The underlying behavioral process that might lead to the adoption of some alternative choice mechanism is not well understood by economists. In this paper, we use neuroeconomic analysis to explore the effects that decision conflict may have on the valuation process of individuals and on the measured value of public goods that economists obtain. We hypothesize that there may be areas in the brain that dually encode decision conflict and valuation assessment which would then suggest an intermediate step to the decision outcome. This research provides the opportunity to improve existing models for economic behavior through a better understand of mediating choice mechanisms that are distinct from valuation and the encoding of decision conflict.

Previous research identifies specific regions associated with certain types of valuation processes (e.g., Camerer, et al. (2005)). For instance, the *goal value* that individuals compute for the expected reward of a good (e.g., willingness-to-pay and marginal utility) is implicated in the orbital prefrontal cortex (OPFC) and the anterior cingulate cortex (ACC) (Plassmann, et al. (2007); Hare, et al. (2008); Smith, et al. (2010), FitzGerald, et al. (2009)). Other studies identify the ventromedial prefrontal cortex (vmPFC) (Chib, et al. (2009); Hare, et al. (2009)), ventral striatum (Crosson, et al. (2009)), and the dorsal striatum (Pine, et al. (2009)) as having strong associations with the goal valuation process implicit in economic decision making.<sup>29</sup> A value of a different sort, *decision value*, is the computed net benefits of the different “goals” for a particular outcome. Hare, et al. (2008) find neural activation in the central orbital prefrontal cortex to be associated with decision values. Similarly, Smith, et al. (2010) identify experience value, or the value derived from the consumption of a good, as being correlated with neural activation in specific regions of the brain. There is evidence that the neural mechanisms associated with valuation also vary with the type of good and the choice scenario involved (Chib, et al. (2009)).

Our analysis examines the potential for individuals’ cognitive capacity to affect the valuation process and subsequent choice outcomes when making comparisons between goods. Under the theory of bounded rationality, individuals’ cognitive resource limitations can lead to choices that would be suboptimal in comparison to the choices made under the traditional rational choice framework in economics. Essentially, if

---

<sup>29</sup> Rushworth and Behrens (2008) note other areas in the brain implicated in economic decision-making and provide a summary of current findings.

thought is costly, the role that cognitive resource limitations may have in the valuation and choice processes is unclear. For instance, cognitive resource limitations are likely to have a role in the outcome of choice tasks that involve a high degree of critical thinking or that require the ability to process certain kinds of information. These limitations can potentially cause internal conflict and stress for an individual, and may result in an individual choosing to adopt some simplifying heuristic as the mechanism of choice.

There are a number of psychology studies that investigate decision conflict for simple recognition and information retrieval tasks. These studies find evidence that the neural encoding of decision conflict may take place in the ACC. However, Pochon, et al. (2008) find that the neural encoding of decision conflict in the ACC also appears to accompany higher-level decision tasks, such as economic decision-making. Pine, et al. (2009) find that the estimated utility difference in the two-alternative intertemporal money decision is correlated with activity in the ACC. Botvinick, et al. (2004) and Carter and Van Veen (2007) provide recent evidence about the role of the ACC in cognitive control in resolving competing simultaneous representations. Other areas found to encode decision conflict include the medial frontal gyrus, the anterior insula, the ventral striatum, and the dorsomedial thalamus (Grinband, et al. (2006)).

Individuals may find that decision conflict increases when the alternatives in a choice are relatively similar in terms of their assessed marginal utilities. When individuals are relatively indifferent between the alternatives in the choice, they may find the choice to be more difficult because of the increased mental processing required to differentiate the most preferred. We look for both behavioral and neuroeconomic



evidence of this within our experimental design. Our goal is to understand the neurobiological link that the degree of difficulty of a choice, or the relative indifference among the alternatives in the choice, may have with individuals' decision-processes and their subsequent choice outcomes. In particular, we hypothesize that there may be regions in the brain which dually encode valuation and decision conflict. One hypothesis is that if a choice is deemed relatively too difficult then the neural activation in areas related to valuation may diminish as the importance of making the valuation assessment decreases. As a second hypothesis, increasing difficulty may lead an individual to adopt of some set of choice heuristics, which then provides an individual a "clearer" goal value and increases the activation in some or all of those regions related to valuation. In essence, decision conflict may behave as a contextual factor in the valuation of goods.

In addition to looking for regions in the brain that dually encode valuation and decision conflict, we explore the possibility that age of a participant may modulate these neural linkages. The neural encoding of choice difficulty may be most easily performed by participants who are less constrained in their cognitive resource capacity, which is a neurobiological feature likely to vary with age. Recent studies supports the theory that as certain systems of the brain evolve through an individual's lifespan, so too will there economic behavior. (See Mohr, et al. (2010) for a review.) Thus, we hypothesize that age effects might play a significant role in our study given the potential for cognitive resource limitations to vary with age and affect the resolution of conflict.

Similar to Harbaugh, et al. (2007) and Moll, et al. (2006), the decisions that participants make in our experiment allow us to investigate motives for charitable

donations. We obtain subjective ratings of charitable organizations and use these ratings as subjective goal values in our analysis. We then examine the regions of the brain that are associated with valuation and choice during a task in which participants can choose to give money to the charities or not. Prior evidence suggests that there are several regions of the brain in which we might expect to find such activation.

Our efforts most closely follow those of Pine, et al. (2009) and FitzGerald, et al. (2009). Both of these teams of researchers perform separate analyses for value assessment and decision conflict. For instance, Pine, et al. (2009) investigate intertemporal money decisions and find that the neural encoding of marginal utility is independent from the encoding of the magnitude of the money reward. Marginal utility estimates are obtained from a behavioral model of the choices that individuals make during an experiment. The behavioral model allows for individual estimates of the degree of concavity in the utility function (i.e., a risk aversion parameter) and an intertemporal discounting parameter. Our investigation improves upon statistical fMRI analyses of these existing papers by simultaneously modeling the effects of these two components during the choice process.

## **Materials and Methods**

### *Participants*

Twenty-five males (ages 18-61, mean 42) and twenty-five females (ages 25-66, mean 46) participated and completed the study. Data for six participants are excluded from our statistical analysis either because of compromised fMRI brain imaging data

(i.e., too much head motion) or because of missing behavioral data for the statistical models we employ. We use the data for the remaining 44 participants in our analyses. Payments to individuals took two forms: 1) a guaranteed \$15 per hour while completing the tasks in the experiments and 2) an endowment of \$100 less any actual donations that they are required (or choose) to make to real charities. (We obtained written consent from all participants prior to performing the experiment as required by the The Institutional Review Board at the University of Oregon.)

### *Experiment Stimuli and Tasks*

Prior to the fMRI scanning tasks in the experiment, participants completed a survey that assessed their personality, empathy, endorsement of prosocial norms, volunteering activity, and their financial situation, including their actual charitable giving to different types of charitable organizations. In addition, participants completed a survey in which they rated twenty-four different charities for the subjective importance of the charity's cause and the potential helpfulness of the charity to that cause. Participants also indicate (yes/no) for each of the charities in the survey if they have any friends or relatives who could benefit directly from this work of this charity. For the importance rating, the scale of the rating ranges from unimportant (1) to extremely important (10). These twenty-four organizations are an explicit part of the context for the decision tasks that participants confront during the scanning portion of the experiment. See Table 8 for a list of the twenty-four charities along with the average (standardized) ratings and Figure 7 for an example of the rating questions from the charity rating survey.

For the fMRI portion of the experiment, we recorded the neural activity of participants during 120 trials in which participants made either mandatory or voluntary monetary transfers from their \$100 endowment account to the account of the charitable organizations. The initial run of 48 trials is the mandatory run (see Figure 8). Participants watched a monitor screen that was set up inside the fMRI station for information necessary to perform the task in a trial. Each trial consisted of a sequence of four different screens. During screen 2, subjects are provided the name of the charity, a short description about the cause of the charity, and the dollar amount of the mandatory transfer that will be made from their account to the account of the charity. At the end of the experiment, only one of the 48 trials randomly is actually implemented.

The set of charities that participants see during the mandatory run is a randomly generated twelve charity subset of the full twenty-four possible charities (see Table 8), each with an equal probability of selection. Thus, the subset of twelve charities varies by participant. For each of the charities, participants see four conditions: (i) a pure \$20 gain to their account, (ii) a pure \$20 gain to a charity's account, (iii) a pure \$20 loss to their account, and (iv) a \$20 gain to a charity's account coupled with a \$20 loss from their account. Next, participants acknowledge the mandatory transfer during the presentation of a second image ("Screen 2") by pressing a button on a specialized non-ferromagnetic "button box" located on the lap of the participants as he or she lies within the fMRI machine. After the subject's acknowledgement of the mandatory transfer (participants have no more than four seconds to respond) and a .5 second empty interval, subjects view a third image ("Screen 3") with an option to rate the transfer from either a 1

(highly dislike) to a 4 (highly prefer) in preference. Screen 3 has duration of four seconds. A new trial begins after a randomly assigned six, eight, or ten second fixation image on the monitor.

The set of 72 voluntary-giving trials follows the mandatory set after a very brief break initialize the presentation of the voluntary trials (see Figure 8). The set of voluntary trials extends the conditions of the mandatory trials in two ways. First, participants are informed on the monitor that their decisions will be observed (by the experimenter) or made privately (and not monitored by experimenter). For the mandatory run, all decisions made are privately. Second, participants choose to accept or reject the proposed monetary transfer from their account to the account of a charitable organization. The decision of a participant thus reflects their preference for their own well-being, the well-being of others, and the avenue via which other individuals will benefit. The possible transfer amounts are \$10, \$20, or \$40. Each participant sees a unique random sequence of the 72 voluntary trials that consist of a random sequence of pairings of each of the twenty-four charities with each of the three monetary transfers. The last screen for the voluntary run is identical to the last screen of the mandatory run in which subjects provide a rating (1-4) of the proposed monetary transfer. A new trial occurs after six, eight, or, ten seconds of the presentation of the fixation image.

### *fMRI Image Acquisition*

While participants performed the mandatory and voluntary runs of the experiment, we acquired functional magnetic resonance imaging (fMRI) data on their

neural activity. Functional MRI detects changes in the blood-oxygenation-level-dependent (BOLD) signal within ranges of a brain. BOLD signals have a positive correlation with neural activation. The imaging of blood oxygenation levels for each of the runs are obtained at a precision of 2mm x 2mm x 2mm voxels using a Seimens Allegra 3T MRI unit located in the Lewis Center of Neuroimaging at the University of Oregon. The repetition time (TR) of imaging is 2 seconds and each 2-second repetition consists of 32 slices. In addition to the functional MRI, we also acquire high-resolution anatomical images for each participant.

### *Behavioral Model*

We use a random effects binary probit specification to model the decisions of participants to accept or reject a transfer of money from their account to the account of a charitable organization. That is, we model the latent propensity to choose to accept a proposed transfer as

$$Accept^* = X_{it}\beta + v_i + \varepsilon_{it}, \quad (1.1)$$

where  $X_{it}$  is a set of trial-specific attributes that varies by participant  $i$  and trial  $t$ ,  $\varepsilon_{it}$  is an unobserved random error component that is independently and identically distributed across participants and trials, and  $v_i$  is an unobserved participant-specific random error that is uncorrelated with  $X_{it}$ . The estimated parameters of the model,  $\beta$ , represent the marginal effects on the latent propensity to choose to accept the proposed transfer and these individual preference parameters are assumed to be homogeneous across participants.

Individuals accept a proposed transfer when  $X_{it}\beta + v_i + \varepsilon_{it} > 0$ . The unobserved error term,  $\varepsilon_{it}$ , is assumed to be normally distributed such that for participant  $i$  on trial  $t$  the probability of accepting is given by,

$$\Pr[\text{Accept}_{it} = 1] = \Phi(X_{it}\beta + v_i). \quad (1.2)$$

With multiple observations on choices for each participant in our experiment (i.e., 72 choices for each participant), we model the random effects for different participants to be normally distributed. In our results section, we report the average “partial” effects instead of the marginal propensities to accept, with the later represented by  $\beta$ . Conditional on the estimated marginal propensities, an average partial effect represents the average effect across all observations of a change in a particular attribute on the *probability* of accepting, holding other attributes constant at their observed values.

Our behavioral analysis allows us to predict for each trial the probability that an individual will choose to accept or reject a proposed transfer of an amount of \$10, \$20, or \$40 to one of the charitable organizations. The attributes of the trial that we include in the model to predict individuals decisions includes participant-specific characteristics, information given to the participants during trial, and the estimated probit coefficients. For example, the attributes of a trial could include the type of charity organization, the transfer amount, and the age of the participant. These attributes can thus vary by trial and by participant. We construct our measure of the difficulty for each choice task of a trial with these predicted probabilities and use this constructed difficulty measure to model decision conflict in the fMRI analysis. We define  $D_{it}$  to be the absolute value of the probability distance between the predicted probability of accepting and the equal odds, or

indifference probability. That is, denoting  $\hat{P}_{it} = \hat{Pr}[Accept_{it} = 1]$ , we set  $D_{it} = 1 - |\hat{P}_{it} - 0.5|$ .

With this construction, the decision conflict or difficulty of a choice,  $D_{it}$ , increases as the absolute value of the probability difference increases.

### *fMRI Data Analysis*

The fMRI analysis of the neural activity of participants is performed with the FSL (Analysis Group, FMRIB, Oxford, UK) statistical analysis software. We generate voxel-wise parameter estimates of the hemodynamic (blood-oxygen-level-dependent) responses to the different stimuli in the experiment using a generalized linear model. These voxel-wise parameter estimates represent the change in the blood-oxygenation level for a given stimulus compared to the baseline neural activation of no stimulus presentation. The six- to ten-second fixation cue between trials represents the majority of the baseline in our statistical models. Other than those subjects who had severe head motion and are not included in the analysis, we correct for less-severe forms of head motion with a realignment of our time-series voxel-wise data (Jenkinson, et al. (2002)). The brain extraction tool (Smith, et al. (2004a)) is used to remove information about non-brain-related areas before performing the analysis. Functional data are registered to the Montreal Neurological Institute (MNI) standard.

For seven regressors, we model the change in neural activation from baseline activation during the voluntary run with a generalized linear model (GLM) for the hemodynamic response of voxels. The first three regressors are binary indicator variables for the presentation of three of the stimulus screens during the voluntary portion of the



experiment. These binary regressors include the observed cue of screen 1 (Sc1), the accept/reject decision at screen 3 (Sc3), and the rating at screen 4 for the proposed transfer (Sc4) (see Figure 8). Three of the four remaining regressors are ratings for the individual's subjective ranking of the relative importance of the charity causes. We construct these charity importance indicators by normalizing the importance ratings within participants into three mutually exclusive bins. These ratings represent the goal values that participants have for the various charity causes. Each participant's charity ratings are classified (1-unimportant, 10-extremely important) into low importance (C1), moderate importance (C2), or high importance (C3) categorical bins. The design of these bins is such that each participant's set of ratings uniquely determine the width of bins. The one constraint in the design of the bins is that the widths of the three bins are equal for a given participant. We leave three subjects out of the fMRI analysis because there was not enough variation across charities within each participant's ratings to classify the ratings into three mutually exclusive bins. The last of our regressors is for the predicted difficulty (D) of each trial. This is an estimated linear parametric regressor that modulates the onset of screen 2 with the value  $D_{it}$  obtained from the behavioral model.

The full set of regressors in the model thus includes Sc1, Sc3, Sc4, D, C1, C2, and C3. Each of these regressors is modeled as a unit impulse function with duration equal to that of its respective stimulus screen. For D, C1, C2, and C3, the duration is 7 seconds. Each unit impulse regressor is convolved with a canonical hemodynamic response function before performing the GLM analysis. FSL produces from the GLM analysis one-sided t-statistical images for the estimated change in baseline activation (averaged

across subjects) for each regressor at every voxel in the brain. These t-statistical images can be *contrasted* to generate images for testing other relations among the regressors. For our analysis, key contrasts include a test for significant increases in activation over baseline (averaged across subjects) for an increased level of difficulty of a trial (D+) and for significant increases in activation for a decreased level of difficulty (D-). We also test for a positive trend in the activation as a function of charity importance ratings (C+) across the three bins and test for a negative trend in activation levels for increasing importance (C-).

Our key contrasts are for D+, D-, C+, and C-. However, in an across-subject analysis, we test for increases in activation effects of these contrasts with increasing participant age. Thus, we generate the contrasts D+A+, D-A+, C+A+, and C-A+. Each of the contrasts for the four main effects, and the four age interactions, are Gaussianized into z-statistical images and thresholded at  $z > 2.3$  with a cluster-corrected significance threshold of  $p < .05$ .

## **Results**

### *Behavioral Factors That Affect a Decision to Give*

The difficulty of trial is a probability-space measure based upon the absolute distance of the predicted probability of accepting the proposed transfer from the equally-probable or “indifference” value of 0.5. Predicted probabilities for giving money to a charity are obtained from binary probit models to explain the accept decision which include participant-specific random effects, attributes of the trial, individual

characteristics, and information about participant behavior during the mandatory portion of the experiment. Model 1 of Table 9 show that basic characteristics—such as the proposed transfer amount (\$10, \$20, \$40), age of the participant, observation of the participant’s decision, and the participant-specific ratings of the importance of a charity—have strong significant effects on the latent propensity to accept.<sup>30</sup> In Figure 9 we show how the average acceptance varies by payouts and monitoring for each of the charities. Model 1 clearly supports the negative relationship between increasing proposed transfer amounts and average propensity accept to proposed transfer. In particular, an increase of \$1 in the proposed transfer amount decreases the probability of acceptance by 0.031, on average, across the observed data. In Model 2 and Model 3, we expand the behavioral determinants to include separate within-participant averages of the ratings of the pure monetary transfers to self and to charities during the mandatory portion of the experiment. The estimated parameters suggest that the probability of acceptance is negatively related to increasing satisfaction with money to self and positively related to money given to the charity. In addition, Model 2 considers the effects of gender, age by gender, observation differentiated by age, and observation differentiated by gender. Of these, only observation differentiated by age appears to have a significant effect on the probability of acceptance.

Model 3 of Table 9 retains only those factors that significantly affect giving. We report the marginal effects of these factors on the probability of giving in Table 10. We use the specification of Model 3 of Table 9 to generate the predicted probabilities of

---

<sup>30</sup> See Table 10 for estimated average partial effects on the probability of accepting and Appendix E for alternative random-utility models for predicted the probability of accepting.

accepting the transfer and to construct the difficulty measure,  $D_{it}$ , that varies by participant  $i$  and trial  $t$ . In a separate analysis testing the relationship in response time and our constructed measure of difficulty, we find response time decreases with a decreasing level of difficulty after controlling for whether an individual accepted or rejecting the proposed transfer. Figure 10 shows the frequency distribution of trials with a given difficulty level over the full range of  $D_{it}$ . The frequency distributions are roughly uniform for each participant. This suggests that, once participant effects are controlled for, the variation in difficulty used in the fMRI analysis is not primarily due to the observed characteristics of participants such as age. In the fMRI analysis, participant age is included as a covariate in the across-subject analysis of neural activity.

#### *Changes in Neural Activation Due to Difficulty and Charity Importance*

The fMRI analysis allows us to determine areas in the brain where there appears to be neural activation correlated with both decision conflict and goal valuation. We make the fMRI analysis of decision conflict operational via our constructed difficulty measure and treat the charity importance ratings as representative goal values. With our regressor of difficulty, we were able to test for neural activation that is positively correlated with an increase in difficulty (D+) and activation that is positively correlated with decreases in difficulty (D-). Neither of these contrasts for difficulty revealed any significant activation. In a confirmatory analysis, we explored several alternative specifications of the behavioral model used to generate the predicted probabilities for the difficulty measure and considered the alternative definitions of “entropy” (see Swait and

Adamowicz (2001a)) and the utility-space distance between alternatives for the construction of the difficulty. None of these alternative difficulty measures revealed significant activation changes in the contrasts of D+ and D-.

In an across-subject analysis, we explored the differential effects of age on neural activity due to increasing levels of difficulty. Specifically, we construct contrast that identify regions where either neural activation positively correlated with increasing difficulty either increases with age (D+A+) or neural activation positively correlated with decreasing difficulty increases with age (D-A+). Here, the results are more interesting. We find activation in the anterior cingulate cortex (see Figure 11; [0, 38, 40],  $z=3.54$ ,  $p<.0001$  corrected) for increasing levels of difficulty to be positively correlated with the age of participants. That is, older participants have a greater amount of activation in this area when the participant is relatively more indifferent between accepting and rejecting the proposed transfer. A number of other areas are also significantly active in this contrasted image of neural activation (see Table 11), although none of the other areas are implicated by existing neural evidence for decision conflict. There was no statistically significant evidence to support the reverse relationship between decreasing difficulty and increasing age for our contrasted image D-A+.

We examine the areas of the brain that encode goal values by utilizing the self-reported charity importance ratings by individuals. Regions activated by increasing charity importance (C+) include the caudate (see Figure 12 and Table 12; [-16, 14, 48],  $\max Z=3.31$ ,  $p=.0491$  corrected). In the construction of the regressors for charity importance (C1, C2, C3), ratings were binned within participants, which suggests that the

increasing activation is related to the increases in the *relative* importance of charities. An additional area in the occipital lobe appears to have neural activation that is correlated with increasing levels of charity importance. No areas are found to be correlated with the decreasing in charity importance (C-) and differentiation in activation for increasing or decreasing importance in the age of participants (C+A+ and C-A+).

## **Discussion**

The goal of our analysis has been to locate regions in the brain which might dually encode valuation and decision conflict. For valuation we focus on identifying areas correlated with goal values. We obtain measures of one type of goal value from ratings of the importance of twenty-four different charity causes. Goal values have been previously implicated in the orbital prefrontal cortex (OPFC), the anterior cingulate cortex (ACC), the ventromedial prefrontal cortex, the ventral striatum, and the dorsal striatum. In testing for activation correlated with decision conflict, we use a measure that maps into probability-space the similarity in utility-space between accepting and rejected a proposed transfer based on attributes of the trial and individual characteristics. Our analysis confirms previous evidence that decision conflict appears to be encoded by the ACC and the dorsal striatum as a region that encodes goal values. In addition, our analysis supports existing studies in the implication of the ACC in encoding decision conflict.

With our fMRI analysis, we do not find statistically significant evidence to support a unifying neural connection between our constructed decision conflict and our

measure of valuation that might affect the behavioral outcomes that take place in our experiment. Thus, we are unable to provide support for an economic theory that would suggest that the outcome of economic decision-making should jointly depend on valuation processes and the decision conflict that can arise during a decision. Even though our evidence does not support a potential mediating region, this does not necessarily mean that activation between the two regions is unrelated. We intend to address this issue in the future analysis by determining whether the areas we identify as important to the choice process, the ACC and dorsal striatum, may be functionally related to each other via correlated activation. To extent that cognitive capacity likely has a role in the ability of individuals to make decisions, age appears to be important. Specifically, we are encouraged by our evidence that age has a role in the activation levels for decision conflict and we intend to examine these effects to a greater degree.

## **CHAPTER III**

### **EXTREME WEATHER RISKS AND MIGRATION**

#### **Introduction**

Climate change is predicted to exacerbate risks to human health and alter the economic systems of vulnerable countries. In turn, vulnerability depends on the capacity of individuals and countries to mitigate and adapt to these risks. Some of the most vulnerable regions of the world are expected to have increased levels of inter- and intra-county migration, with changes in severe weather events as the main catalyst. An improved understanding of the relationship between changes in severe weather events and the decision to migrate will provide important information to businesses and governments as they consider mitigation and adaptation policies. At a micro level, however, little is understood about the evolution of economic and health risk perceptions due to severe weather events. Fluctuations in the spatial pattern, frequency, and severity of extreme weather events may induce different rates and patterns of migration as an adaptation mechanism. Some individuals will find the perceived economic and health risks to be too great and, thus, they will decide to move out of high risk areas. Others will choose to stay purely because of differences in risk tolerance and/or the presence of overriding constraints on migration.

Our goal is to predict potential adaptation in terms of new patterns of migration by different segments of society in response to perceived changes in the risks of extreme



weather events. We rely upon domestic data for U.S. migration as a function of tornado events as a “proof of concept.” Via these data and a number of alternative empirical characterizations, we are able to address the systematic nature of migration as a response to perceived economic and health threats from severe weather hazards. This study provides the most comprehensive investigation to date of the potential for migration as a response to changes in the frequency and pattern of one type of severe weather event likely to be affected by climate change in the U.S.

Existing economic studies characterize the potential consequences of a changing climate in the U.S. by using the historic relationship between economic and health outcomes and climate measures. Climate-related variables include, for example, measures of annual and mean daily temperatures (Deschenes and Greenstone (2007); Deschenes and Moretti (2009)) and hurricane events (Smith, et al. (2006)). Economic and health outcome variables include agricultural production (Deschenes and Greenstone (2007)), mortality (Deschenes and Moretti (2009)), birth weights (Deschenes, et al. (2009)), and, to a limited extent, migration (Deschenes and Moretti (2009)).<sup>31</sup>

We focus our empirical inquiry on US migration between 1992 and 2005 as a function of tornado activity in the United States, with controls for broad range of other natural hazards during the time period of the analysis. Individual tornadoes represent a

---

<sup>31</sup> Using data from the 2000 Census, Deschenes and Moretti consider individual-level migration decisions (aggregated to the state-level and grouped by age) from one’s birth-state to his or her current-state of residence. The authors note that they their findings “do not necessarily provide a causal interpretation” for the effects of extreme cold-weather events on migration due to the “many unobserved determinants of mobility” not accounted for in their empirical modeling. Our analysis is annual migration aggregated to the county-level that allows for a more complete set of controls for unobserved factors that may be important to the decision to migrate. Thus, any effects that we find for tornado activity on migration behavior are more likely to represent a causal effect.

relatively small threat to the economic livelihoods and health of individual Americans each year. However, roughly 800 tornados strike the U.S. each year and cause numerous injuries and deaths.<sup>32</sup> Figure 13 shows that tornado hazards over the last 58 years represent a non-zero risk to life, limb, and property for residents in large sections of the US. Furthermore, the estimated aggregate economic damage can run into the billions of dollars.

The historical spatial randomness in local tornado activity allows us to infer migration rates and spatial patterns *if* tornadoes everywhere were to become either more (or less) widespread, frequent, or severe as a result of weather trends due to climate change. Given that existing research concerning patterns of injuries and fatalities from extreme weather events has tended to concentrate on single localized events such as hurricanes, an important part of this research is the sheer breadth of community types and demographics within the United States that are represented in our data. Our results show modest but statistically significant changes in migration patterns in response to tornado occurrences within a county.<sup>33</sup>

The ability of the vulnerable portions of the population to restore tolerable levels of health and economic risks via migration is one of the key questions which motivates this research. We note that few existing studies consider the differential effects of extreme weather events for in- and out-migration separately. Instead, most focus on *net*

---

<sup>32</sup> The National Oceanic and Atmospheric Administration reports that “in an average year, 800 tornadoes are reported nationwide, resulting in 80 deaths and over 1,500 injuries. ...Damage paths can be in excess of one mile wide and 50 miles long.”

<sup>33</sup> In subsequent work, we plan to use our estimated models to simulate prospective future migration patterns for different demographics and communities in the face of alternative forecasts of changes in these extreme weather events as a potential consequence of climate change.

migration decomposed across demographic groups. We build a variety of statistical models and show that heterogeneity in tornado activity across space and time corresponds to changes in the patterns of both out- and in-migration across counties. The extent of the migration response to tornado activity depends upon the sociodemographic makeup of the origin county for out-migrants and the destination county for in-migrants. To a certain extent, tornado activity also seems to affect the distribution of distances traveled by migrants.

### **Extreme Weather Events and Perceived Risks**

Some of the objective risks of climate change include adverse impacts upon life, health, and welfare via changes in the pattern and severity of extreme weather events. For example, Ebi, et al. (2006) and Ebi (2009) draw attention to the prospect of increases in mortality and morbidity such as heart attacks, disease-borne illness, and diagnoses of posttraumatic stress disorder from changes in extreme weather events. The objective economic and health risks depend empirically on the underlying vulnerability of the population to these potential risk exposures. Objective demographic and socioeconomic characteristics of the population such as age composition and the level of poverty in an area can affect the overall outcome of exposure. Income levels are important because they affect individuals' abilities to engage in risk-mitigating behaviors, such as more building more wind-resistant housing structures, storm cellars, and so on.

Subjective risks of climate change are also likely to play an important role in understanding individuals' adaptation and migration behavior. Hunter (2005) concludes

that environmental factors can play a role in shaping migration decisions, particularly among those who are most vulnerable, and that perception of risk acts as a “mediating factor.” Recent evidence supports this notion for the risks of climate-change-related natural disasters. Carbone, et al. (2006) conclude that Hurricane Andrew affected perceived hurricane risks and caused a larger decrease in housing prices after the event in Dade County, the affected county, than in the less-affected Lee County. Smith (2008) finds that individuals directly affected by Hurricane Andrew appear to treat the storm as an information signal about long-term risk exposure. As a result, they update their longevity expectations. More recently, Baker, et al. (2009) use a stated-preference survey to examine residential location choices based on the subjective risks of hurricanes.

If the patterns and severities of extreme weather events could be expected to remain constant in the long run, people would spatially sort themselves across different risk zones to balance the discomfort of perceived risk from the natural hazards with the compensating positive features of all the different places in the U.S. where they might choose to live. However, if the *perceived* patterns and severity are altered by climate change, there will be adjustment costs as people adapt to these new patterns. With respect to the findings of Smith (2008), any migration in response to extreme weather events might suggest an adjustment in risk expectations, holding all else constant, as individuals gain experience with the different types of events over their lifetime. As one part of our analysis, we characterize historic risk levels of tornado activity at the county level and investigate the differential migration response to contemporaneous tornado activity within a county to the historic tornado risks of the county. The general perspective of our

analysis is in line with Baker, et al. (2009) who use a stated-preference survey to examine residential location choices based on the subjective risks of hurricanes.<sup>34</sup>

### **The Perceived Risks of Tornado Activity**

The direct health and economic impacts of tornadoes represent an ongoing and potentially changing source of climate-related mortality risk in the US. Tornadoes can strike almost anywhere across a wide range of Midwestern, Southern, and Eastern U.S. states, and with very short notice, providing little opportunity for evacuation as a mitigating behavior. Our logic in starting with tornadoes as an exemplar of extreme weather risks stems from the relative spatial and temporal exogeneity of tornado risks, conditional on individuals choosing to live in tornado-prone areas.

#### *Spatial and Temporal Variability*

For other types of weather-related hazards, such as floods or hurricanes, risk exposure can be mitigated locally to some extent by a judicious choice of residence and workplace locations. For tornadoes, however, the risks are relatively uniform over much wider areas. Whether you will be struck by a tornado is much less controllable by minor adjustments to residential location choices. This feature makes tornados an attractive case, empirically, because there is less of a concern with perceived tornado risk and residential locations being jointly endogenous at a fine level of spatial disaggregation.

---

<sup>34</sup> The Baker et al. study uses stated location decisions of individuals who had moved away from New Orleans in response to Hurricane Katrina. The sample in the study is limited to set of 78 individuals and has a narrow range of socioeconomic characteristics.

This relative spatial exogeneity within a county and for some broader regions of the US also suggests that the decision to relocate may entail significant travel costs to sufficiently lower the perceived risk of tornado activity. Our analysis points to this being the case. Tornado events that are the most severe appear to have the strongest effect on observed migration patterns.

The temporal exogeneity of tornadoes is also a benefit to our analysis. Temporary displacement (i.e. evacuation) is unlikely as a mitigation option for tornado activity given the relatively short lead times for advance warnings for tornadoes, and in some cases, the absence of any advance warning. In contrast, hurricanes have a more limited geographical extent in the US and typically have large lead times for preparation or evacuation warnings. To some extent, the familiarity and perception of historic levels of tornadoes in an area, and recognition of warning signs in the weather, will play a role in likely mitigation options. Recent retirees who move to a new area, perhaps on a lower retirement income, may need to adapt to unfamiliar weather-related risks and different warning procedures.<sup>35</sup> According to Ashley (2007) and Merrell, et al. (2005), tornadoes are more likely to produce fatalities when a tornado hits at night.<sup>36</sup> Likewise, Simmons and Sutter (2005) find that the availability and sophistication of storm cell detection technology, specifically the introduction of WSR-88D (Doppler) Radar in the 1990's, resulted in a substantial increase in the frequency and mean lead time of tornado

---

<sup>35</sup> This subpopulation may also be less aware of modern advanced warning technologies via new media. Some of these advanced systems include National Weather Service ATOM or CAP/XML formats/feeds for tornado watches, warning, and advisories.

<sup>36</sup> Merrell, et al. (2005) find both the time of day and season of a tornado to be important predictors of fatalities and injuries.

warnings. Thus improved availability and sophistication of detection technology has the ability to affect perceived risk exposure. Although we do not include time of day of a tornado event in our analysis, we do include state-year indicators for the period of our analysis which allows us to identify effects on migration patterns controlling implicitly for any overall regional changes in radar technology over the period our analysis.

#### *In-flow and Out-flow Migration Asymmetry*

Our research focuses on migration decisions as *ex ante* responses to perceived increases in risks, not just reactions to individual tornado events themselves, such as evacuation beforehand or displacement due to damage to dwellings directly in a tornado's path. Further, our analysis will treat in-migration and out-migration separately and enable us to determine if those who are most deterred from regions of perceived higher risk are current residents who leave or potential in-migrants who choose other destinations instead. Landry, et al. (2007), explore return-migration decisions by Hurricane Katrina survivors. Smith, et al. (2006) find that migration by Florida households in response to Hurricane Andrew in 1992 is largely predicted by the "economic capacity" of households. Likewise, Paul (2005) and Myers, et al. (2008) consider the behavior of victims after the disaster has occurred. However, these survivors have already been victims of a severe weather event. In contrast, we consider migration not just by "refugees" for specific-events, but by the general population in response to elevated *ex ante* risk perceptions based on events that have directly affected others in their county.

In- and out-migrants will have different reasons, such as employment or family considerations, for remaining in, or relocating to, a particular region. More specifically, Myers, et al. (2008) examine post-disaster migration specifically in the wake of Hurricanes Katrina and Rita along the U.S. Gulf Coast. They find that the county-level net out-migration caused by these two hurricanes was significantly greater among groups with lower socioeconomic status, for areas that suffered greater property damage, and for areas that were originally more densely populated. However, net out-migration may increase after severe-weather events because households move out in increasing numbers, or because potential new residents fail to move into these regions, choosing other destinations instead. Likewise, potential in-migrants and out-migrants will have different levels of experience (or none, maybe, for some in-migrants) with the type of extreme weather event in question. These varying risk experiences and different reasons for migration will likely produce asymmetric migration responses for these two groups.

#### *Demographic Differences*

The objective risks of tornadoes are likely to have impacts that differ across demographic groups in the population. Those more at risk include older seniors and the disabled because of their lesser mobility (Merrell, et al. (2005); Ebi, et al. (2006)) and individuals with compromised immune systems are more susceptible to the risks of illness as result of contamination of water or air in the aftermath of tornadoes. A common view is that divisions might exist along lines of ethnicity and housing types when it comes to the potential mortality and health risks of tornadoes and other extreme weather



events. However, there is considerable overlap between areas with high risk and areas with a high percentage of housing that consists of mobile homes. Mobile and manufactured homes tend to be occupied, on average, by households with lower incomes and fewer assets.<sup>37</sup> A greater proportion of mobile home occupants will tend to belong to either low-income and/or minority groups.<sup>38</sup> In addition, Merrell, et al. (2005) find that the percent of housing consisting of mobile homes within a census tract, and the intensity of a tornado, are important predictors of fatality and injury counts.

Our data do not allow us to control at an individual or household level for all the potential factors that might affect migration behavior. However, we are able to account for county-level heterogeneity in other time-varying weather conditions, prevailing unemployment rates as an indication of local economic activity, and the overall demographic composition of the county's population.<sup>39</sup>

## **Data**

We combine four different data sources to investigate migration as an adaptive behavior to perceived increases in tornado risks. These data provide a panel of county-year observations from 1992 through 2005 that consists of roughly 3,141 U.S. counties

---

<sup>37</sup> Newer manufactured homes are considerably less likely to be leveled than older homes, due to stricter construction and tie-down codes enacted by the Department of Housing and Urban Development after Hurricane Andrew.

<sup>38</sup> As evidence against the ethnic disparity hypothesis, Smith, et al. (2006) find that middle-income white households were more damaged than poor, minority households in the aftermath of Hurricane Andrew.

<sup>39</sup> In subsequent work, we plan also to explore the likely influences of long-term trends in several other types of hazards and recent disaster events on individuals' migration decision by using more extensively the data from Federal Emergency Management Agency (FEMA) on the Presidential Disaster Declarations. Over the longer term, these types of events may also result in an increase or decrease in risk-averting behaviors such as migration.

across 14 years and yield 43,974 observations. The data we obtained data on migration flows from the Internal Revenue Service (IRS) provides annual county-to-county out-migration and in-migration data for the entire set of 3,141 U.S. counties. “Matched” migrant household and individual returns across tax years are based on social security numbers (SSN) for the primary filers.<sup>40</sup> To establish migration, residential address information is extracted from the domestic household tax forms 1040, 1040A, and 1040EZ, and the foreign tax forms 10NR, 1040PR, 1040VI, and 104SS. The extraction process occurs until the 39<sup>th</sup> week in the IRS’s processing year and covers 95% to 98% of all returns filed for a tax year. Thus the data are not complete, but they represent the most detailed data available on an annual basis for the entire U.S. at the county level.<sup>41,42</sup>

With the county-to-county level migration flows and the ArcGIS software package, we construct a set of five basic dependent variables for our analysis: (1) the number of out-migrants (or in-migrants) from county  $i$  to any other county based on county  $i$ ’s population, (2) the number of out-migrants (or in-migrants) from county  $i$  to any other county based on county  $i$ ’s population, (3) the mean distance traveled from county  $i$  to any other county (with the distance from county  $i$  to each county  $j$  weighted by number migrants for that county pairing), (4) the standard deviation in distances

---

<sup>40</sup> See Appendix F for a detailed description of the data.

<sup>41</sup> The IRS also censors migration flows to protect the individual identities of taxpayers, so only “non-trivial” flows (i.e. ten or more migrant household tax returns) are identified at the county-to-county level going back to 1992. All other flows are aggregated up to a larger geographic region as their destination (i.e. state-to-state or region-to-region).

<sup>42</sup> Tax years lead the actual year of filing such that a “match of tax years 2003 and 2004 produces 2004 to 2005 migration estimates.” Individuals could experience a tornado between January 1<sup>st</sup>, 2003 and December 31<sup>st</sup>, 2003, move to a different county on January 1<sup>st</sup>, 2004, and subsequently be classified as a migrant for the year of 2004. For this reason, we combine current year tornado activity with the activity of the previous year in some of our empirical models of migration flows in our results section.

traveled, and (5) the skewness of the distribution of distances traveled. Table 13 provides the summary statistics for the ten (out-migrant and in-migrant) dependent variables.<sup>43</sup>

We utilize data from the National Oceanic and Atmospheric Administration's National Weather Service (NWS) Severe Weather Database Files to construct our tornado activity variables. The NWS data include path information such as date, severity, and size for each reported tornado in the U.S. going back to 1950.<sup>44</sup> Documented information about the impacts of each tornado includes the number of injuries, fatalities, total monetary damages, and total crop losses. The tornado activity variables we consider in our analysis include the total number of tornadoes, the total number of fatalities, the total number of injuries, the average Fujita-scale intensity strength, and the total monetary loss for property damages.<sup>45</sup> Tornado events in the data include geographical information on the starting and ending latitudinal and longitudinal coordinates. With this information, we construct mean and aggregate level amounts of activity (e.g, mean severity and total count of tornadoes) within each county and buffer zone for each of our

---

<sup>43</sup> Ideally the summary statistics for any single county pair would be identical across out-migrants from the first county and in-migrants to the second county, because county-to-county observations can be thought of as just a symmetric flow matrix. However, the data the IRS provides comes in separate files for in-migration and out-migration. As apparent in Table 1, our processing of these data reveal some inconsistencies in the raw data across the two migrant types. Instead of choosing one or the other set of migration flows to use for our analysis, we proceed by using both.

<sup>44</sup> Until the adoption of Doppler radar by the N.W.S., tornado reports were more likely to occur in areas with larger populations. We utilize data on population size and incorporate a time specific indicator for technology adoption to control for this reporting bias over time. However, if a tornado is not even recorded, it may be safe to assume that it would not have a significant impact upon perceived risk and thus little effect on migration behavior.

<sup>45</sup> Theodore Fujita developed The Fujita Scale in 1971. The scale assigned to a tornado is an estimate of the wind-speed of a tornado based on an assessment of the damage to buildings. The scale is as follows (in miles per hour): F0 (40-72), F1 (73-112), F2 (113-157), F3 (158-206), F4 (207-260), F5 (261-318). The Enhanced Fujita-scale (introduced in 2007) addresses a number of problems with the original scale. See *A Recommendation for an Enhanced Fujita Scale (EF-Scale)*. Lubbock, Texas: Wind Science and Engineering Research Center, Texas Tech University. October 10, 2006, Rev. 2 and Doswell Iii, et al. (2009).

tornado variables. Specifically, we associate our tornado variables with each of the spatial jurisdictions, i.e., counties and our constructed county buffers, by first generating the paths of the tornadoes using the starting and ending coordinates and then spatially intersecting the paths with the corresponding jurisdictions. We make the assumption that the paths move unidirectionally and approximately linearly, based on the starting and ending coordinates given by the NWS.<sup>46</sup>

We create spatial measures of tornado activity in proximity to a county for a county-level spatial analysis of tornado risk. With the U.S. Census Bureau's Census 2000 County and County Equivalent Areas cartographic boundary files, we map tornado activity within each county and within two buffer zones around county boundary, a 0-to-20-mile zone and a 20-to-50-mile zone. These buffers zones are created with GIS software and are geographical buffers that extend outward from the perimeters of the county boundaries. These spatially delineated tornado activity variables enable us to investigate the effects of tornado activity in proximity to a county's jurisdiction.<sup>47</sup> Figure 14 shows an example of the county boundaries and buffers zones.

---

<sup>46</sup> Early accounts of tornadoes were often from direct observation. Recorded tornadoes have increased over time with increases with population size and detection technology. These time trends can also be spatially distinct because of differences in population growth and changes in detection technology across geographic regions. In addition to the upward trends in the data, which we control for with state-year fixed effects in our empirical analyses, some of the recorded tornadoes have missing values for the ending coordinates of the path of the tornado. This is also related to detection capabilities. These tornadoes still contain relevant spatial information, mainly the county of occurrence, and sometimes information on injury and fatality impacts. We do not drop them from the analysis but, instead, impute a short path (approximately 100m) based on the starting coordinates to make the tornadoes "operational" in the GIS calculations.

<sup>47</sup> Available at: <http://www.census.gov/geo/www/cob/co2000.html>. The structure of these boundary files is such that, for some counties with non-contiguous geographic boundaries, e.g., mainland and an island, the individual elements of the county are not treated as observationally equivalent. We recode or "dissolve" these separate entities for a single county into a single county boundary observation.

Other county-level characteristics come from the Bureau of Labor Statistics (BLS) (annual level data from 1990 to 2005) and the Neighborhood Change Database (NCDB) (decadal data from 1970 to 2000) produced by GeoLytics. The NCDB database includes a subset of the Census long-form data on counts for different of sociodemographic groups by census tract. Census tracts are designed to be relatively homogeneous with respect to population size, household characteristics, economic status, and living conditions. However, as populations grow over decades, tracts are split to maintain as much as possible the intended number of persons per tract. The advantage of the NCDB is that the spatial extents of the 1970, 1980, and 1990 census years are normalized to the 2000 census tracts. We use the 1990 and 2000 NCDB data to build a spatially and time-indexed panel of the NCDB data via linear interpolation for the years 1990 to 2000, and linear extrapolation back to 1985 and forward to 2005. We then aggregate the year-wise tract-level data in the demographic information for the 65,232 census tracts up to the county level. The linear interpolation and extrapolation involves strong assumptions. However, it provides us with a fourteen-year period (1992-2005) of evenly spaced proxies for annual time series observations for every county in the U.S. Table 14 provides a summary table of the full set of independent variables for our empirical specifications.

The last of our data sources is the Presidential Disaster Declarations (PDD), obtained directly from the Federal Emergency Management Agency (FEMA). This database catalogs the entire set of formal disaster declarations dating back to 1964 with their afflicted counties, and provides the specific dates for declared disasters. Table 15

shows the total number of county-level declarations in the U.S. since 1964. These declarations represent only the most extreme and widespread types of natural disasters. Under the somewhat strenuous assumption that political factors never influence the probability of a disaster declaration, we use these data to control coarsely for other major weather-related events other than tornados, such as severe storms and hurricanes, which may also affect migration behavior.<sup>48</sup>

### **Modeling Migration in Response to Changes in Perceived Tornado Risks**

Models of migration in the economics literature have largely developed within the fields of population and development, labor, international trade (via the influence of gravity-type empirical model specifications), and environmental economics. In the environmental literature, migration can be a response to spatially-differentiated environmental quality. For example, Sieg, et al. (2004) and Smith, et al. (2004b) show migration as part of the process of re-equilibration in response to a large change in air quality over space. Cameron and McConaha (2006) and Banzhaf and Walsh (2008) contribute further empirical support for the argument that a change in perceived environmental health risks affects the locational equilibrium of households and thus the sociodemographic and income-level composition of communities.<sup>49</sup>

---

<sup>48</sup> Although a two-stage least squares approach could be adopted in the empirical modeling, the broad range of socioeconomic and demographic factors that we include in our empirical specifications diminishes the extent to which the political endogeneity of the disaster declarations might affect our results with respect to tornados.

<sup>49</sup> In the social science literature, Hunter (2005) reviews classic migration theories and interprets them specifically in the context of an environmental hazard, conceptualized in her case as a local disamenity such as a toxic waste disposal facility. Existing economic studies typically employ some sort of hedonic analysis to infer evidence of aversion to disamenities or natural disasters by analyzing changes in housing prices in areas near or around the negatively affected site(s). Of the economic studies that consider natural

*Count of Migrants*

Our most basic model of migration behavior is given by

$$\log(M_{it}) = \beta T_{it} + \gamma S_{it} + \alpha_i + \alpha_{st} + \varepsilon_{it} \quad (1.1)$$

where  $\log(M_{it})$  is the logarithm of the number migrant household tax returns identified by the IRS for either in-migration or out-migration.  $T_{it}$  is some measure of tornado activity permitted to affect migration rates in that county and year. In our results section, we report estimates for a range of possible measures of tornado activity that include the number, type, geographic scope, severity of tornadoes in a county. Given this basic log-linear specification, the key coefficient  $\beta$  can be interpreted roughly as the (decimal) percent change in migration numbers as a result of a one-unit increase in the tornado variable.

Other variables that may affect migration decisions are captured by  $S_{it}$ , a vector of other time-varying information about county  $i$  in year  $t$  including socioeconomic and demographic characteristics, county population (in log form), and county unemployment levels (in log form). To minimize any potential heterogeneity bias that may stem from omitted variables that are correlated with  $T_{it}$ , our analysis includes county- and state-year fixed effects, the terms  $\alpha_i$  and  $\alpha_{st}$ , respectively. County-level fixed effects account for any potential differences in unobservable features that are constant within a county over time. Important unobservables may be the county's time-invariant geography and

---

disasters, the disasters in question are primarily single hurricane events, such as Hurricane Andrew in Florida, or Hurricanes Katrina and Rita on the Gulf Coast.

topology. Conditional on  $T_{it}$ ,  $S_{it}$ ,  $\alpha_i$  and  $\alpha_{st}$ , we assume the error terms,  $\varepsilon_{it}$ , are independently and identically distributed with mean zero. We use robust standard errors clustered on counties in our hypothesis testing of the estimated parameters of the model.

When modeling out-migrant tax returns,  $M_{it}$  is specifically the aggregate number of household tax returns with an address in county  $i$  in year  $t-1$  that had an address in any other county  $j$  in year  $t$ .  $M_{it}$  We also employ a specification analogous to equation (1.1) in separate models for in-migrants with an address in county  $i$  in year  $t$  who had an address in county  $j$  in year  $t-1$ . If scaled by the average number dependents per household tax return (which we will assume is roughly constant across counties in the US)  $M_{it}$  can be thought of as simply the number of migrant individuals.

Concerning the different fixed effects,  $\alpha_i$  and  $\alpha_{st}$ , the county fixed effects capture whatever may be constant over time for each different county. For example, it may be easier to detect the occurrence of a distant tornado in relatively flat counties but, for the same reason, exact geographical information on the path and the starting and ending coordinates of the tornado may be harder to verify. Thus, relatively flat counties may be expected to have systematically higher counts of tornadoes but, conditionally, a lower percentage with complete data on both the starting and ending coordinates. State-time fixed effects capture overall national changes in population size and tornado detection technology that may vary across time and regions larger than the county-level. The state-year fixed effects also capture time-varying differences in state-level disaster preparedness and relief policies. We are also careful to control for non-tornado climate-



related or weather-related disasters that may be correlated with tornado activity and might also affect migration decisions. Other disasters, such as severe storms or hurricanes, also have the potential to vary as the climate changes. We control for these confounding events with the time-wise and spatially varying Presidential Disaster Declarations.

Our general empirical strategy relies upon separate models for in-migration and out-migration.<sup>50</sup> As a consequence, the interpretation of the explanatory variables and their corresponding effects in all our empirical models depends on type of migration considered. With out-migration, the explanatory variables are the characteristics of the origin county. The explanatory variables for in-migrant flows describe characteristics about the destination county.

### *Historic Risk and Temporal Dynamics*

We extend equation (1.1) to include lagged variables for tornado activity such that

$$\log(M_{it}) = \beta_0 T_{it} + \beta_1 T_{i,t-1} + \dots + \beta_k T_{i,t-k} + \gamma S_{it} + \alpha_i + \alpha_{st} + \varepsilon_{it} \quad (1.2)$$

The  $k$  lagged variables allow us to identify the apparent effects on perceived future risk due to a tornado event in the current period. One would anticipate that the coefficients on these variables to diminish over increasing lags until, for some value of  $n$ , the effect is no longer apparent.

Further, the extent to which the history of severe weather event(s) may affect in-migration and out-migration via perceived risk has not been studied statistically.

---

<sup>50</sup> An alternative approach we may seek in future analyses when modeling migration includes spatially-dependent models (LeSage and Pace (2008)). Likewise, Davies, et al. (2001) provide a discrete-choice conditional logit approach for modeling U.S. state-to-state migration for an interval 5-year period.

Anecdotal evidence suggests that some people may view a single local tornado occurrence as a one-time random event, but after multiple tornadoes in their vicinity, the cost of relocating may be overwhelmed by updated subjective probabilities concerning the likelihood of a future tornado, resulting in a move to another location which is perceived to be safer. To explore this possibility, we construct a variable that captures elapsed time since the last previous tornado in a county (see Table 14 for summary information) and explore the usefulness of this variable instead of the lags in equation (1.2). To investigate the extent to which long-term locational sorting behavior by individuals due to historic risk may affect the overall observed response of the current population in a county to current tornado activity, we model the interaction (not shown in equation (1.2)) between the mean number of recorded tornadoes in a county from 1950 to 1985 with contemporaneous tornado activity during the 1992 through 2005 sample period of our analysis.

### *Spatial Dynamics*

Using only own-county tornado activity to model migration may be an overly simplistic assumption about individuals' capacities to perceive and internalize spatially-delineated severe weather risks. Thus, we also explore broader spatial patterns of tornado activity that might affect changes in migration. An obvious possibility is that tornadoes in neighboring or nearby counties may matter. Our use of GIS software allows us to develop variables that provide more useful measures of distance in this regard. Our primary set of 'lagged' spatial variables is the additional tornado activity that occurs in buffers (of

varying widths) around an origin county. We include these buffer activity variables in our specification such that

$$\log(M_{it}) = (\beta_0 T_{it} + \beta_1 T_{i,t-1} + \dots + \beta_k T_{i,t-k}) + (\delta_1 T_{i20,t} + \delta_2 T_{i50,t}) + \gamma S_{it} + \alpha_i + \alpha_{st} + \varepsilon_{it} \quad (1.3)$$

where the subscripts  $i20$  and  $i50$  indicate, respectively, 0-to-20-mile and 20-to-50-mile buffer zones around county  $i$  in year  $t$ , not including county  $i$ . These variables are constructed such that  $T_{i50,t}$ , for example, may represent the additional tornado counts located within a 50-mile buffer of the origin county that do not already cross the origin county or the 20-mile buffer zone, i.e. the path must be located within the 20-mile and 50-mile buffer rings. Thus, the coefficients on these variables are the added out-migration due to tornado activity that occurs only within a buffer zone of a given width outside an origin county.<sup>51</sup>

### *Conditional Distribution of Migration Distances*

The second type of migration behavior we wish explain is the distance migrants may move, conditional on moving out of their current county. We calculate this by taking the distance from county  $i$  to each other county to which anyone moved (when at least ten families moved to that other county), weight by the number of people moving to that county, and sum. This weighted (mean) measure of distance moved for significant out-migrants flows,  $D_{it}^m$ , provides us with our primary measure of distance. We use this aspect

---

<sup>51</sup> In future work, by varying the numbers and distances for these non-overlapping buffer zones, we may be able to infer the distance outside a particular county beyond which the effect of an additional tornado declines essentially to zero.

of migration behavior as an alternative dependent variable, in a model analogous to equation (1.1):

$$D_{it}^{mean} = (\beta_0 T_{it} + \beta_1 T_{i,t-1} + \beta_2 T_{i,t-2}) + (\delta_1 T_{i20,t} + \delta_2 T_{i50,t}) + \gamma_{it} S_{it} + \alpha_i + \alpha_{st} + \varepsilon_{it} \quad (1.4)$$

Finally, it is possible that mean distance traveled by migrants would be unaffected by tornado activity but that other higher moments in the distribution of distances could change.<sup>52</sup> We thus model separately the standard deviation and skewness of the distribution of distances traveled,  $D_{it}^{stdev}$  and  $D_{it}^{skew}$ , in addition to  $D_{it}^{mean}$ .

## Results

Our analyses of in-migration and out-migration behavior allow us to test for factors that jointly determine the net migration: in-migration, and out-migration.<sup>53</sup>

Whereas previous research has mostly focused on the typically negative changes in net outmigration migration for afflicted counties, we acknowledge that a negative net migration could result from an increase in out-migration and/or a decrease in in-migration. In this section, we review our findings for empirical specifications of equations (1.1) - (1.4). Our analysis examines several types of outcome migration response variables (*# of migrant tax returns, # of claimed exemptions, the weighted-average distance moved from any one origin county to the range of destination counties, the standard deviation of these distances, and their skewness*). We also explore in our

<sup>52</sup> For instance, a positively skewed distribution has a “tail” that is pulled in a positive direction. Skewness in moving distances might decrease as a result of current or recent tornadoes. The mass of destinations would tend to become farther away, resulting in a negative coefficient for tornado activity.

<sup>53</sup> We could treat these as seemingly unrelated regressions, but our reliance on identical regressors implies that there will be no gain in efficiency from joint estimation.

analyses two different subsets of specific types of migration flows to assess differences in responsiveness for origins and destinations with very different levels of historic tornado activity.

### *Frequency and Intensity Effects on the Size of Migrant Flows*

Model 1 in Table 16 reports the results for our most basic specification for examining the effects of current-period tornado activity on out-migrant behavior. The number of out-migrants who leave a county and move to another county increases by 0.13% for each additional tornado in that same county and year. With an average of roughly 2,400 migrant households per county in the US in the 2005, this translates into an additional 3 households, or roughly 7 individuals, who move all the way out of that county for each additional tornado occurrence in a county. In comparison to the positive effect that we find for generally worsened local economic conditions (captured by increases in the logarithm of unemployment levels), the effect of additional tornado occurrences on counts of out-migrants is relatively small. However, in an absolute sense, aggregated to the national level across over 3,100 counties, the overall increase in the number of migrants is sizeable at the US level. In 2005, 750 counties had a least one tornado occurrence, which is about quarter of all US counties. If 10% of these counties had at least one additional tornado occurrence, then 525 additional individuals would suffer sufficient disutility from these tornadoes to induce them to move all the way to another county.

The results of Model 1 include county-level fixed effects to control for historical risk levels across counties which might be correlated with the average types of individuals who self-select to live in counties with a higher historical tornado risks. Hence, these results are robust to long-run equilibrium residential location choice that would result from historical activity. However, if the socioeconomic composition is changing over time then so might those individuals that are most likely to move in response to an event. In Model 2, where we account for an assortment of time-varying demographic changes within counties, we find the magnitude of the effect of current year tornadoes on migration rates seems to diminish slightly.<sup>54</sup>

Models 1 and 2, in addition to all of the rest of our models, include state-year effects. The inclusion of these effects is important for two reasons. Simmons and Sutter (2005) find that the introduction of advanced Doppler radar technology dramatically decreased the level of fatalities and injuries, and increased the advanced lead time of warnings and detection of tornadoes. These advanced radar systems were introduced at different periods both across and within states such that the exclusion of state-year effects would potentially lead to the conclusion that increases in tornado activity had little or no effect on migration behavior. Thus, a downward bias would be expected. Second, tornado activity within a county might be correlated with other conditions or events at the state-level that would also be expected to have an effect on migration behavior within the county. Such an event may be a year with severe weather that produces not only a potential increase in tornado activity for a particular state within a given year but also an

---

<sup>54</sup> The time-varying demographic variables account for the composition of race, education, worker type, housing characteristics, and income levels in a county. See Table H5.1 in Appendix H for a full list of socioeconomic variables included in the model specification.

increase in flooding, both of which could lead to changes in migration behavior. These effects would lead to an upward bias in our estimated results. Again, the state-year effects provide a mechanism for which to control for these potential confounds.

Model 3 in Table 16 further addresses the concern for time-varying weather events that could lead to an upward bias in the estimated effect for the number of tornado occurrences within a county. This model includes county-year specific controls for presidential disaster declarations by incident type (see Table 13). (We report the full result of the rest of the models of Table 16 in Appendix H.) The effect for the number of tornado occurrences diminishes slightly to a .095% change in the number of households who leave the county. Thus, conditional on the other covariates in the model, the controls for county-level variation of other severe weather events within a state have only a modest impact on our estimates of the influence of tornado activity on migration behavior.

Patterns of migration might differ in important ways across out-migration and in-migration decisions. For instance, in-migrants from farther away may be less likely to be aware of the recent tornado history of potential destinations, and thus more likely to move in to a recently tornado-afflicted county than if they were moving from a nearby county. Models 4 through 6 consist of analogous specifications for in-migration behavior. Our findings suggest little if any change in in-migration behavior as a result of tornado activity in the destination county. However, while they are not statistically significant, the signs of the estimated effects suggest that counties may indeed have fewer in-migrants with each additional tornado event.

Model 1 in Table 17 is identical to Model 3 of Table 13 except that the tornado activity variable, *# of tornado occurrences*, has been recalculated to include both current-year activity and previous-year activity. We do this because the timing of the tax return starting and ending filing dates might result in some households experiencing a tornado in a year for which the tax year is the next calendar year.<sup>55</sup> The remainder of the models in Table 17 and subsequent tables in this paper employ this two-period cumulative tornado occurrence variable. In general, Table 17 shows that, while controlling for the *number* of tornado events in a given year, the recorded intensity/severity tornadoes in a county (i.e. the number of tornados at each Fujita-scale rating) has a negligible impact on out-migration behavior and in-migration. An exception appears to exist for the strongest of tornadoes—those with an F5 rating. The strong tornadoes have a large positive effect on out-migration and narrowly miss significance at the 10% level.

#### *Effects by Household Types*

Although not shown in any of the tables discussed thus far, our analyses reveal that migration behavior seems to differ according to the demographic and economic mix in the origin county for out-migrants and in the destination county for in-migrants. In Model 1 of Table 18 we control for the sociodemographic characteristics of the affected counties and simultaneously interact these sociodemographic characteristics with our key variable of interest, *# of tornado occurrences*. This permits us to examine the degree to

---

<sup>55</sup> For example, if an individual experienced a tornado on March 1<sup>st</sup>, 1999, moved on April 1<sup>st</sup>, and filed a tax return prior to April 15<sup>th</sup>, then his or her “migrant year” would be recorded as 1999. However, if the person waited until after April 15<sup>th</sup> to move and file a tax return (in the following IRS filing period) then the migrant year would be recorded as 2000. Of course, a household could wait any number years or time before deciding to move *in response* to a tornado event.



which the average out-migration might differ systematically with the socioeconomic characteristics of the origin county (i.e. with these variables as intercept-shifters), as well as how the response of out-migration to tornado events might differ with the same variables (i.e. as slope shifters). Each model in the table includes two columns of coefficients. The first column reports the effects of the sociodemographic variables on expected migration rates. The second column reports the effects of these variables on the derivative of log-migration with respect to tornado activity. Model 1 shows that the sociodemographic variables exert many statistically significant effects on average migration rates. However, when all of these (somewhat correlated) sociodemographic variables are employed to shift the slope coefficient on the tornado variable, none of the coefficients on these interaction terms is individually statistically significant other than the “% w/ some h.s.” and “% farm, fish, or forest workers” interactions. Model 1 in Table 18 can be compared to the corresponding model without sociodemographic interaction terms in Model 3 of Table 13. Acknowledging that multicollinearity among these sociodemographic shift variables will tend to produce some bias in a more parsimonious model, we nevertheless find that the more limited specification in Model 2 displays four individually statistically significant shifters, with mixed signs, that could be offsetting if this heterogeneity were not permitted to manifest itself.

These results reveal several interesting insights for the estimated number of out-migrants that leave a county due to increasing levels of tornado activity. Model 2 shows that counties that have larger portions of college educated individuals (i.e. “% w/ some college” and “% w/ college degree”) all else equal, have a larger number of out-migrants

in general, but also more out-migrants as a result of increased tornado activity. Likewise, those counties with a larger share of individuals without a high school diploma have, on average, less out-migration, but tend to have an increasing amount of out-migration with higher levels of tornado occurrences. We find that counties with larger shares of farming, fisheries, or forestry workers have a statistically significant larger amount of out-migrants in response to tornado activity, although the baseline effect of this share for the level of out-migration from a county does not seem to matter. This finding might reflect that for workers in these industries that consist of higher levels of outdoor exposure and, perhaps, relatively fewer adequate places to take immediate shelter may perceive themselves to be more at risk from tornadoes. We also find that counties with a larger percentage of occupied mobile housing are likely to have fewer out-migrants in general and (counter to our initial conjectures) to have a lower response in terms of out-migration when an additional tornado strikes. This result is present even though we control for income in this model (which itself increases expected migration rates but does not influence the effect of tornadoes on migration rates).

For in-migration flows, we look at a comparable set of socioeconomic terms in a fully general specification in Model 3 and a parsimonious specification in Model 4. In Model 4, we find a statistically significant greater in-migration to counties that have a higher proportion of college-educated individuals or larger shares of farming, fisheries, or forestry workers. In-migration is lower to destination counties that have a larger share of individuals without a high school diploma, have higher average household income and house value, or have a larger population share in renter-occupied or mobile housing. Less

in-migration also occurs where the destination county has more workers who do not reside in that county.

Several socioeconomic variables also seem to affect the extent to which in-migration varies with the number of tornadoes afflicting the destination county. In-migration in response to an additional tornado is less where there are more people with less than a high school diploma, but it is also less when more people in the destination county have a college degree. However, there is more in-migration in response to tornadoes with a higher proportion of individuals who have attended (or are attending) college but have not finished their degree. Destination counties with a higher proportion of blacks and that have higher house values are more likely to receive in-migrant flows.

#### *Changes in Distances Traveled and Variation by Subflows*

In this subsection we broaden our analysis to consider different dimensions of migration behavior. The first of these additional outcome variables, *the number of claimed exemptions* by migrant households, is potentially a more useful measure if the outcome of interest is a more precise measure of the number of individuals who move (instead of assuming some constant average household size for the number of tax returns, as we did in the previous sections). In our models, we take the logarithm of the total tax returns that are considered migrant and the logarithm of the total claimed exemptions. This means that the estimated effects for unit changes in our independent variables represent proportional changes in these migration behavior outcomes.

We examine the estimates reported in Models 3 and 6 in Table 13 of the effects from the number of tornado occurrences on the total count of out-migrants and in-migrants in a similar fashion for our four other outcome variables. The specific coefficient of interest in every one of the models for these outcome variables is the effect of an additional tornado event on the outcome variable in question. Table 19 economizes on space by reporting only this coefficient for thirty different models. Outcome variables are listed on the left side of the table. The columns of Table 19 report the estimated effects (and standard errors) across the different samples of the data. The first column reports the results for the full set of county-to-county out-migrant flows in the conterminous U.S. The second column includes only those out-migrant flows in which individuals departed from an origin county in one of the twenty historically most tornado-prone states for a destination county strictly outside this set of states.<sup>56</sup> The third column reports the results for out-migrants from the ten historically most tornado-prone (a subset of those states used for the second column) that had a destination county outside those states. The fourth through sixth columns are identical to the first three except that in these models the analysis is for in-migrant flows.

The rows of Table 19 correspond to different candidates for the role of dependent variable in all of these models. The first two rows show that the estimated fractional change for an additional tornado occurrence are roughly similar between the *# of migrant tax returns* and *# of claimed exemptions* for each of the different sample types. This is

---

<sup>56</sup> States we identify as the 20 most-prone states are: MS, KS, OK, NE, AR, IN, IA, AL, IL, WI, GA, TN, NC, OH, LA, KY, PA, MI, MO, and TX. The 10 most-prone states include: MS, KS, OK, NE, AR, IN, IA, AL, IL, and WI.

consistent with the rough uniformity of average household sizes across different counties in the U.S.

The remaining rows of Table 19 show the estimated proportional changes (and standard errors) for the three different moments we calculate for the distribution of distances moved for the out- or in-migrants associated with each county. In these specifications, very little evidence is found for out-migrants to suggest that an additional tornado has any systematic effect on the mean, the standard deviation, or the skewness of the distribution of distances to which people move. For in-migrants, however, there is some suggestion that when there has been an additional tornado in the destination county, in-migrants are likely to arrive from relatively nearer origin counties and the standard deviation of arrival distances will be lower. However, these effects are statistically significant only in the case of in-migrants arriving in a county in one of the twenty most tornado-prone states from somewhere outside this domain.<sup>57</sup>

Table 20 shows results for a set of models analogous to those covered in Models 3 and 6 of Table 17 for each of the five outcome variables. The tornado variable employed as the key regressor in this table is “Any F5 tornado occurrence” even though the models for tornado occurrences of other intensity levels and the total number of occurrences in a county. These models show that there are many more instances of statistically significant coefficients on the impacts of the most severe tornado events. An F5 tornado in a county can lead to a considerably larger, roughly 3%, increase in out-migrants and decrease in-

---

<sup>57</sup> We exercise caution in noting these tendencies because differences in estimated effects across the different samples might pertain to the level of censoring of insignificant flow (i.e., the flows with less than ten tax returns). We cannot obtain accurate distances traveled for these flows because the migration data only reports a more spatially aggregated destination region (for out-migrants) and origin region (for in-migrants).

migrant returns. Further, an F5 occurrence positively affects the expected distance that individuals move (by an amount appearing to be between 6 and 11 miles), with an increase also in the dispersion of the distances traveled. The opposite appears to be true for in-migrants. An F5 occurrence in county results in a set of in-migrants who arrive from areas more heavily concentrated nearby (perhaps from other historically tornado prone areas, where familiarity with tornadoes makes the in-migrants less wary of the risks).

The results in Table 21 convey the dynamic relationship between county-level migration flows and tornado occurrences. Instead of reporting dynamic models with lagged effects (see Table G.2 in Appendix G), we report effects for an increase in the elapsed time since the most recent tornado event for all the outcome variables. Longer durations appear to reduce individuals' perceptions of tornado risks in those areas and thus the degree of migration. Specifically, the estimated effects suggest that a longer period of time since the last tornado decreases the number out-migrants leaving a county and increases the number-migrants choosing to live in that county. For out-migrants, there is some evidence that a greater time since the last tornado actually decreases the mean distance moved, but only for the subsample of migrants moving from the twenty most tornado-prone states to other states outside this area.<sup>58</sup> For in-migration, a longer vacation from tornados in the destination county appears to make household more likely to move to those destinations from farther away, with a corresponding increase in the

---

<sup>58</sup> For earthquakes, there might be a physical basis for the expectation that the longer a county has gone without an earthquake, the more likely an earthquake is to strike in the current period (i.e. positive duration dependence, or an increasing hazard rate). Arguments of a similar nature probably cannot be made for tornadoes.

standard deviation in the distances from which they arrive (i.e. they do not completely stop coming from closer origin counties). Our results for the third moment of the distance distributions (skewness) for both the out-migrant and the in-migrant samples are inconclusive.

### *Spatially Displaced Tornado Activity*

In Table 22, we model the effects of tornado activity in neighboring regions around origin (destination) counties for out-migrants (in-migrants). This allows us to determine if risk perceptions of tornado activity also varies with the “proximity” of other nearby tornado activity just outside the origin or destination county in the same year. Models 1 and Model 3 of Table 22 are the identical specifications as Models 3 and 6 in Table 13, respectively. Models 2 and 4 of Table 22 include are for out-migrants and in-migrants, respectively, and include the additional variables for tornado activity in the 0 to 20 mile buffer region around the county and tornado activity in the 20-50 mile buffer region. For out-migrants, these two variables are significant and negative, which suggests less out-migration for increasing levels of activity near the origin county. One possible explanation for these effects is that the frequency of activity adjacent to a county has its strongest influences on the out-migration to those affected areas. The magnitude of the effect appears to be diminishing with the distance away from the origin county. We also find that by accounting for nearby activity, the positive effect of own-county activity on out-migration increases by a factor of two. For in-migration, tornado activity near the

destination county and own-activity in the destination county are insignificant determinants of in-migration flows.

The negative effects on out-migration of tornado activity in regions near the origin county are suggestive that the role of distance between origin and destination counties may be an important determinant for migration that needs further exploration in our analysis. In particular, it appears that from Model 2 of Table 22 that the fact that migration is more frequent between nearby counties leads to observed negative effects of tornado activity near the origin county on out-migration. Future research will examine influence of “strong” local migration patterns on our results by replicating the models in Table 13 to only migration flows between counties that are relatively close to each other.

## **Discussion**

The full costs of climate change will depend on the capacity of individuals and countries to mitigate, or adapt to, the potential health and economic risks that may ensue from climate change. For domestic and international policies to address these risks, we need a much clearer understanding of individuals’ adaptive strategies in response to changes in extreme weather risks. Migration is one such form of adaptation. Those who migrate may be able to avoid adverse extreme weather impacts of climate change, while those who do not migrate may have to bear the costs of this extreme weather. The existing research in this area has tended to concentrate upon case studies involving specific extreme weather events. In contrast, our analysis uses tornadoes as a case where there are many events each year over a wide geographic area, and these events tend to be



spatially random, unpredictable, and “surgical” in the pattern of damages they cause. Our results characterize migration behavior as a function of perceived risk and potential changes in risk from extreme weather, assumed to be influenced by individuals’ recent experiences with nearby instances of such extreme weather. We find effects of tornado activity on migration behavior even after controlling for the many confounding factors that might also lead to household relocation. Specifically, we find evidence that the level of household and individual migration between US counties increases with the frequency of tornado occurrences in the origin county. Our strongest estimated effects are for out-migration. However, these migration effects appear to be relatively short-lived, lasting roughly a year. One drawback to our investigation is that we do not explicitly model the costs of migration. This prevents us from developing useful welfare measurements for the disutility of tornado occurrences based on (destination and origin) county-level variation in travel-costs. Instead, we can characterize the welfare loss in the U.S. based on average distance traveled and typical moving related costs. With roughly 1,500 tornado strikes in the US in 2005 and an average \$5,000 dollars spent by a migrant household in moving related costs that include the transportation costs of themselves and household belongings, and forgone work-wages, our estimated results suggest that Americans spend roughly \$22.5 million dollars a year on moving-related expenditures due to the presence of tornadoes.<sup>59</sup> Of course, this would only represent the total welfare loss of tornado occurrences under the simplifying assumptions that the only effect of tornadoes is a change in migration behavior and that non-movers do not incur disutility from tornado

---

<sup>59</sup> We assume a constant moving cost of \$5000 to move belongings and oneself to a new residence. When combined with our estimated effect of 3 migrant households per tornado occurrence, the total moving expenditures nationwide related to tornado activity is  $3 * 1500 * \$5000 = \$22.5$  million dollars.

events. Many other households are likely to experience disutility from increased tornado activity, but if this disutility may not be great enough to overcome the transactions cost of moving to another county with a lower perceived tornado risk. For established households, such moves would likely involve changes of job and schools and perhaps the need for real estate transactions and the associated commissions, along with the psychic costs of moving away from friends (and perhaps family) which would be experienced by everyone in the household. Clearly, the monetized value of disutility from increased tornado risks would have to exceed a substantial threshold before migration would be observed. The aggregate pecuniary costs of migration for households who actually move are very much just a lower-bound estimate for total welfare loss from such increases in tornado activity.

The National Center for Atmospheric Research (2001) reports that the estimated cost in property damages was \$1.1 billion in 2001 from tornadoes in the US. The probabilistic loss of approximately 80 lives per year, if evaluated at conventional U.S. EPA estimates of the so-called “value of a statistical life” of approximately \$7 million, suggests that society’s willingness to pay to avoid the loss of these 80 lives is about \$560 million per year. Willingness to pay to reduce the “over 1,500” tornado injuries each year would depend on the nature and severity of those injuries, but would also be substantial.

Thus, we find that migration-related expenditures increase the estimated welfare loss of tornado activity in the US by 10%. Our treatment of the migration related-expenditures as lower bound estimate could mean that the actually percentage increase is much larger when one also accounts for changes in the expected distance moved by all

migrants. As an alternative to quantifying the total welfare loss of the average tornado activity in the US, one could develop an estimate for the loss of welfare for particular forecast scenarios for changes in tornado activity due to climate change. We have not found any scientific studies that make these predictions. Instead, we consider the simple scenario in which 10% of the counties that have a least one tornado occurrence in a given year have at least one additional occurrence. Under this scenario, we find the estimated expected migration expenditures to be \$1.125 million (2005) dollars.

The annual frequency of our data does not allow us to accurately model dynamic temporal responses less than a year in length.<sup>60</sup> However, unlike the dynamic time response of health outcomes from extreme weather events, there is no reason to believe that the decision to move in response to the occurrence of one tornado should take place over an extended period of time. Conditional on other covariates related to tornado activity, and barring any capital or physical constraints to moving, the decision to move is likely be immediate although there may be delays in the execution of such a plan. The absence of any empirical evidence for migration reactions to lagged tornado events supports this theory. However, when examining migration behavior for areas that are most prone, we do find a recency effect on out-migrant behavior. Increases in duration since the last tornado event for a county decreases the out-migration response to contemporaneous activity. This evidence suggests that frequent, cumulative exposure may trigger larger migration responses via individuals' perceived risk.

---

<sup>60</sup> A potential mechanism by which one could analyze these shorter time periods with the same migration data would be to examine the change migrant status as function of the number of days before the (if one) tornado occurs before the last day of filing a tax return. A likely problem with this method would be to deal with tax filers who ask for an extension in their filing. Our data does not provide information of the number of requested extensions.

The exact relationship between tornado activity and a migration response appears also to involve individuals' spatial perceptions of risk and differences in the perceived changes in risk for different segments of society. We find that, on average, out-migration in response to a tornado event is less likely to occur when there is other tornado activity in regions outside, but in close proximity to, ones' own jurisdiction. This may be best explained by noting over a 5-year period, 90% of individuals in the US remain in the same state. Relocation frequently occurs between relatively close areas within the same county (i.e. as households move up or downsize their dwellings without necessarily changing jobs or schools). This would suggest that tornado occurrences in these areas would likely have the strongest effects on the number of out-migrants. Interestingly, we do not find the pattern of these spatial effects to hold for in-migrants. In addition, the negative effects on out-migration of tornado activity in regions near the origin county suggest that the role of distance between origin and destination counties may be an important determinant for migration that needs further exploration in our analysis.

Our findings also suggest that migration varies across different socioeconomic groups. This could be due to individuals having different perceptions of the economic and health risks or having different constraints and opportunities that we are unable to account for in our models. We find that the level of out-migration is responsive to increasing tornado activity, with the most prominent effects related to the education levels of origin and destination counties. Origin counties with higher levels of an educated populace are more likely to experience further amounts of out-migration for increasing levels of tornado activity in the origin county. For in-migrants, the education

effects on the level of migration from increasing tornado activity are less systematic across the levels of education but nonetheless suggest that education is an important mediating factor to the decision of in-migrants to move to a destination county when that county has had recent tornado occurrences.

The disaggregation of our average effects by the mix of household types in the origin or destination counties is important to the task of understanding the possible effects of income and/or risk preferences on migration behavior. Differences in the initial demographics of household types within a community may affect the types of households who leave and the types of household who replace them. For example, Myers, et al. (2008) examines post-disaster migration specifically in the wake of Hurricanes Katrina and Rita along the U.S. gulf coast. They find that the county-level net out-migration caused by these two hurricanes was significantly greater among groups with lower socioeconomic status, for areas which suffered greater property damage, and for areas which were originally more densely populated. Our results support the existence of these differences.

Importantly, our results are likely to have implications for human adaptation behavior in response to severe weather events in other countries besides the US. The southern Prairie Provinces of Canada suffer tornadoes as well, as does the southern portion of Ontario. Tornado activity is most common in the Northern and Southern hemispheres between the latitudes of 30° and 50°. Besides the US and Canada, specific regions that have regular activity include northern Europe, western Asia, Japan, China,

South Africa, Argentina, and Bangladesh.<sup>61</sup> There is also a growing literature on the political-economy consequences of changing international migration patterns and population density due to climate change related natural disasters. These consequences are most important to less-developed countries that lack the government and economic resources necessary to handle large changes in sub-populations of particular ethnic/cultural types and/or occupations/industries.

One of the next steps in our research agenda is to construct a range of reasonable forecast scenarios concerning potential risk changes for the frequency, severity, and spatial extent of tornadoes in the face of climate change. Simmons and Sutter (2007) find evidence that injuries and fatalities are greater when people are less well-prepared for unseasonable tornadoes. Furthermore, changes in the duration or timing of tornado “seasons” have been suggested as a potential consequence of climate change. Our analysis already captures some “atypicality” of tornadoes, to a degree, in that we estimate models using the years since the most recent tornado activity based on yearly aggregate level counts of tornadoes for each county. However, our data will also potentially permit us to study the unseasonableness of tornado activity (by season of year). There is also significant variation across geographical areas in the seasonal patterns of tornado occurrences. The data behind Figure 15 (state-by-state differences in seasonal tornado risks) capture expected seasonal tornado risks in different parts of the country. Individuals may respond less to tornadoes that are more typical than to tornadoes that have an atypical time of occurrence during the year.

---

<sup>61</sup> See <http://www.ncdc.noaa.gov/oa/climate/severeweather/tornadoes.html>.

One of our key interests is to determine whether changes in severe weather risks caused by climate change will result in significant socioeconomic disparities in migration as an adaptation mechanism. Although our study is limited to the sociodemographic and economic conditions of the continental United States, we are able to find effects for socioeconomic dimensions for which it is reasonable to expect migration disparities might manifest in other countries or regions of the world under similar weather calamities. We find that some groups in society may be more or less likely to respond to unseasonable tornadoes than are others. In particular, that the proportion of college educated, percent of individuals living in mobile housing, and the percent of individuals in industries that entail working outdoors for an origin county, and the educational mix, proportional of blacks, and the average household income of a destination county have some effect on the responsiveness of migration to tornado events in that origin or destination county.

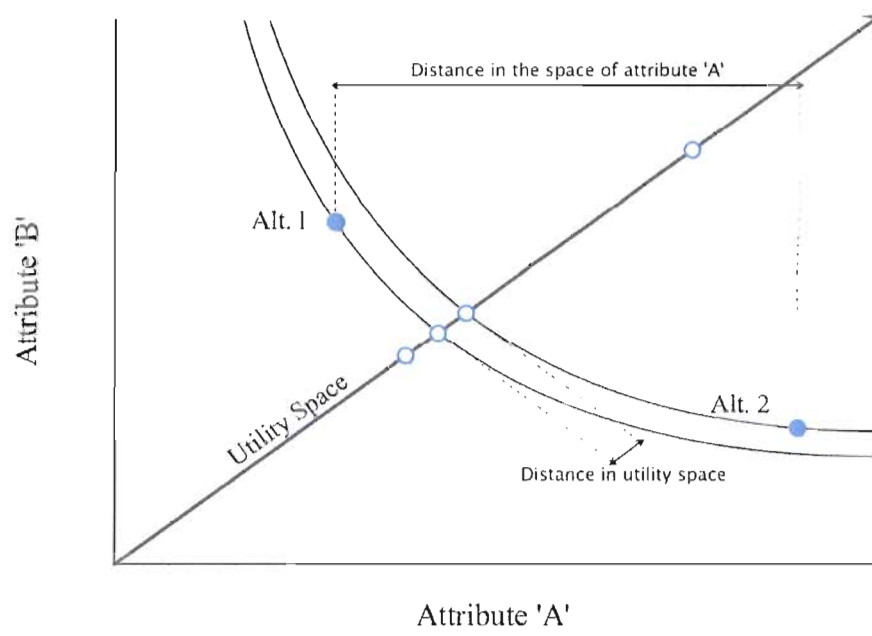
Over the longer term, we also expect to broaden the scope of our inquiry to address several other types of severe-weather-related hazards (e.g. floods, heat waves, wildfires, severe winter storms, etc.) that can be expected to alter their patterns due to climate change. The timing, geographic scope and severity of these events can in many cases be quantified in a fashion similar to our data on tornadoes. The extent to which we can apply our results to regions not at risk for tornadoes will depend, in part, on our ability to replicate these sorts of migration-related model for different types of hazards. However, as pointed out by Smith et al. (2006), it is difficult to draw transferable lessons from a single analysis of adjustment to one large disaster. It is first necessary to study

tornadoes separately because the mechanisms for adjustment and adaptation may be much different than for hurricanes and other disaster types.



APPENDIX A  
FIGURES AND TABLES

Figure 1: Attribute- versus Utility-Space Complexity



**Figure 2: Example of a Choice Scenario**

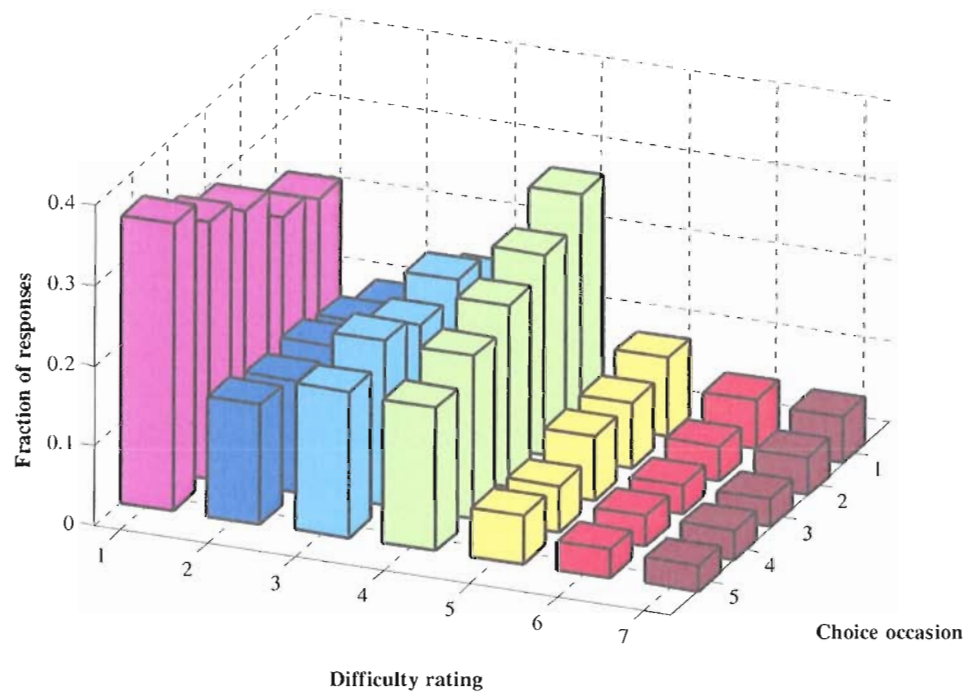
Choose the program that reduces the illness that you most want to avoid. But think carefully about whether the costs are too high for you. If both programs are too expensive, then choose Neither Program.

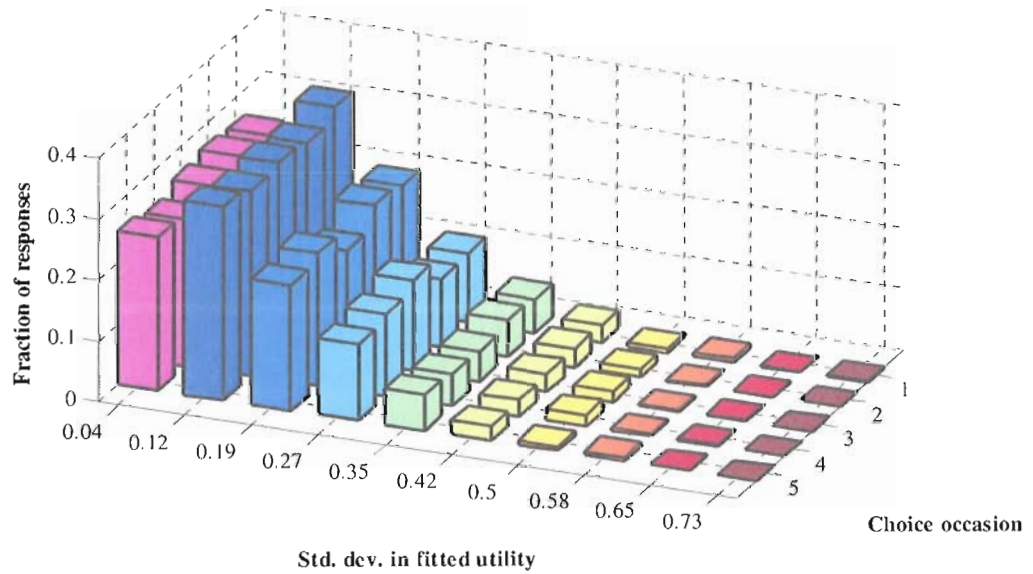
If you choose “neither program”, remember that you could die early from a number of causes, including the ones described below.

	<b>Program A for Diabetes</b>	<b>Program B for Heart Attack</b>
<b>Symptoms/ Treatment</b>	Get sick when 77 years-old 6 weeks of hospitalization No surgery Moderate pain for 7 years	Get sick when 67 years-old No hospitalization No surgery Severe pain for a few hours
<b>Recovery/ Life expectancy</b>	Do not recover Die at 84 instead of 88	Do not recover Die suddenly at 67 instead of 88
<b>Risk Reduction</b>	10% From 10 in 1,000 to 9 in 1,000	10% From 40 in 1,000 to 36 in 1,000
<b>Costs to you</b>	\$12 per month [ = \$144 per year]	\$17 per month [ = \$204 per year]
<b>Your choice</b>	<input checked="" type="checkbox"/> <b>Reduce my chance of diabetes</b>	<input type="checkbox"/> <b>Reduce my chance of heart attack</b>
	<input type="checkbox"/> <b>Neither Program</b>	

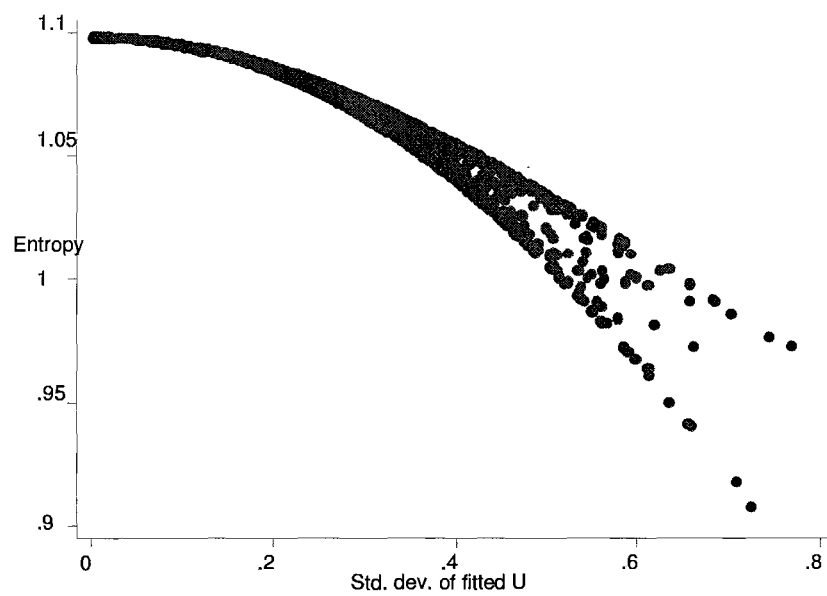


**Figure 4: Subjective Choice Difficulty Response Frequencies  
by Choice Occasion**



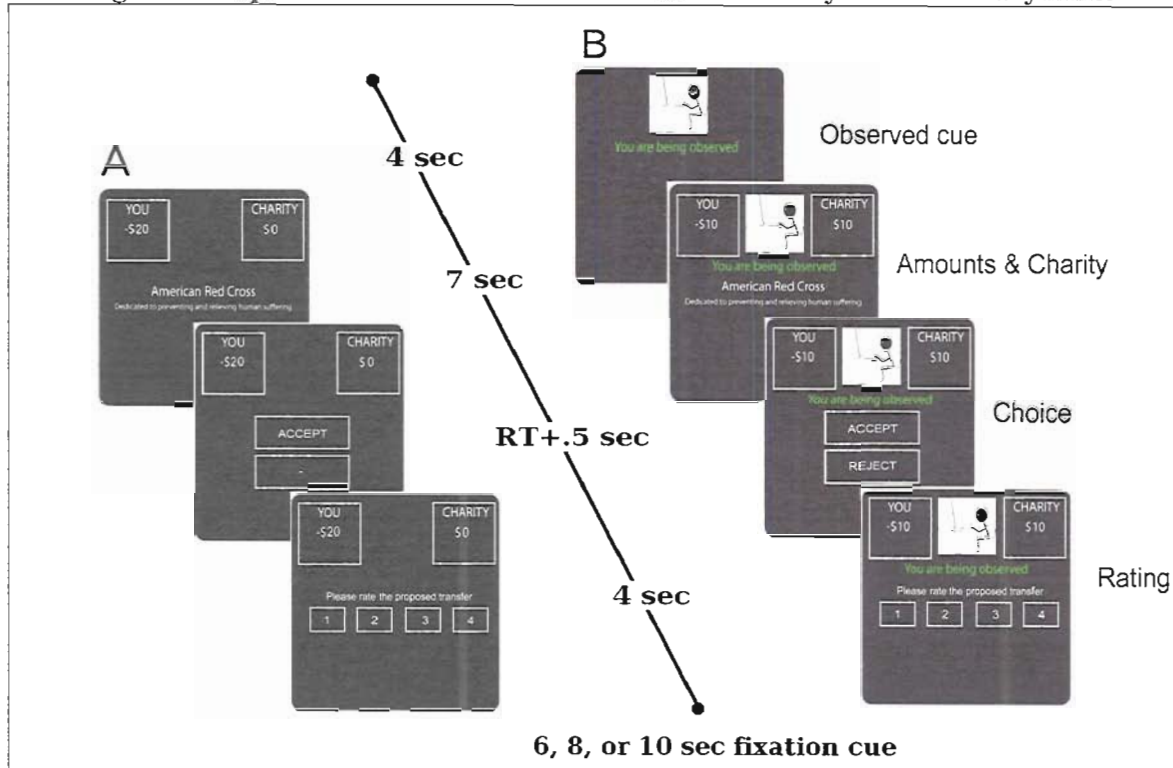
**Figure 5: Pattern of *Std. dev. of fitted U* by Choice Occasion**

**Figure 6: Relationship between *Entropy* and *Std. dev. of fitted U***





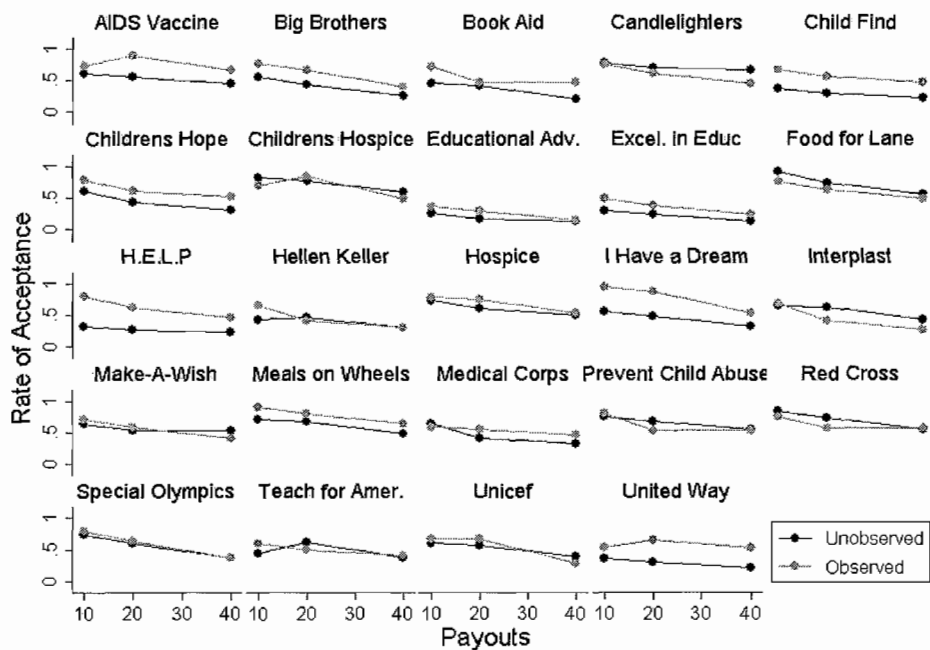
**Figure 8: Sequence of Screen Presentations in Mandatory and Voluntary Runs**



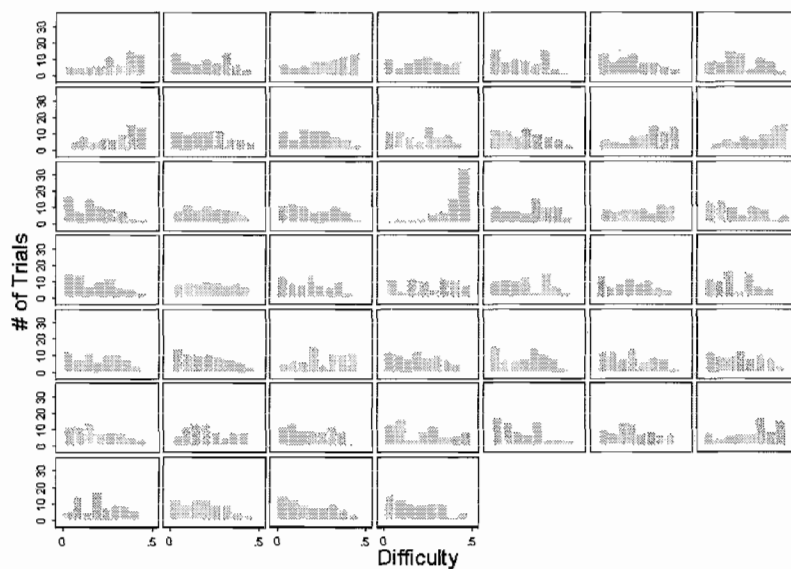
*Notes:* There are a total of three runs in the experiment. In the first run, participants must mandatorily accept the proposed transfers. For the next two runs, participants voluntarily choose to accept or reject a proposed transfer. Participants see three screens for the trials in the mandatory run (sequence A). This includes a screen for the presentation of information on the attributes of the choice (Amounts & Charity (7 seconds)), a screen for acknowledging the transfer (Choice (0-4 seconds)), and providing of transfer (Rating (0-4 seconds)). Prior to the beginning of the mandatory trials, participants are told that all trials are unobserved. For the voluntary trials (sequence B) there are four screens: Observed cue (4 seconds), Amounts & Charity (7 seconds), Choice (0-4 seconds), and Rating (0-4 seconds). The duration between all trials is randomly chosen to be 6, 8, or 10 seconds.



**Figure 9: Average Rate of Acceptance for Payout Conditions**

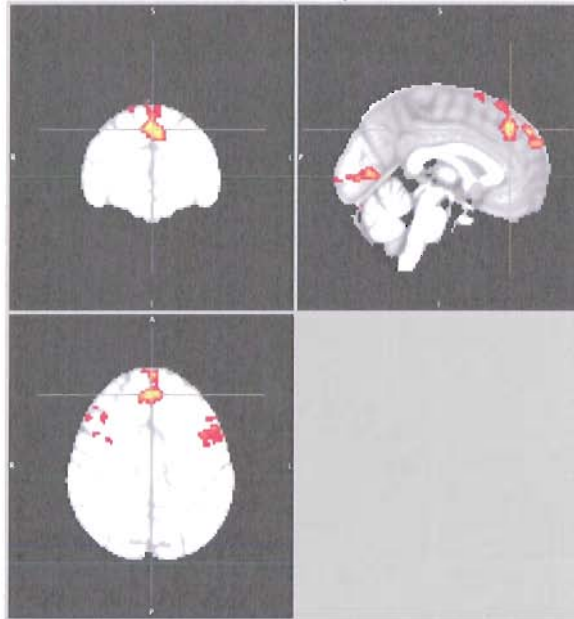


*Notes:* The rate of acceptance across all voluntary trials for a given charity is, on average, decreasing in payouts and increasing when the decision to accept or reject by participants is observed by the experimenter.

**Figure 10: Frequency Distribution of Fitted Choice Difficulty Trials**

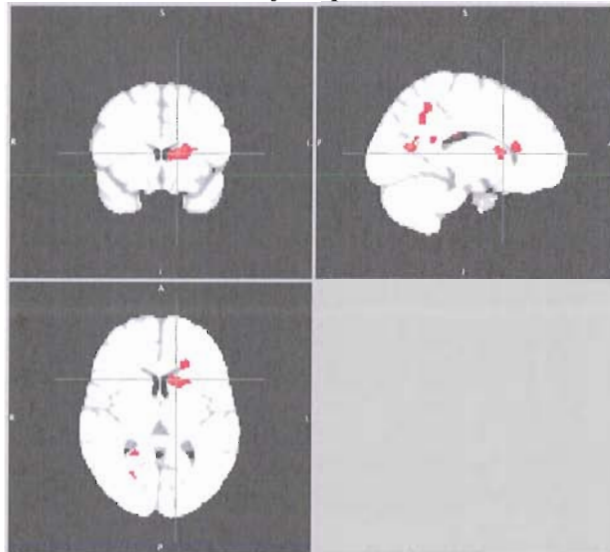
*Notes:* For each participant, the simple frequency distribution of the fitted choice difficulty of trials is roughly uniform. Each pane in the figure is the frequency distribution of trials over difficulty for a particular participant. By construction, the difficulty measure is bounded between 0 and .5.

**Figure 11: Neural Activation for Increasing Difficulty**



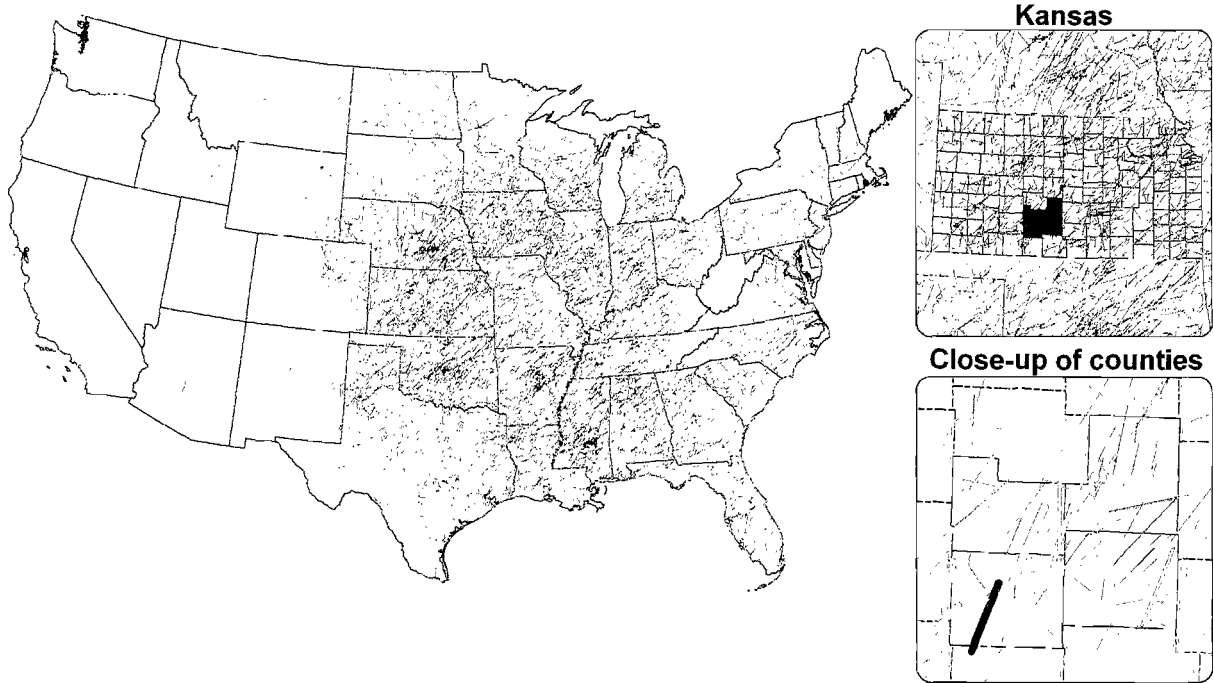
*Notes:* Neural activation for more difficulty decisions is increasing with increasing levels of age in the cerebral cortex. ( $x=0$ ,  $y=38$ ,  $z=40$ ;  $p=.00012$ ,  $\max Z=3.54$ ).

**Figure 12: Neural Activation for Increasing Charity Importance**



*Notes:* Neural activation for increasing charity importance is increasing in the left caudate ( $x=-16$ ,  $y=14$ ,  $z=48$ ;  $p=.0491$ ,  $\max Z=3.31$ ).

**Figure 13: Tornadoes - Conterminous U.S., Single State; Single County**  
**Distribution of Tornadoes, 1950-2008**

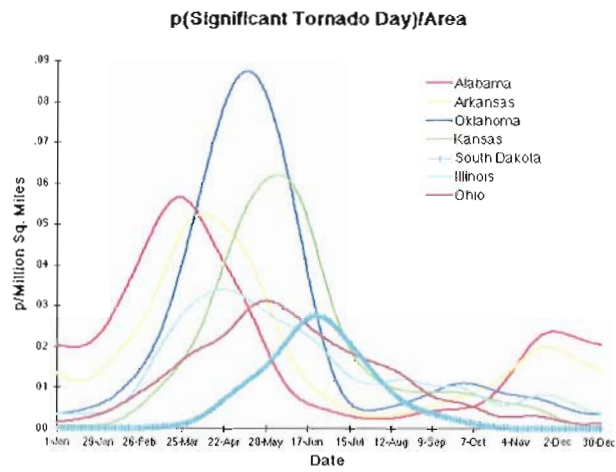


**Figure 14: Within- and Adjacent-to-county Tornado Risks**



*Notes:* Within- and adjacent-to-county tornado risks for a 20-mile buffer around an example county in Kansas. The darkest paths are those tornadoes since 1950 associated as a within-county event. The lightest paths are those associated tornado events for the buffer zone.

**Figure 15: Spatial Variation in Peak Hazard**



Notes: Data from the NOAA -  
[http://www.nssl.noaa.gov/users/brooks/public\\_html/tornado](http://www.nssl.noaa.gov/users/brooks/public_html/tornado)

**Table 1: Summary Statistics (n=22176)**

	Mean	Std. dev.	Min.	Max.
<b>Dependent variable</b>				
How difficult (1 very easy - 7 very hard)	2.87	1.69	1	7
<b>Measures of choice set complexity</b>				
<i>(a) In utility space:</i>				
Std. dev. of fitted U	0.18	0.08	8e-4	0.48
Entropy	1.09	0.01	1.03	1.1
<i>(b) In attribute space:</i>				
<i>(1) Within-alternative attrib. variability (across alts.):</i>				
Mean std. dev.	1.12	0.11	0.82	1.31
Disp. of std. dev.	0.33	0.08	0.07	0.58
<i>(2) Across-alternative attrib. variability (Ad hoc)</i>				
Std. dev. of montly costs	0.81	0.62	0.06	2.87
" risk difference	1.1	0.34	0.47	1.53
" latency	0.96	0.46	0.09	2.42
" years sick	0.82	0.66	0	4.28
" lost life years	0.85	0.6	0	3.19
<i>(3) Across-alternative attrib. variability (Structural)</i>				
Std. dev. of linear net income term	0.85	0.59	5e-3	2.95
" quadratic income term	0.69	0.69	0.01	6.92
" $\Delta\Pi_i^{AS} \log(pdv_i^A + 1)$	0.89	0.59	0	3.16
" $\Delta\Pi_i^{AS} \log(pdv_r_i^A + 1)$	0.52	0.87	0	6.18
" $\Delta\Pi_i^{AS} \log(pdvl_i^A + 1)$	0.89	0.61	0	3.19
<b>Observable proxies of sociodemographics</b>				
Income (in \$1000)	50.27	33.5	5	150
Age	50.7	15.2	25	93
1(Female)	0.52	-	0	1
1(Divorced)	0.11	-	0	1
1(Black)	0.09	-	0	1
1(Other ethnicity)	0.04	-	0	1
1(Hispanic)	0.06	-	0	1
Household size	2.57	1.26	1	8
# of kids	0.52	0.95	0	5
1(Dual income household)	0.65	-	0	1
1(Single parent)	0.02	-	0	1
<b>Observable proxies measures of cognitive capacity</b>				
1(Less than high school)	0.11	-	0	1
1(High school degree)	0.34	-	0	1
Avg. duration on other choice occasions	45.97	26.4	0	202
1(Valid duration)	0.99	-	0	1
<b>Attention behavior controls</b>				
1(All status quo)	0.15	-	0	1



**Table 1 (continued):**

	Mean	Std. dev.	Min.	Max.
1(No change in difficulty rating)	0.21	-	0	1
<b>Survey-specific health characteristics</b>				
Illness experience count (0-13):	9.08	3.78	0	13
Avg. subj. risk of future experience (0-4):	-0.24	0.86	-2	2
Subjective controllability or risks (0-4):	-0.3	1.02	-2	2
1(Missing health)	0.09	-	0	1

**Table 2: Invariant Difficulty Ratings**

Variable: <i>How difficult</i> (rating)	Total # of responses	# with no change in rating	% with no change
1	8310	3414	41.08%
2	3597	342	9.51%
3	4800	480	10.00%
4	5793	963	16.62%
5	1845	93	5.04%
6	1050	90	8.57%
7	1056	258	24.43%
Total	26451	5640	21.32%

**Table 3: Simple Preliminary Conditional Logit Models**

COEFFICIENT	Model 1	Model 2
<b>Ad hoc attributes:</b>		
Annualized costs	-0.007*** (-9.29)	-
Risk difference	-50.920*** (-4.40)	-
Latency	0.002 (1.30)	-
Years sick	0.009*** (3.92)	-
Unexpected lost life years	0.012*** (7.27)	-
<b>Structural attributes:</b>		
Linear net income term	-	5.355*** (9.19)
Quadratic net income term	-	-2.193*** (-4.68)
$\Delta\Pi_i^{AS} \log(pdv_i^A + 1)$	-	-24.793*** (-4.23)
$\Delta\Pi_i^{AS} \log(pdv_r^A + 1)$	-	-22.166** (-2.37)
$\Delta\Pi_i^{AS} \log(pdvl_i^A + 1)$	-	-30.717*** (-6.02)
Observations	22485	22485
LogL	-11662.73	-11687.13

*Notes:* Conditional logit models are for three-way choices between Program A, Program B, and Neither Program (N). z statistics in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 4: Utility-Space Determinants of Choice Difficulty**

COEFFICIENT	Model 1 Ordered Probit	Model 2 Ordered Probit	Model 3 Ordered Probit
<b>Measures of choice set complexity:</b>			
<i>(a) In utility space:</i>			
Std. dev. of fitted U	-0.353** (-2.40)	-	0.355 (0.61)
Entropy	-	3.094*** (2.64)	5.831 (1.26)
Number of choice sets	7392	7392	7392
Number of respondents	1694	1694	1694
Maximized LogL	-12851.81	-12851.20	-12851.01

*Notes:* The intercept is normalized to zero for ordered probit models. Incidental threshold parameter estimates are not reported. The sample size of all three models has been reduced somewhat to match smaller sample with complete data for all of the variables in the models in Table 6.

**Table 5: Utility-Space and Attribute-Space Determinants of Choice Difficulty**

COEFFICIENT	Difficulty			
	Model 2a Ordered Probit	Model 2b Ordered Probit	Model 2c Ordered Probit	Model 2d Ordered Probit
<b>Measures of choice set complexity:</b>				
<i>(a) In utility space:</i>				
Entropy	3.202*** (2.73)	3.139*** (2.68)	3.202*** (2.73)	2.948** (2.50)
<i>(b) In attribute space:</i>				
<i>(1) Distrib. of within-alt. attrib. variability</i>				
Mean std. dev.	-0.304*** (-2.69)	-	-0.304* (-1.81)	5.068** (1.97)
(Mean std. dev.) <sup>2</sup>	-	-	-	-2.401** (-2.09)
Disp. of std. dev.	-	0.322** (1.99)	-0.002 (-0.01)	-
Number of choice sets	7392	7392	7392	7392
Number of respondents	1694	1694	1694	1694
Maximized LogL	-12847.58	-12849.21	-12847.58	-12845.39

*Notes:* The intercept is normalized to zero for ordered probit models. Incidental threshold parameter estimates are not reported. The sample size of all three models has been reduced somewhat to match smaller sample with complete data for all of the variables in the models in Table 6.

**Table 6: Additional Determinants of Subjective Choice Difficulty**

COEFFICIENT	Model 4 Ordered Probit	Model 5 Ordered Probit	Model 6 LS Fixed Effects
<b>Measures of choice set complexity:</b>			
<i>(a) In utility space:</i>			
Entropy	2.844** (1.99)	4.077*** (2.58)	4.542** (2.42)
<i>(b) In attribute space:</i>			
<i>(1) Distribution of within-alt. attrib. variability</i>			
Mean std. dev.	4.842* (1.87)	4.873* (1.86)	6.822** (2.24)
(Mean of std. dev.) <sup>2</sup>	-2.294** (-1.99)	-2.247* (-1.93)	-3.111** (-2.29)
<i>(2) Across-alt. attrib. variability (Ad hoc model)</i>			
Std. dev. of annualized costs	0.028 (1.00)	0.092*** (2.98)	0.073* (1.89)
" risk difference	-0.051 (-1.35)	-0.063 (-1.64)	0.053 (1.19)
" latency	0.034 (1.07)	-0.080 (-1.59)	-0.061 (-1.04)
" years sick	0.071*** (3.64)	0.032 (1.42)	0.045* (1.72)
" lost life years	-0.015 (-0.57)	-0.051 (-1.62)	-0.029 (-0.78)
<i>(3) Choice occasion indicators</i>			
2nd choice occasion	-	-0.148*** (-3.89)	-0.332*** (-8.30)
3rd "	-	-0.283*** (-7.35)	-0.342*** (-8.40)
4th "	-	-0.350*** (-9.00)	-0.372*** (-8.98)
5th "	-	-0.435*** (-11.13)	-0.479*** (-11.53)
<b>Obj. measures of cognitive capacity:</b>			
1(Less than high school)	-	0.128*** (2.89)	n/a
1(High school degree)	-	0.060** (2.07)	n/a
Avg. duration on other choice occasions	-	0.002*** (4.39)	-0.033*** (-16.78)
1(Valid duration)	-	-0.504*** (-3.62)	-1.481*** (-8.14)
<b>Obj. measures of sociodemographics:</b>			
Income (in \$1000)	-	0.001** (2.01)	n/a

**Table 6 (continued):**

COEFFICIENT	Model 4 Ordered Probit	Model 5 Ordered Probit	Model 6 LS Fixed Effects
Age	-	-0.018*** (-2.68)	n/a
Age <sup>2</sup>	-	0.009 (1.56)	n/a
1(Female)	-	0.046* (1.85)	n/a
1(Divorced)	-	-0.024 (-0.56)	n/a
1(Black)	-	-0.185*** (-4.05)	n/a
1(Other ethnicity)	-	-0.004 (-0.06)	n/a
1(Hispanic)	-	0.093* (1.74)	n/a
Household size	-	0.023 (1.17)	n/a
# of children	-	-0.055** (-2.26)	n/a
1(Dual income household)	-	-0.050* (-1.65)	n/a
1(Single parent)	-	-0.048 (-0.46)	n/a
<b>Survey-specific respondent characteristics:</b>			
Illness experience count (0-13):	-	-0.010*** (-2.83)	n/a
Avg. subj. risk of future experience (0-4):	-	0.091*** (5.65)	n/a
Subjective controllability of risks (0-4):	-	0.046*** (3.44)	n/a
1(Missing health)	-	0.024 (0.52)	n/a
<b>Respondent attention behavior:</b>			
1(All status quo)	-	-0.068* (-1.72)	n/a
1(No change in difficulty rating)	-	-0.632*** (-18.70)	n/a
Constant <sup>b</sup>	-	-	-2.438 (-0.96)

**Table 6 (continued):**

COEFFICIENT	Model 4 Ordered Probit	Model 5 Ordered Probit	Model 6 LS Fixed Effects
Number of choices	7392	7392	7392
Number of respondents	1694	1694	1694
Maximized LogL	-12463.901	-10115.086	-10116.586

*Notes:* All models are pooled ordered-probit except for Model 6, which uses a linear fixed-effects estimator and clusters on errors at the respondent level. Robust t-statistics in parentheses; \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . The intercept is normalized to zero for ordered probit models. Incidental threshold parameter estimates are not reported.



**Table 7: Comparison of Ad hoc and Structural Specifications**

COEFFICIENT	Model 6 LS Fixed Effects	Model 7 LS Fixed Effects
<b>Measures of choice set complexity:</b>		
<i>(a) In utility space:</i>		
Entropy	4.542** (2.42)	2.729*** (2.58)
<i>(b) In attribute space:</i>		
<i>(1) Distribution of within-alt. attrib. variability</i>		
Mean std. dev.	6.822** (2.24)	7.476** (2.47)
(Mean std. dev.) <sup>2</sup>	-3.111** (-2.29)	-3.403** (-2.52)
<i>(2) Across-alt. attrib. variability (Ad hoc model)</i>		
Std. dev. of annualized costs	0.073* (1.89)	-
" risk difference	0.053 (1.19)	-
" latency	-0.061 (-1.04)	-
" years sick	0.045* (1.72)	-
" lost life years	-0.029	-
<i>(3) Across-alt. attrib. variability. (Structural model)</i>		
Std. dev. of linear net income term	-	0.055 (0.99)
" quadratic net income term	-	-0.001 (-0.02)
" $\Delta\Pi_i^{AS} \log(pdv_i^A + 1)$	-	0.072*** (2.77)
" $\Delta\Pi_i^{AS} \log(pdvr_i^A + 1)$	-	0.016 (0.95)
" $\Delta\Pi_i^{AS} \log(pdvl_i^A + 1)$	-	-0.009 (-0.34)
<b>Choice occasion indicators</b>	Yes	Yes
<b>Observable proxies for cognitive capacity</b>	"	Yes <sup>b</sup>
<b>Obs. measures of sociodemographics</b>	"	n/a
<b>Survey-specific health characteristics</b>	"	n/a
<b>Attention behavior controls</b>	"	n/a
Number of choices	7392	7392
Number of respondents	1694	1694
Maximized LogL	-10115.086	-10116.586

*Notes:* All models use a linear fixed-effects estimator and cluster on errors at the respondent level. Robust t-statistics in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. To minimize endogeneity, the average duration is for *other* choice occasions only.

**Table 8: The Set of Charitable Organizations**

<b>Charitable organization</b>	<b>Brief description</b>	<b>"Importance of cause" rating</b>
Children's Hospice International	Provides support for children with life-threatening illnesses.	-1.556 (least important)
Hospice Foundation of America	Dedicated to providing leadership in hospice and its philosophy of care.	-0.888
Educational Advancement Fund International	Provides students from rural China with experiences that build character.	-0.316
Prevent Child Abuse America	Dedicated to valuing children and strengthening families.	-0.314
Teach for America	Dedicated to eliminated educational inequality in America.	-0.273
Center for Excellence in Education	Dedicated to careers in science and technology for promising students.	-0.206
Hellen Keller International	Dedicated to combating causes of blindness and malnutrition.	-0.133
International Medical Corps	Dedicated to providing health care training and relief programs.	-0.116
Meals on Wheels of America	Dedicated to nutritional and economic betterment of Americans.	-0.097
Big Brothers Big Sisters of America	Dedicated to helping children through relationships with mentors.	-0.058
Children's Hope International Foundation	Dedicated to giving physical, medical, and humanitarian help to orphans.	-0.052
H.E.L.P. International	Provides health, education, and literacy programs worldwide.	0.055
Child Find of America	Dedicated to the prevention and resolution of child abduction.	0.061
Food for Lane County	Dedicated to eliminating hunger locally by creating access to food.	0.067
Unicef	Dedicated to establishing children's rights and strong principles.	0.159
American Red Cross	Dedicated to preventing and relieving human suffering of all forms.	0.172
Make-A-Wish Foundation	Dedicated to creating hope for children with life-threatening illnesses.	0.209
Interplast	Provides free surgery for those with clefts and disabling burns.	0.216
Candlelighters Childhood Cancer Foundation	Provides support for children with cancer and their families.	0.343
Book Aid International	Dedicated to promoting literacy in developing countries.	0.345
I Have A Dream Foundation	Dedicated to helping low-income children reach their life goals.	0.528
United Way	Dedicated to improving lives by mobilizing the power of communities.	0.530
Special Olympics	Dedicated to empowering people with disabilities to physical fitness.	0.605
International AIDS Vaccine Initiative	Dedicated to developing safe, effective, and accessible HIV vaccines.	0.719 (most important)

*Notes:* Each participant rates the importance of the cause that a charity serves from unimportant (1) to extremely important (10). These charity ratings shown above are standardized (z-scored) by participant and then averaged across participants

**Table 9: Estimated Marginal Effects on the Propensity to Give**

	RE Probit (1)	RE Probit (2)	RE Probit (3)
Proposed transfer (\$10, 20, 40)	-0.031*** (0.000)	-0.031*** (0.000)	-0.031*** (0.000)
Age	0.031** (0.010)	0.044*** (0.006)	0.028** (0.019)
Observed	0.412*** (0.000)	1.012*** (0.001)	0.829*** (0.000)
Importance rating (1-10) ^	0.062*** (0.001)	0.077*** (0.002)	0.060*** (0.001)
Importance * Observed	-	-0.025 (0.381)	-
Avg. rating for self payment (1-4) ^^	-	-0.418* (0.069)	-0.395* (0.096)
Avg. rating for charity payment (1-4) ^^^	-	0.528** (0.039)	0.473* (0.074)
Female	-	0.910 (0.365)	-
Age * Female	-	-0.030 (0.178)	-
Age * Observed	-	-0.009** (0.043)	-0.009** (0.038)
Constant	-0.755 (0.195)	-1.700 (0.137)	-1.024 (0.322)
Charity-specific indicators	Yes	Yes	Yes
Observations	3312	3312	3312
Number of subjects	44	44	44
Log-likelihood	-1482.6	-1476.0	-1478.3

*Notes:* A columns model the decision to accept (=1) the proposed monetary transfer to the charity with a random effects binary probit specification and include charity-specific indicators. (A Hausman-test failed to reject the null of no significant difference between random-effects and fixed-effects logit specifications with covariates specified as those in Model 3.) Coefficients represent the increases in the latent propensity to give for an increase in the independent variable. P-values are in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. ^ Participant ratings for "... rate the importance of the cause this charity serves" (1 – unimportant, 10 – extremely important). ^^ The average rating for self reward (by subject; 1 – extremely dissatisfied, 2 – extremely satisfied) during the the mandatory portion of the experiment. ^^^ The average rating for a charity receiving a payment (1 – unimportant, 10 – extremely important) during the mandatory portion of the experiment.

**Table 10: The Marginal Effects on the Probability of Giving**

Proposed transfer (\$10, 20, 40)	-0.031*** (0.000)
Age	0.023** (0.046)
Observed	0.415*** (0.000)
Importance rating (1-10) ^	0.060*** (.001)
Avg. rating for self reward (1-4)	-.396* (0.095)
Avg. rating for charity reward (1-4)	.473 (0.074)

*Notes:* Estimated effects are based on Model 3 of Table 2A. Estimated coefficients are average partial effects and represent the average increase in the probability of giving for an increase in the independent variable. P-values are in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

**Table 11: Peak Voxels for the D+A+ Contrasted Image**

<b>Region</b>	<b>Side</b>	<b>MNI - x</b>	<b>MNI - y</b>	<b>MNI - z</b>	<b>P-value</b>	<b>Max Z</b>
Anterior cingulate cortex, paracingulate gyrus	B	0	38	40	0.000123	3.54
Inferior frontal gyrus	R	54	22	6	6.23E-05	3.36
Middle frontal gyrus	L	-54	18	30	5.9E-06	3.72
Cerebellum	R	10	-78	-32	7.75E-07	3.41

*Notes:* Coordinates for peak voxels are in Montreal Neurological Institute (MNI-152) space. Peak voxels are in areas that have increasing activity for increasing difficulty and age.

**Table 12: Peak Voxels for the C+ Contrasted Image**

<b>Region</b>	<b>Side</b>	<b>MNI - x</b>	<b>MNI - y</b>	<b>MNI - z</b>	<b>P-value</b>	<b>Max Z</b>
Doral striatum, nucleus caudate	L	-16	14	48	0.0491	3.31
Occipital lobe, supracalcarine cortex	L	-16	-66	14	0.000386	3.19

*Notes:* Coordinates for peak voxels are in Montreal Neurological Institute (MNI-152) space. Peak voxels are in areas that have increasing activity for a positive trend of charity importance.

**Table 13: Dependent Variables: Migration Response Variables, 1992-2005**

	Total Obs.	Mean	Std. Deviation	Min	Max
<b><u>In-migrants</u></b>					
Total Migrant Tax Returns	42,896	2,143.14	5,820.84	13.00	155,822
Total Exemptions for Migrants	42,896	4,107.73	11,152.72	19.00	326,847
(Weighted) Mean Distance Traveled*	41,891	52.31	76.49	0.10	840.27
Std. Deviation of Distance*	41,891	96.75	141.35	0.23	900.06
Skewness of Distance*	41,891	3.59	2.00	0.45	16.08
County population**	42,588	88,342	289,782	47	9,847,753
<b><u>Out-migrants</u></b>					
Total Tax Returns	42,809	2150.89	5,419.34	11.00	99,392
Total Exemptions	42,809	4,125.02	9,800.89	28.00	168,717
(Weighted) Mean Distance Traveled*	41,239	50.77	79.71	0.06	925.76
Std. Deviation of Distance*	41,239	93.65	136.86	0.19	1,010.07
Skewness of Distance*	41,239	3.66	2.26	0.61	19.76
County population**	42,588	88,342	289,782	47	9,847,753

*Notes:* Data constructed from the IRS County-to-County Migration Data. Reported means with standard deviations in parentheses. In and out-migration *Exemptions* is based on the reported exemption on household tax returns and provides an approximate estimate of the number of individuals that migrate by household. *Mean of Distance* is a group-weighted average based on the number of tax returns between each pair of counties. \*Subject to censoring when less than 10 migrant tax returns between county i and county j. Some of empirical specifications are per capita models for Total Tax Returns and Total Exemptions.

**Table 14: Tornado Activity Variables, Presidential Disaster Indicators, and Socioeconomic Characteristics; Summarized for all U.S. Counties, 1992-2005**

	All counties		Counties with 0 tornado events (in a given year)		Counties with 1 or more tornado events (in a given year)	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
	<b>Tornado Activity</b>					
# of tornadoes	0.5	(1.1)	0	(0.0)	1.8	(1.4)
Years since previous tornado	5.8	(7.3)	6.5	(7.9)	3.9	(4.9)
Avg. # of tornadoes pre-1985	0.3	(0.3)	0.2	(0.2)	0.4	(0.4)
# of tornadoes in 0-20 mile buffer	2.1	(2.9)	1.4	(2.0)	3.9	(3.8)
# of tornadoes in 20-50 mile buffer	5.6	(5.8)	4.5	(4.9)	8.6	(6.9)
Average F-scale intensity	0.2	(0.5)	-	-	0.6	(0.8)
Average length of tornado path (in miles)	1.1	(5.5)	-	-	4.1	(10.1)
Average width of tornado path (in yards)	28.0	(106.6)	-	-	105.2	(185.7)
Total # of injuries	0.7	(9.3)	-	-	2.5	(17.8)
Total # of fatalities	0.0	(0.7)	-	-	0.1	(1.3)
<b>Socioeconomic Characteristics</b>						
County population (per 1,000)	91.0	(294.7)	84.7	(247.8)	108.4	(395.6)
Unemployment level (per 1,000)	2.5	(9.3)	2.3	(7.5)	2.9	(13.0)
% age 17-	26.0	(3.4)	26.0	(3.4)	26.2	(3.3)
% age 18-24	8.9	(3.4)	8.9	(3.5)	9.0	(3.4)
% age 25-64	50.4	(3.5)	50.5	(3.5)	50.0	(3.3)
% age 65+	14.9	(4.2)	14.8	(4.1)	15.1	(4.5)
% w/ no h.s.	10.3	(5.9)	10.3	(6.1)	10.1	(5.5)
% w/ some h.s.	14.6	(4.9)	14.7	(4.9)	14.4	(4.9)
% w/ h.s. diploma	36.9	(6.8)	37.1	(6.9)	36.5	(6.4)
% w/ some college	21.1	(5.0)	20.8	(5.0)	21.7	(4.7)
% w/ college degree	17.1	(8.1)	17.0	(8.3)	17.2	(7.6)
% Black	9.5	(15.0)	9.3	(15.0)	9.9	(14.8)
% Hispanic	5.9	(11.9)	5.7	(12.0)	6.3	(11.8)
% White	85.5	(16.1)	85.8	(16.1)	84.8	(15.8)
Avg. Hh. income (per \$1,000)	42.7	(11.4)	42.7	(11.6)	42.7	(10.7)
House value (per \$1,000)	43.5	(28.2)	44.5	(28.7)	40.9	(26.3)
% of households in poverty	14.4	(6.8)	14.5	(6.9)	14.3	(6.6)
% occupied housing	86.1	(9.2)	85.8	(9.7)	86.9	(7.9)
% renter occupied housing	2.5	(1.8)	2.5	(1.8)	2.4	(1.7)
% mobile occupied housing	14.6	(9.6)	14.7	(9.7)	14.2	(9.2)
% farm, fish, or forest workers	5.7	(6.3)	5.6	(6.2)	6.1	(6.5)
% workers w/ residence in county	67.9	(17.8)	67.2	(17.8)	70.0	(17.5)
<b>Presidential Disaster Declarations (=1)</b>						
Flood	0.05	-	0.04	-	0.06	-
Snow	0.02	-	0.02	-	0.02	-
Hurricane	0.05	-	0.04	-	0.08	-
Earthquake	0.00	-	0.00	-	0.00	-
Fire	0.01	-	0.01	-	0.01	-
Ice storm	0.01	-	0.01	-	0.02	-
Severe weather storm	0.14	-	0.13	-	0.18	-
Other	0.00	-	0.00	-	0.01	-
Observations		39,270		28,802		10,468

**Table 14 (continued):**

*Notes:* Sample of observations is restricted to those counties that have had at least one tornado during since 1950. Less than 0.6% of all counties have not had at least one tornado during this period.



**Table 15: County-Level Presidential Disaster Declarations,  
by Incident Type and General Category, 1964-2008**

<u>Category of disaster</u>	<u>Incident type</u>	<u># of declarations by incident</u>	<u># of declarations by category</u>
Storm	<i>Severe Storm(s)</i>	10,347	11,746
	<i>Tornado</i>	1,237	
	<i>Coastal Storm</i>	162	
Flood	<i>Flood</i>	7,507	7,512
	<i>Mud/Landslide</i>	3	
	<i>Dam/Levee Break</i>	2	
Hurricane	<i>Hurricane</i>	3,206	3,315
	<i>Typhoon</i>	109	
Ice	<i>Snow</i>	1,343	2,093
	<i>Severe Ice Storm</i>	666	
	<i>Freezing</i>	84	
Drought	<i>Fire</i>	701	785
	<i>Drought</i>	84	
Seismic	<i>Earthquake</i>	83	131
	<i>Volcano</i>	48	
Other	<i>Fishing Losses</i>	34	38
	<i>Human Cause</i>	2	
	<i>Terrorist</i>	1	
	<i>Toxic Substances</i>	1	
Total declarations		25,620	25,620

*Notes:* Data from the Federal Emergency Management Agency (FEMA). Prior to December 24, 1964, 179 declarations did not have county designations.

**Table 16: The Effects of Tornado Occurrences on County-level Migrant Household Tax Return Flows**

	Out-migration Flows			In-migration Flows		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
# of tornado occurrences	0.00130** (0.000611)	0.00101* (0.000575)	0.000979* (0.000574)	-0.000240 (0.000689)	-0.000280 (0.000625)	-0.000297 (0.000625)
Log population	0.973*** (0.0380)	1.067*** (0.0435)	1.067*** (0.0435)	1.033*** (0.0526)	1.153*** (0.0575)	1.153*** (0.0574)
Log unemployment	0.0587*** (0.00936)	0.0648*** (0.00827)	0.0653*** (0.00827)	-0.0497*** (0.0100)	-0.0469*** (0.00905)	-0.0466*** (0.00905)
Constant	-3.912*** (0.431)	-4.801*** (0.511)	-4.810*** (0.510)	-3.679*** (0.622)	-4.510*** (0.697)	-4.510*** (0.696)
County demographics	No	Yes	Yes	No	Yes	Yes
Controls for disaster decls.	No	No	Yes	No	No	Yes
County fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
State-year fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	41089	41089	41089	40999	40999	40999
R-squared	0.560	0.588	0.589	0.446	0.482	0.482
Number of counties	2978	2978	2978	2978	2978	2978

*Notes:* Dependent variable is log of total (out- or in-) migrant tax returns for a given county-year. Coefficients represent the fractional change in migrant household tax returns for a one unit change in the independent variable. The explanatory variables describe the county of origin for out-migration flows and the destination county for in-migration flows. Standard errors clustered by county are reported in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. All models include county-specific and state-year-specific fixed effects. Sample of observations includes all counties with a population less than 1 million residents. Observations are weighted by average county population over the sample time period from 1992 to 2005.

**Table 17: The Effects of Tornado Intensity on County-level Migrant Household Tax Return Flows**

	Out-migration Flows			In-migration Flows		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
# of tornado occurrences	0.0010*	0.0009*	0.0010*	0.0001	0.0000	0.0005
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Any F3, F4, or F5 (=1)	-	0.0030	-	-	0.0025	-
		(0.003)			(0.004)	
Any F0 (=1)	-	-	0.0008	-	-	-0.0018
			(0.002)			(0.002)
Any F1 (=1)	-	-	-0.0019	-	-	-0.0018
			(0.002)			(0.002)
Any F2 (=1)	-	-	0.0019	-	-	-0.0038
			(0.002)			(0.003)
Any F3 (=1)	-	-	0.0023	-	-	0.0009
			(0.003)			(0.004)
Any F4 (=1)	-	-	-0.0033	-	-	0.0029
			(0.005)			(0.007)
Any F5 (=1)	-	-	0.0354	-	-	-0.0078
			(0.022)			(0.009)
Constant	-4.7945***	-4.7951***	-4.7944***	-4.4998***	-4.5003***	-4.4921***
	(0.509)	(0.508)	(0.508)	(0.695)	(0.695)	(0.694)
<i>Log pop. &amp; Log unemp.</i>	Yes	Yes	Yes	Yes	Yes	Yes
County demographics	Yes	Yes	Yes	Yes	Yes	Yes
Controls for disaster decls.	Yes	Yes	Yes	Yes	Yes	Yes
County fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
State-year fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	41,089	41,089	41,089	40,999	40,999	40,999
R-squared	0.589	0.589	0.589	0.482	0.482	0.482
Number of counties	2,978	2,978	2,978	2,978	2,978	2,978

*Notes:* Dependent variable is log of total (out- or in-) migrant tax returns for a given county-year. Coefficients represent the fractional change in migrant household tax returns for a one unit change in the independent variable. The explanatory variables describe the county of origin for out-migration flows and the destination county for in-migration flows. Standard errors clustered by county are reported in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. All models include county-specific and state-year-specific fixed effects. Sample of observations includes all counties with a population less than 1 million residents. Observations are weighted by average county population over the sample time period from 1992 to 2005. Miles per hour wind-speeds of the Fujita-scale: F0 (40-72), F1 (73-112), F2 (113-157), F3 (158-206), F4 (207-260), F5 (261-318).

**Table 18: Socioeconomic Response Heterogeneity: Baseline Effects with Socioeconomic Interactions**

	Out-migration flows				In-migration flows			
	Model 1		Model 2		Model 3		Model 4	
# of tornado occurrences	-0.039*	X	-0.016*	X	-0.018	X	-0.0014	X
	(0.021)		(0.0084)		(0.025)		(0.0083)	
% age 17-	-0.00025	0.00052	-0.000070	-	0.0056	0.000048	0.0056	-
	(0.0034)	(0.00042)	(0.0034)		(0.0040)	(0.00042)	(0.0040)	
% age 18-24	0.011***	0.00029	0.011***	-	0.0042	0.00051	0.0044	-
	(0.0041)	(0.00027)	(0.0041)		(0.0054)	(0.00032)	(0.0054)	
% age 65+	-0.012***	0.00016	-0.012***	-	0.0051	0.00039	0.0056	-
	(0.0034)	(0.00032)	(0.0034)		(0.0056)	(0.00035)	(0.0057)	
% w/ no h.s.	-0.0022	-0.000098	-0.0022	-0.000036	-0.0075***	-0.000040	-0.0074***	-
	(0.0020)	(0.00025)	(0.0020)	(0.00030)	(0.0026)	(0.00030)	(0.0026)	
% w/ some h.s.	-0.0081***	0.00073**	-0.0080***	0.00062*	-0.018***	-0.00072*	-0.018***	-0.00072**
	(0.0024)	(0.00035)	(0.0024)	(0.00032)	(0.0029)	(0.00043)	(0.0030)	(0.00032)
% w/ some college	0.0048**	0.00024	0.0048**	0.00027*	0.0013	0.00027	0.0014	0.00040**
	(0.0022)	(0.00020)	(0.0023)	(0.00016)	(0.0030)	(0.00025)	(0.0030)	(0.00018)
% w/ college degree	0.0058**	0.00017	0.0058**	0.00018	0.0081***	-0.00052***	0.0080***	-0.00019*
	(0.0025)	(0.00014)	(0.0025)	(0.00012)	(0.0030)	(0.00020)	(0.0031)	(0.00011)
% Black	-0.0047***	-0.000078	-0.0047***	-0.000060	-0.00049	0.00015**	-0.00045	0.00014**
	(0.0016)	(0.000058)	(0.0016)	(0.000053)	(0.0018)	(0.000066)	(0.0018)	(0.000061)
% Hispanic	-0.0071***	-0.000057	-0.0071***	-	-0.0017	-0.000011	-0.0016	-
	(0.0017)	(0.000094)	(0.0017)		(0.0022)	(0.00011)	(0.0022)	
Avg. Hh. income (\$10k)	0.0024*	0.000092	0.0024*	-	-0.0055***	0.00035**	-0.0054***	-
	(0.0013)	(0.00012)	(0.0013)		(0.0014)	(0.00018)	(0.0014)	
House value (\$10K)	-0.0024***	0.000013	-0.0024***	0.000026	-0.0013**	0.000040	-0.0013**	0.000063***
	(0.00060)	(0.000021)	(0.00060)	(0.000020)	(0.00062)	(0.000025)	(0.00062)	(0.000024)
% renter occupied housing	0.0036	-0.00047	0.0035	-	-0.014*	0.0011	-0.013*	-
	(0.0055)	(0.00079)	(0.0055)		(0.0076)	(0.0011)	(0.0075)	
% mobile occupied housing	-0.0038**	-0.00017	-0.0037**	-0.00026***	-0.012***	-0.00010	-0.012***	0.00012
	(0.0016)	(0.00016)	(0.0016)	(0.000083)	(0.0020)	(0.00023)	(0.0020)	(0.00011)
% farm, fish, or forest workers	0.00046	0.00041*	0.00047	0.00042**	0.0098***	-0.00024	0.0097***	-
	(0.0017)	(0.00022)	(0.0017)	(0.00021)	(0.0026)	(0.00026)	(0.0026)	
% workers w/ residence in county	-0.0028***	0.000011	-0.0028***	-	-0.0051***	0.000017	-0.0051***	-
	(0.00088)	(0.000036)	(0.00088)		(0.0013)	(0.000046)	(0.0013)	

**Table 18 (continued):**

*Notes:* Dependent variable is log of total (out- or in-) migrant tax returns for a given county-year. Coefficients represent the fractional change in migrant household tax returns for a one unit change in the independent variable. The explanatory variables describe the county of origin for out-migration flows and the destination county for in-migration flows. Standard errors clustered by county are reported in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. All (county-level) conditional fixed effects models include controls for *Log population*, *Log unemployment*, county demographics, binary indicators for declared presidential disasters by type and year, and state-year (unconditional) fixed-effects. The sample of observations includes all counties with a population less than 1 million residents—a total of 2,978 counties. Out-migration models have 41,089 and account for approximately 58.9%. In-migration models have 40,999 observations and account approximately 48.3% of the total variation in the dependent variable. Observations are weighted by average county population over the sample time period from 1992 to 2005.

**Table 19: The Effect of an Additional Tornado by Migration Response Variables and Tornado Prone Regions**

(Independent Variable of Interest: Number of Tornado Occurrences)

Dependent variable	Out-migrants			In-migrants		
	All counties	20 states	10 states	All counties	20 states	10 states
Log migrant tax returns	.001* (.0006)	.0005 (.0008)	.0013 (.0011)	-.0003 (.0006)	-.001 (.0009)	.0014 (.0011)
Log migrant exemptions	.0008 (.0006)	.0004 (.0008)	.0015 (.0011)	-.0004 (.0007)	-.0016 (.001)	.001 (.0013)
Expected distance traveled	.1601 (.1471)	.674 (.5172)	.5815 (.3909)	.0122 (.1951)	-.957* (.5098)	.1014 (.4482)
Std. dev. of distance	.1971 (.2047)	.3758 (.5003)	.1032 (.4796)	-.1007 (.2354)	-1.0366* (.5712)	-.6695 (.593)
Skew of distance	-.001 (.006)	-.0095 (.0073)	-.0028 (.008)	.0062 (.008)	.0173 (.0092)	-.012 (.0107)

*Notes:* All coefficients represent the effect on the dependent variable (entries of first column) from an additional tornado from separate county-level conditional fixed effects linear regression models. Robust standard errors clustered on county are in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Each separate model controls for *Log population*, *Log unemployment*, county demographics, binary indicators for declared presidential disasters by type and year, and state-year (unconditional) fixed-effects. Additional each model weights observations by a county's average population size over the duration of the estimation sample (1992-2005). For *Log migrant tax returns* (*Log claimed exemptions*) by migrant type of a county, these effects are the percent change in household migrant tax returns (claimed exemptions by migrants). Coefficients for *Expected distance traveled* (weighted-average), *Standard deviation of distance traveled*, and *Skew of distance traveled* are level-changes in miles. Each column, excluding the first, delineates a different estimation sample. The first three columns consider only out-migrant flows from a county. The last three columns use in-migrant flows into a county. "All counties" indicates all (conterminous) flows are used during estimation. "20 states" ("10 states") indicates that the estimation sample is only for those migrants that either left the 20 (10) most tornado prone state, for the case of out-migrants, or that came into the 20 (10) most prone tornado states, for the case of in-migrants.

**Table 20: The Effect of an F5 Tornado by Migration Response Variables and Tornado Prone Regions**  
 (Independent Variable of Interest: Any F5 Tornado Occurrence)

Dependent variable	Out-migrants			In-migrants		
	All counties	20 states	10 states	All counties	20 states	10 states
Log migrant tax returns	.0354 (.0224)	.0192 (.0161)	.0371** (.0169)	-.0078 (.009)	-.0386** (.0159)	-.0253*** (.0094)
Log migrant exemptions	.0393* (.0211)	.0241 (.0172)	.0394** (.0189)	-.0134* (.0074)	-.0414*** (.0121)	-.0344*** (.0079)
Expected distance traveled	6.308** (3.0256)	10.6252 (9.8743)	7.6735 (7.3314)	-9.6401** (3.7718)	-7.9619 (13.53)	-1.9281 (7.7961)
Std. dev. of distance	8.4984 (5.773)	5.1094 (10.265)	4.128 (8.4388)	-18.0779** (7.8364)	-7.0106 (7.9355)	-3.9531 (9.5214)
Skew of distance	-1.1574 (.1414)	-.1563 (.1609)	-.0461 (.1028)	-.5031 (.4479)	-.252 (.1971)	-.1955 (.173)

*Notes:* All coefficients represent the effect on the dependent variable (entries of first column) from an F5 tornado event from separate county-level conditional fixed effects linear regression models. Robust standard errors clustered on county are in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Each separate model controls for the # of tornado events (current and previous year), Log population, Log unemployment, county demographics, binary indicators for declared presidential disasters by type and year, and state-year (unconditional) fixed-effects. See the notes of Table 7 for further interpretation of this table.

**Table 21: Recency Effects on Contemporaneous Migration Response**

(Independent Variable of Interest: Duration Since Last Tornado Event)

Dependent variable	Out-migrants			In-migrants		
	All counties	20 states	10 states	All counties	20 states	10 states
Log migrant tax returns	.0003 (.0002)	-.0005* (.0003)	-.0006 (.0005)	.0001 (.0003)	.0001 (.0004)	.0001 (.0006)
Log migrant exemptions	.0003 (.0003)	-.0006** (.0003)	-.0008 (.0006)	.0001 (.0004)	.000 (.0004)	-.0001 (.0007)
Expected distance traveled	.0372 (.0895)	-.4622** (.2135)	-.1353 (.1913)	.1452** (.0572)	-.184 (.169)	-.0273 (.222)
Std. dev. of distance	-.0289 (.1163)	.0483 (.2284)	-.0389 (.2663)	.1551* (.0871)	-.1364 (.216)	.1331 (.3372)
Skew of distance	-.0018 (.0026)	-.0036 (.0041)	-.0037 (.0049)	.0026 (.0025)	.0013 (.0033)	-.0021 (.0056)

*Notes:* All coefficients represent the effect on the dependent variable (entries of first column) from an increase in the duration since the previous tornado in the county. Each coefficient (and standard error) comes from a separate county-level conditional fixed effects linear regression models. Robust standard errors clustered on county are in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Each separate model controls for the # of tornado events (current and previous year), Log population, Log unemployment, county demographics, binary indicators for declared presidential disasters by type and year, and state-year (unconditional) fixed-effects. See the notes of Table 7 for further interpretation of this table.



**Table 22: The Effects of Spatially Displaced Tornado Occurrences on County-level Migration**

	Out-migration Flows		In-migration Flows	
	Model 1	Model 2	Model 4	Model 5
# of tornado occurrences	0.00098* (0.00057)	0.0015*** (0.00058)	-0.00030 (0.00063)	0.000052 (0.00063)
# of tornadoes in 0-20 mile buffer region	-	-0.00058** (0.00028)	-	-0.00039 (0.00032)
# of tornadoes in 20-50 mile buffer region	-	-0.00041** (0.00016)	-	-0.00026 (0.00023)
Constant	-4.81*** (0.51)	-4.81*** (0.51)	-4.51*** (0.70)	-4.51*** (0.69)
County demographics	Yes	Yes	Yes	Yes
Controls for disaster decls.	Yes	Yes	Yes	Yes
County fixed-effects	Yes	Yes	Yes	Yes
State-year fixed-effects	Yes	Yes	Yes	Yes
Observations	41089	41089	40999	40999
R-squared	0.589	0.589	0.482	0.482
Number of counties	2978	2978	2978	2978

*Notes:* Dependent variable is log of total (out- or in-) migrant tax returns for a given county-year. Coefficients represent the fractional change in migrant household tax returns for a one unit change in the independent variable. The explanatory variables describe the county of origin for out-migration flows and the destination county for in-migration flows. Standard errors clustered by county are reported in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. All models include county-specific and state-year-specific fixed effects. Sample of observations includes all counties with a population less than 1 million residents. Observations are weighted by average county population over the sample time period from 1992 to 2005.

## APPENDIX B

### CALCULATION DETAILS OF THE OBJECTIVE COMPLEXITY MEASURES

#### *B.1 Entropy and the Standard deviation in Fitted Utilities*

Using a random utility framework, we assume individual  $i$ 's expected utility of the program alternative  $j$  (Program A, Program B, status quo) on choice occasion  $k$  to be of the form:

$$V_{ijk} = x_{ijk}\beta + \varepsilon_{ijk}, \quad (1.1)$$

where  $\varepsilon_{ijk}$  represents the unobserved stochastic component of the utility.  $x_{ijk}\beta$  is a linearly and additively separable utility-index consisting of each program attribute (e.g.,  $x_{ijk} \in$  monthly cost, risk reduction, latency, years sick, unexpected lost life years) and their respective marginal utility, which we parameterize within the coefficient vector  $\beta$ . Under the assumption that the errors are independently and identically distributed with a type-I extreme value distribution, we estimate the marginal utilities for the program attributes using a conditional logit choice model.

We obtain the predicted values for the marginal utilities,  $\hat{\beta}$ , and then calculate the fitted systematic portion of utility,  $x_{ijk}\hat{\beta}$ , and the predicted probability of each alternative,

$\hat{P}_{ijk} = \exp(x_{ijk}\hat{\beta}) / \sum_J \exp(x_{ijk}\hat{\beta})$ . This enables us to calculate our desired utility-space

choice complexity measures:

$$\text{Std. dev. of fitted } U_{ik} = \sqrt{\left[ \sum_J \left( x_{ijk}\hat{\beta} - \overline{x_{ijk}\hat{\beta}} \right)^2 \right] / (J-1)} \quad (1.2)$$

and

$$Entropy_{ik} = -\sum_j \hat{P}_{ijk} \log(\hat{P}_{ijk}). \quad (1.3)$$

### *B.2 The Standard Deviations of “Standardized” Attribute Levels Within Each*

#### *Alternative: Mean and Dispersion across Alternatives.*

These two variables mimic the “Average  $SD_j$ ” and “Disp.  $SD_j$ ” variables described in DeShazo and Fermo (2002).  $SD_j$  is the standard deviation in the levels of the attributes for alternative  $j$ . In the DeShazo and Fermo study, just three different monotonically increasing qualitative levels were used for most of the attributes, and these were denoted as three evenly spaced integers in the analysis. They conjecture that “alternatives with attribute levels that are all either desirable, undesirable, or in-between do not require the respondent to make intra-alternative tradeoffs, and thus are cognitively easier to process than alternatives with highly dispersed attribute levels.” Again, choice difficulty is linked to attributes of the choice set.

Choice difficulty is presumed to show up in the error variance associated with the residuals in the choice model. It is helpful to combine the information in the  $SD_j$  calculations so that it describes a particular choice set, rather than just single alternatives. DeShazo and Fermo propose the use of two moments of the distribution of this measure across the alternatives in a choice set—the mean and the standard deviation (dispersion). For ease of exposition, we sometimes refer to attributes as “rows” in a choice table, and alternatives as “columns”:

The necessary computations include:

1. Standardize each attribute featured in choice set  $k$  using the  $J$  different levels of the attribute appearing within the choice set. This normalization ensures that each attribute has the same mean and variance in its levels, so that the scale of measurement does not function as a weight in the calculations. For each attribute, this normalization step first requires the calculation of both the mean and the standard deviation of the levels *of that attribute* across alternatives  $j=1, \dots, J$  (i.e. calculate the row mean and row standard deviation). For example,

$$\overline{years\ sick}_{i,k} = \frac{1}{J} \sum_J years\ sick_{ijk}$$

$$SD\ years\ sick_{i,k} = \sqrt{\frac{1}{(J-1)} \sum_J \left( years\ sick_{ijk} - \overline{years\ sick}_{i,k} \right)^2}.$$

Starting from each raw attribute levels, subtract the mean and divide by the standard deviation.

$$years\ sick_{ijk}^* = \frac{years\ sick_{ijt} - \overline{years\ sick}_{i,t}}{SD\ years\ sick_{i,t}}$$

Now each attribute has been “standardized.”

2. *Within* each alternative  $j$  (i.e. for each column), calculate the standard deviation across attributes (rows) of these standardized attributes (e.g.  $years\ sick_{ijk}^*$ ) for individual  $i$ , which we denote as  $s_{ijk}$ , which has one value for each alternative in the choice set.

3. Next, look *across* alternatives. Summarize the distribution of  $s_{ijk}$  values in that particular choice set by using the across-alternative mean and standard deviation (which we call “dispersion” to minimize confusion):

$$Mean\ SD_{i,k} = \bar{s}_{i,k} = \frac{1}{J} \sum_J s_{ijk} \quad (1.4)$$

$$Disp\ SD_{i,k} = \sqrt{\frac{1}{J-1} \left[ \sum_J (s_{ijk} - \bar{s}_{i,k})^2 \right]}. \quad (1.5)$$

As a measures of this particular type of choice set complexity, we can include the calculated values of  $Mean\ SD_{i,k}$  and  $Disp\ SD_{i,k}$  for each choice set for each individual in our models to explain subjective choice difficulty.

## APPENDIX C

### REVIEW OF THE BASIC “STRUCTURAL” SPECIFICATION FOR UTILITY

We rely on the model developed in Cameron and DeShazo (2009) to guide our specification of a “structural” utility model to use for the utility parameters that permit us to construct utility-space measures of the closeness of alternatives in attribute space. In this appendix, we review the model that is developed there. Note that we stop short of the full fourteen-parameter specification used in that paper, relying instead on a simplified five-parameter version with a greater degree of linearity and less heterogeneity with respect to respondents’ ages.

To sketch the intuition behind the choice model, consider just the pairwise choice between Program A and the status quo “numeraire” alternative (N).<sup>62</sup> Define the discount rate as  $r$  and let  $\delta^t = (1 + r)^{-t}$ . Let  $\Pi_i^{NS}$  be the probability of suffering the adverse health profile (i.e. getting “sick”) if the status quo alternative is selected for person  $i$ , and let  $\Pi_i^{AS}$  be the reduced probability of suffering the adverse health profile if Program A is chosen. The difference between  $\Pi_i^{NS}$  and  $\Pi_i^{AS}$  is  $\Delta\Pi_i^A$ , which is a negative number—the risk reduction to be achieved by Program A.

The sequence of health states that makes up an illness profile is captured by a sequence of mutually exclusive and exhaustive (0, 1) indicator variables associated with each future time period. These are defined as  $1(pre_u^A)$  for pre-illness years,  $1(ill_u^A)$  for

---

<sup>62</sup> This model can be readily generalized to a three-alternative case including Program A, Program B, and Neither Program (N).

illness-years,  $1(rcv_u^A)$  for post-illness recovered years, and  $1(lyl_u^A)$  for life-years lost. All profiles in the survey involve just one “spell” of sickness (i.e. moderate, or severe pain or disability).

The present discounted number of years making up the remainder of the individual’s nominal life expectancy,  $T_i$ , is given by  $pdvc_i^A = \sum_{t=1}^{T_i} \delta^t$ . Other relevant discounted spells, also summed from  $t=1$  to  $t = T_i$  include  $pdve_i^A = \sum \delta^t 1(pre_u^A)$ ,  $pdvi_i^A = \sum \delta^t 1(ill_u^A)$ ,  $pdvr_i^A = \sum \delta^t 1(rcv_u^A)$ , and  $pdvl_i^A = \sum \delta^t 1(lyl_u^A)$ . The four different health states exhaust the individual’s nominal life expectancy, so  $pdve_i^A + pdvi_i^A + pdvr_i^A + pdvl_i^A = pdvc_i^A$ . Finally, to accommodate the assumption that individuals expect to pay program costs only during the pre-illness or post-illness recovered periods, we define  $pdvp_i^A = pdve_i^A + pdvr_i^A$ . Absent any other data about future income expectations, we assume that individuals choose programs as though they expect to maintain their income at current real levels until they die.

To further simplify notation, let  $cterm_i^A = [(1 - \Pi_i^{AS})] pdvc_i^A + \Pi_i^{AS} pdvp_i^A$  and let  $yterm_i^A = [-pdvc_i^A + \Pi_i^{AS} pdvi_i^A + \Pi_i^{NS} pdvl_i^A]$ . Assume that expectations are taken across the two possible health outcomes,  $S$  = gets sick (suffers this illness) and  $H$  = remains healthy (does not suffer the illness). Then the present discounted value of the expected utility-difference that drives the individual’s choice between Program A and the status quo can be written succinctly as:

$$\begin{aligned}
PDV(E_{s,H}[\Delta V_i^A]) &= \beta_0 \{(Y_i - c_i^A) cterm_i^A + Y_i yterm_i^A\} \\
&\quad + \beta_1 \{(Y_i - c_i^A)^2 cterm_i^A + Y_i^2 yterm_i^A\} \\
&\quad + \alpha_1 \{\Delta \Pi_i^{AS} pdvi_i^A\} + \alpha_2 \{\Delta \Pi_i^{AS} pdvr_i^A\} + \alpha_3 \{\Delta \Pi_i^{AS} pdvl_i^A\} + \varepsilon_i^A
\end{aligned} \tag{1.1}$$

We refer to the five terms in braces as the “structural” attributes associated with Program A. These structural attributes can be constructed from the descriptions of the (raw) program attributes for each alternatives described in each choice set, given a specific assumption about the discount rate.<sup>63</sup> However, Cameron and DeShazo (2009) find that this simple five-parameter model is dominated by a specification that is not simply *linear* the three types of present-discounted health-state years. Instead, a shifted log formulation of last three terms (e.g.  $\{\Delta \Pi_i^{AS} \log(pdvi_i^A + 1)\}$ ) appears to explain respondents’ choices better, so we use this modification as well.<sup>64</sup>

In the utility-theoretic specification, the structural variables modified by the coefficients  $\beta_0$  and  $\beta_1$  in equation (1) will be referred to simply as the “Linear net income term” and the “Quadratic net income term” in the lower panel of Table 3. These are the same two parameters in the linear and squared terms in net income in the underlying single-period indirect utility function.

In this paper, we will work with this relatively simple five-parameter specification in order to keep the dimensionality of the problem manageable, although we

---

<sup>63</sup> In this paper, we assume a common discount rate of 5%. In Cameron and DeShazo (2009), for an expanded structural specification, the consequences of assuming either a 3% discount rate or a 7% discount rate are also explored.

<sup>64</sup> An even better-fitting specification is obtained when the term in lost life-years also enters quadratically and when the terms in sick-years and lost life-years are interacted. Also, several of the marginal utility parameters appear to vary systematically as quadratic functions of the respondent’s current age.



acknowledge that the functional form represents a local approximation under an assumption of homogeneous (average) preferences.<sup>65</sup>

---

<sup>65</sup> The opportunity for longer durations in each health state is also correlated with the youth of the respondent, so there is some advantage to allowing some of the  $\alpha$  coefficients to be quadratic functions of the respondent's current age. Other papers in this series have also explored a minor correction for sample representativeness that involves a measure of the deviation of the subject's survey response probability from the average for the half-million panelist recruitment contacts for the survey research firm.

**APPENDIX D**  
**THE STANDARD DEVIATION OF FITTED UTILITY AS A DETERMINANT**  
**OF SUBJECTIVE CHOICE DIFFICULTY**

**Table D.1: Additional Determinants of Subjective Choice Difficulty**

COEFFICIENT	Model 4 Ordered Probit	Model 5 Ordered Probit	Model 6 LS Fixed Effects
<b>Measures of choice set complexity:</b>			
<i>(a) In utility space:</i>			
Std. dev. of fitted U	-0.371** (-2.09)	-0.509*** (-2.60)	-0.527** (-2.26)
<i>(b) In attribute space:</i>			
<i>(2) Distribution of within-alt. attrib. variability</i>			
Mean std. dev.	-0.320* (-1.82)	-0.045 (-0.24)	0.143 (0.65)
Disp. of std. dev.	-0.026 (-0.11)	0.230 (0.90)	0.546* (1.84)
<i>(2) Across-alt. attrib. variability (Ad hoc model)</i>			
Std. dev. of annualized costs	0.021 (0.78)	0.087*** (2.84)	0.069* (1.80)
" risk difference	-0.051 (-1.35)	-0.065* (-1.68)	0.048 (1.06)
" latency	0.034 (1.08)	-0.091* (-1.77)	-0.084 (-1.41)
" years sick	0.073*** (3.73)	0.030 (1.32)	0.037 (1.40)
" lost life years	-0.015 (-0.56)	-0.060* (-1.86)	-0.050 (-1.33)
<i>(4) Choice occasion indicators</i>			
2nd choice occasion	-	-0.151*** (-3.96)	-0.336*** (-8.40)
3rd "	-	-0.282*** (-7.33)	-0.341*** (-8.38)
4th "	-	-0.349*** (-8.98)	-0.371*** (-8.95)
5th "	-	-0.435*** (-11.13)	-0.480*** (-11.53)
<b>Obj. measures of cognitive capacity:</b>			
1(Less than high school)	-	0.128*** (2.88)	n/a
1(High school degree)	-	0.059** (2.05)	n/a

**Table D.1 (continued):**

COEFFICIENT	Model 4 Ordered Probit	Model 5 Ordered Probit	Model 6 LS Fixed Effects
Avg. duration on other choice occasions	-	0.002*** (4.41)	-0.033*** (-16.78)
1(Valid duration)	-	-0.506*** (-3.63)	-1.486*** (-8.16)
<b>Obj. measures of sociodemographics:</b>			
Income (in \$1000)	-	0.001** (1.98)	n/a
Age	-	-0.018*** (-2.70)	n/a
Age <sup>2</sup>	-	0.009 (1.53)	n/a
1(Female)	-	0.046* (1.85)	n/a
1(Divorced)	-	-0.021 (-0.50)	n/a
1(Black)	-	-0.184*** (-4.03)	n/a
1(Other ethnicity)	-	-0.002 (-0.04)	n/a
1(Hispanic)	-	0.093* (1.73)	n/a
Household size	-	0.023 (1.19)	n/a
# of children	-	-0.055** (-2.25)	n/a
1(Dual income household)	-	-0.050 (-1.63)	n/a
1(Single parent)	-	-0.052 (-0.50)	n/a
<b>Survey-specific respondent characteristics:</b>			
Illness experience count (0-13):	-	-0.010*** (-2.83)	n/a
Avg. subj. risk of future experience (0-4):	-	0.092*** (5.69)	n/a
Subjective controllability of risks (0-4):	-	0.046*** (3.42)	n/a
1(Missing health)	-	0.021 (0.46)	n/a
<b>Respondent attention behavior:</b>			
1(All status quo)	-	-0.067* (-1.70)	n/a
1(No change in difficulty rating)	-	-0.633***	n/a

**Table D.1 (continued):**

COEFFICIENT	Model 4 Ordered Probit	Model 5 Ordered Probit	Model 6 LS Fixed Effects
Constant <sup>b</sup>	-	- (-18.73)	6.011*** (15.36)
Observations	7392	7392	7392
Number of caseid	1694	1694	1694
LogL	-12838.89	-12465.671	-10117.361

*Notes:* All models are pooled ordered-probit except for Model 6, which uses a linear fixed-effects estimator and clusters on errors at the respondent level. Robust t-statistics in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. The intercept is normalized to zero for ordered probit models. Incidental threshold parameter estimates are not reported.

**Table D.2: Using *Std. de. Of fitted U*, Comparison of Ad hoc and Structural Specifications**

COEFFICIENT	Model 6 LS Fixed Effects	Model 7 LS Fixed Effects
<b>Measures of choice set complexity:</b>		
<i>(a) In utility space:</i>		
Std. dev. of fitted U	-0.527** (-2.26)	-0.503*** (-2.84)
<i>(b) In attribute space:</i>		
<i>(4) Distribution of within-alt. attrib. variability</i>		
<i>Mean SD<sub>i,k</sub></i>	0.143 (0.65)	0.143 (0.68)
<i>Disp. SD<sub>i,k</sub></i>	0.546* (1.84)	0.569** (1.97)
<i>(5) Across-alt. attrib. variability (Ad hoc model)</i>		
Std. dev. of annualized costs	0.069* (1.80)	-
" risk difference	0.048 (1.06)	-
" latency	-0.084 (-1.41)	-
" years sick	0.037 (1.40)	-
" unexpctd. lost life years	-0.050 (-1.33)	-
<i>(6) Across-alt. attrib. variability. (Structural model)</i>		
Std. dev. of linear net income term	-	0.061 (1.09)
" quadratic net income term	-	-0.003 (-0.05)
" $\Delta\Pi_i^{AS} \log(pdv_i^A + 1)$	-	0.085*** (3.17)
" $\Delta\Pi_i^{AS} \log(pdvr_i^A + 1)$	-	0.015 (0.92)
" $\Delta\Pi_i^{AS} \log(pdv_i^A + 1)$	-	-0.007 (-0.24)
<b>Observable proxies for cognitive capacity</b>	Yes	Yes <sup>b</sup>
<b>Obs. measures of sociodemographics</b>	"	n/a
<b>Survey-specific health characteristics</b>	"	n/a
<b>Attention behavior controls</b>	"	n/a
Constant	6.011*** (15.36)	5.874*** (15.00)
Number of choices	7392	7392

**Table D.2 (continued):**

Number of respondents	1694	1694
Maximized LogL	-10117.36	-10117.49

*Notes:* All models use a linear fixed-effects estimator and cluster on errors at the respondent level. Robust t-statistics in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

**APPENDIX E**  
**ADDITIONAL BEHAVIORAL MODEL SPECIFICATIONS**

**Table E.1: Logit modeling of decision to accept proposed monetary transfers**

	Conditional Logit (1)	Conditional Logit (2)	Conditional Logit (3)	Mixed Logit (4)	
COEFFICIENTS				Mean	SD
Payouts	-0.035*** (0.000)	-0.038*** (0.000)	-0.038*** (0.000)	-0.055*** (0.000)	-
Giving alt. specific constant <sup>^</sup>	-2.081*** (0.001)	-1.794** (0.018)	-1.898** (0.013)	-0.862 (0.273)	1.201*** (0.000)
Giving a.s. constant <i>interactions</i>					
Age	0.039*** (0.000)	0.041*** (0.000)	0.040*** (0.001)	0.715*** (0.000)	0.538*** (0.000)
Observed (=1)	0.568*** (0.000)	0.519*** (0.000)	0.526*** (0.000)	0.044 (0.135)	0.019*** (0.010)
Importance rating (1-10) <sup>^^</sup>	0.129** (0.015)	0.163** (0.015)	0.091 (0.189)	0.133*** (0.005)	0.092*** (0.000)
Empathy rating (1-10) <sup>^^^</sup>	-	-	0.091 (0.137)	-	-
Charity-specific indicators	No	Yes	Yes	Yes	-
Log-likelihood	-2001.4	-1893.0	-1887.3	-1470.1	
Observations	3,168	3,168	3,168	3,168	-
Number of participants	44	44	44	44	-

*Notes:* Coefficients represent the marginal propensity to accept a proposed monetary transfer. Standard errors clustered on participants. P-values are in parentheses; \*\*\*p<.01, \*\*p<.05, \*p<.1. <sup>^</sup>The alternative specific constant for accepting is interpreted as the average effect of unobserved factors on utility relative to rejecting. <sup>^^</sup> Participant ratings for "... rate the importance of the cause this charity serves" (1 – unimportant, 10 – extremely important). <sup>^^^</sup> Participant ratings for "It's easy for me to picture a person who would be helped by this charity" (1 – unimportant, 10 – extremely important). Models 1-3 are conditional logit models with choice occasion (or trial) as the unit of observation. The conditional fixed effects are individual-specific. Model 4 is a mixed logit model with choice occasion as the unit of observation and observations grouped by individual. Random parameters in Model 4 are normally distributed.



## APPENDIX F

### MIGRATION DATA CONSTRUCTION

#### *F1.1 Detailed Description the Migration Data*

The IRS calculates the number of migrant household and individual returns that move from county based on matched social security numbers (SSN) across tax years. Residential address information is extracted from the domestic household tax forms 1040, 1040A, and 1040EZ, and the foreign tax forms 10NR, 1040PR, 1040VI, and 104SS. The extraction process occurs until the 39<sup>th</sup> week in the IRS's processing year and covers 95% to 98% of all returns filed for a tax year.

The IRS and the Census Bureau jointly geocode the addresses provided on tax returns. An individual or household's tax return in a given year is classified as being a migrant return if the address listed on the return is different from the preceding tax year's address for that household or individual. Tax years lead the actual year of filing such that a "match of tax years 2003 and 2004 produces 2004 to 2005 migration estimates." The migration year refers is the same as the filing year. Individuals could experience a tornado between January 1<sup>st</sup>, 2003 and December 31<sup>st</sup>, 2003, move to a different county on January 1<sup>st</sup>, 2004, and subsequently be classified as a migrant for the year of 2004. For this reason, in some of our empirical specifications that model migration flows, we combine current year tornado activity with the activity of the previous year.

The IRS produces three categories of match status: (1) matched social security numbers across the two years, (2) unmatched, with the previous tax year missing, and (3)

unmatched, with the current tax year is missing. Unmatched tax returns are not recorded in the migration data. On occasion, an individual that physically migrates could potentially not be recorded as a migrant. This could occur when two individuals become a joint filer or split from a joint filing to two individual returns and move during the same year. In these cases, the secondary filer is given an unmatched status and is not directly recorded as an (additional) migrant return in the data. However, the total exemptions would roughly stay the same for the pair of individuals across years. The IRS migration data includes information on exemptions claimed on tax returns such that total exemptions can be used as an alternative dependent variable as a robustness check for our findings. There are similar procedures taken for filed returns which have the primary or secondary filer listed as deceased. Zero exemption returns, mostly children who are dependents on another return but earn a sufficient amount of income to file an individual return, are not counted as returns nor included in the total exemptions for a county.

We also have the potential to use IRS county-to-county migration data from the Interuniversity Consortium for Political and Social Research (ICPSR) Study #2937 in addition to the migration we obtained directly from the IRS. The ICPSR data contains in-migration data for the years 1978-1979 and 1980-1992. Out-migration exists only for the years 1980-1992. The directly obtained IRS migration data span the years 1983-2007. Transitions in the data format of files makes conformable data across the entire set of years a difficult task (given our limited resources.) Currently, our analysis focuses on a conformable set of data from 1992 to 2005.

Each of the data sources contains separate files for each state-year. For out-migration, each file records the number of tax returns filed in the previous year for an origin county (within a given state) that file in any other (destination) county in the United States when the number of in- or out-migrant tax returns is greater than or equal to 10. The county-to-county records are censored when the number tax returns is less than 10. However, for a given origin, the censored county-level destination records are summed with the uncensored county-level destination records into an aggregate record of the total number of migrant tax returns. Thus we are able to infer the aggregate level of censored migration from the total migrant tax returns for a given county and the summed uncensored migrant tax returns. The data files also provide total tax returns of a county and the level of aggregate international migrant tax returns, if any. Both the ICPSR and the IRS data needed to be converted into a workable format. In total, the two sources of data contain seven different formats for the raw data, which we reformat into one common form.

### *F1.2 Censoring*

With the IRS data, an important consideration is that migration flows consisting of fewer than ten households between any pair of counties are censored. These minor flows are aggregated up to a coarser jurisdictional level, for confidentiality. This censoring of migration counts would have an expected downward bias on the estimated effects since this measure aggregates migration across all destination counties.<sup>66</sup> For

---

<sup>66</sup> For example, if the *ex-ante* and *ex-post* out-migration level to any other county was below ten households for all of the approximately 3,200 counties then the estimated effect of tornado activity on the distribution

expected distance traveled, the potential bias is more complicated because the direction of the effect depends on the relationship between total migrants and distance traveled. Without direct knowledge of the relationship, one might hypothesize that migration counts would decrease with distance such that changes in the counts for households that move large distances are likely to be censored. This would lead to an over-weighting of households who move shorter distances and a downward bias in the estimated effect of tornado activity on distance traveled.<sup>67</sup> However, there are many counties for which the destinations with the most migrants are also further away, which would thus lead to a potential upward bias in distance traveled.

### *F2.1 The Neighborhood Change Database (NCDB)*

Information on economic indicators and demographic information come from the CensusCD Neighborhood Change Database (NCDB) by Geolytics©. This database contains census tract level information for the entire U.S. for the 1970, 1980, 1990, and 2000 decennial census. The benefit of this database is that all the reported variables are normalized on the 2000 census tracts. For our analysis, we aggregate these data to the county level. Not all counties have information for the 1970 and 1980 decennial census. This is because the censuses for those years were of limited national scope, focusing primarily on metropolitan areas. The number of represented counties for 1990 and 2000

---

of *distances* would be zero, even though we can infer from the data the total *number* of censored households that move.

<sup>67</sup> This is under the assumption that the marginal effect of activity on distance is independent of distance.

is 3,141 counties. For 1970 and 1980, the number of counties is substantially less at 615 and 911, respectively.

We selected NCDB variables in the 1990 and 2000 decennial census (see “Selected NCDB variables” below) in our analysis from the years from 1992 to 2005. We interpolate (and extrapolate) values at the census tract level for the non-decadal years and then aggregate the tract-level information to the county level. This method results in some variables (see “Constructed variables” below) with negative values for the years in which we extrapolate, i.e., 2001-2005, for counties that had relatively large changes during the time period. Table F2.1 reports the number of counties in 2001 and 2005 for our constructed variables that had negative values. As can be seen, the number of such counties amounts to a relatively small fraction of the total number of counties. We recode these negative values to zero for our analysis.

**Selected NCDB variables (available at tract level):**

(\* indicates 9 or 0, corresponding to the years 1990 and 2000.)

**Age Distribution**

PERS517*	Persons 5-17
YTHPOP*	Persons 18-24
PERS64*	Persons 18-64
PERS65P*	Person 65+

**Education**

EDUC8*	Persons 25+ who have 0-9 years of school
EDUC11*	Persons 25+ who have 9-12 years of school but no diploma
EDUC12*	Persons 25+ who have completed h.s. but no college experience
EDUC15*	Persons 25+ who have completed some college, but have no degree
EDUC16*	Persons 25+ who have with a bachelors or graduate/professional degree
EDUCPP*	Persons 25+

**Labor**

INDEMP*	Civilian employed persons 16+ years old
OCC9*	Persons 16+ employed as farm workers or in forestry and fishing

**Family Structure**

NUMHHS*	Total number of households within census tract
NONFAM*	Total number of nonfamily households within census tract

**General Population Characteristics**

TRCTPOP*	Total population
----------	------------------

SHRWHT*N	Total White population
SHRBLK*N	Total Black population
SHRHSP*N	Total Hispanic Population
ADULT*N	Persons 18+ years old
CHILD*N	Persons less than 18 years old
OLD*N	Persons 65+ years old
KIDS*N	Persons less than 5 years old
MILTQ*N	Population in military quarters
ROOMH*N	Population in rooming houses
OTHGQ*N	Population in other group quarters
<b><i>Housing characteristics, tenure, occupancy, costs</i></b>	
OCCHU*	Total occupied housing units
VACHU*	Total vacant year-round housing units
TOTHSUN*	Total (year-round) housing units
TTUNITM*	Total housing units consisting of a mobile home or trailer
REUNITM*	Renter-occupied housing units
AGGVAL*	Aggregate value for spec. owner-occupied noncondo-housing units

***Income and Earnings***

FAVINC*N	Aggregate family income (\$)
FAVINC*D	Total families
AVHHIN*N	Aggregate household income last year (\$)

***Mobility/Transportation***

WRCNTY*N	Workers 14+ working within their county of residence
WRCNTY*D	Workers 14+ reporting place of work
SMHSE*N	Persons 5+ residing in the same house five years ago
SMHSE*D	Persons 5+ years old
LCLMV*N	Persons 5+ residing in different house five years ago but in same county
SMCNTY*N	Persons 5+ residing in same county five years ago
SMST*N	Persons 5+ residing in different house five years ago but in same state

***Poverty/Public Assistance***

POVRAT*N	Persons in families and unrelated individuals below poverty level last year
POVRAT*D	Total population with poverty status determined

***NCDB-specific***

STATECD	State FIPS in 2000
COUNCD	County FIPS in 2000
PCTCOV70	Percentage 2000 blocks covered by 1970 tracts
PCTCOV80	Percentage 2000 blocks covered by 1980 tracts

**Constructed variables (for each county-year observation):*****Age Distribution***

<i>age17L</i>	Proportion of population less than 17 years old
<i>age1824</i>	Prop. of population in county of age 18-24 years old
<i>age2564</i>	Prop. of population in county of age 25-65 years old
<i>age65P</i>	Prop. of population in county of age 65 years or older

***Education***

<i>nohs</i>	Proportion of population 25+ who have 0-9 years of school
<i>somehs</i>	Prop. of population 25+ who have 9-12 years of school but no diploma
<i>diploma</i>	Prop. of population 25+ who have completed h.s. but no college experience
<i>somecoll</i>	Prop. of population 25+ who have completed some college, but have no degree
<i>degree</i>	Prop. of population 25+ who have with a bachelors or graduate/prof. degree

***Labor***

<i>farm</i>	Proportion of population 16+ employed as farm workers or in forestry and fishing
-------------	--

***Family Structure***

<i>tothss</i>	Total number of households
<i>nonfam</i>	Proportion of nonfamily households
<b>General Population Characteristics</b>	
<i>pctwhite</i>	Proportion of population that is White
<i>pctblack</i>	Prop. of population that is Black
<i>pcthispanic</i>	Prop. of population that is Hispanic
<i>miltq</i>	Prop. of population in military quarters
<i>roomh</i>	Prop. of population in rooming houses
<i>othgq</i>	Prop. of population in other group quarters
<b>Housing characteristics, tenure, occupancy, costs</b>	
<i>occ</i>	Proportion of housing units occupied
<i>vacant</i>	Prop. of housing units vacant
<i>mobile</i>	Prop. of housing units consisting of a mobile home or trailer
<i>rentocc</i>	Prop. of housing units renter-occupied
<i>housval</i>	Avg. value for spec. owner-occupied noncondo-housing units
<b>Income and Earnings</b>	
<i>faminc</i>	Average family income (\$)
<i>hhinc</i>	Average household income (\$)
<b>Mobility/Transportation</b>	
<i>wwcnty</i>	Proportion of workers 14+ working within their county of residence
<i>smhouse</i>	Prop. of population 5+ residing in the same house five years ago
<i>smcnty</i>	Prop. of pop. 5+ residing in different house five years ago but in same county
<i>smstate</i>	Prop. of pop. 5+ residing in different house five years ago but in same state
<b>Poverty/Public Assistance</b>	
<i>poverty</i>	Proportion of population in families and unrelated individuals (for whom poverty status is determined) below poverty level last year
<b>NCDB-specific</b>	
<i>avgcov70</i>	Avg. percentage of 2000 blocks covered by 1970 tracts (in the county)
<i>avgcov80</i>	Avg. percentage of 2000 blocks covered by 1980 tracts (in the county)

**Table F2.1 – Number of counties with negative values for socioeconomic variables**

Variables	2001	2005
% age 17-	0	0
% age 18-24	0	2
% age 25-64	0	0
% age 65+	0	0
% w/ no h.s.	1	12
% w/ some h.s.	0	1
% w/ h.s. diploma	0	0
% w/ some college	0	0
% w/ college degree	0	0
% Black	17	31
% Hispanic	6	25
% White	0	1
Avg. Hh. income (\$10k)	0	0
House value (\$10K)	1	1
% of households in poverty	0	1
% occupied housing	0	0
% renter occupied housing	7	36
% mobile occupied housing	0	4
% farm, fish, or forest workers	6	211
% workers w/ residence in county	0	1

Notes: Linear extrapolation of the decennial census variables results in negative values for some variable in the years 2001-2005. We report the number of such counties for the sociodemographic variables in this table for 2001 and 2005. Negative values are recoded to zero in our empirical analysis.



**APPENDIX G**  
**ADDITIONAL (ABBREVIATED) TABLES**

**Table G.1—Temporal Effects of Leading and Lagged Tornado Occurrences on County-level Migration**

	Out-migration Flows			In-migration Flows		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
# of tornado occurrences two years ahead	-	-	-0.00040 (0.00057)	-	-	-0.00075 (0.00073)
# of tornado occurrences one year ahead	-	-	-0.00025 (0.00060)	-	-	-0.0010 (0.00077)
# of tornado occurrences	0.00100* (0.00057)	0.00096* (0.00058)	0.0015** (0.00063)	-0.00028 (0.00062)	-0.00029 (0.00067)	0.000040 (0.00071)
# of tornado occurrences last year	-	0.00097 (0.00061)	0.00047 (0.00068)	-	0.00046 (0.00082)	-0.00067 (0.00089)
# of tornado occurrences two years ago	-	-0.00029 (0.00058)	-0.00028 (0.00057)	-	-0.00065 (0.00099)	-0.00076 (0.00099)
# of tornado occurrences three years ago	-	-0.00028 (0.00064)	-0.00015 (0.00065)	-	-0.00049 (0.00086)	-0.00039 (0.00088)
# of tornado occurrences four years ago	-	0.000037 (0.00066)	0.00017 (0.00069)	-	0.00011 (0.00077)	0.00019 (0.00075)
# of tornado occurrences five years ago	-	-0.000093 (0.00064)	-0.00021 (0.00064)	-	0.00078 (0.00080)	0.00032 (0.00083)
Constant	-4.81*** (0.51)	-4.80*** (0.51)	-4.70*** (0.54)	-4.51*** (0.70)	-4.51*** (0.70)	-5.03*** (0.72)
County demographics	Yes	Yes	Yes	Yes	Yes	Yes
Controls for disaster decls.	Yes	Yes	Yes	Yes	Yes	Yes
County fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
State-year fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	41,089	41,089	37,767	40,999	40,999	37,637
R-squared	0.589	0.589	0.579	0.482	0.482	0.468
Number of counties	2,978	2,978	2,972	2,978	2,978	2,971

Notes: Dependent variable is log of total (out- or in-) migrant tax returns for a given county-year. Coefficients represent the fractional change in migrant household tax returns for a one unit change in the independent variable. The explanatory variables describe the county of origin for out-migration flows and the destination county for in-migration flows. Standard errors clustered by county are reported in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. All models include county-specific and state-year-specific fixed effects. Sample of observations includes all counties with a population less than 1 million residents. Observations are weighted by average county population over the sample time period from 1992 to 2005.

## APPENDIX H

## UNABBREVIATED SOCIODEMOGRAPHIC TABLES

Table H5.1—Demographic Response Heterogeneity for Out-migration

	Model 1		Model 2		Model 3	
	Baseline	Baseline w/ Interactions	Baseline w/ Interactions	Baseline w/ Interactions	Baseline w/ Interactions	Baseline w/ Interactions
# of tornado occurrences	0.00099* (0.00052)	0.0010* (0.00062)	X	0.0010* (0.00060)	X	
% age 17-	-0.000060 (0.0034)	-0.000028 (0.0034)	-0.000043 (0.00024)	-0.000068 (0.0034)	-	
% age 18-24	0.011*** (0.0041)	0.011*** (0.0041)	-0.000066 (0.00018)	0.011*** (0.0041)	-	
% age 65+	-0.012*** (0.0034)	-0.012*** (0.0034)	-0.00028 (0.00021)	-0.012*** (0.0034)	-	
% w/ no h.s.	-0.0022 (0.0020)	-0.0021 (0.0020)	-0.00017 (0.00025)	-0.0022 (0.0020)	-	
% w/ some h.s.	-0.0079*** (0.0024)	-0.0079*** (0.0024)	0.00038 (0.00031)	-0.0078*** (0.0024)	-	
% w/ some college	0.0049** (0.0022)	0.0049** (0.0023)	0.000050 (0.00018)	0.0050** (0.0022)	-	
% w/ college degree	0.0058** (0.0025)	0.0058** (0.0025)	0.000059 (0.00015)	0.0058** (0.0025)	0.000064** (0.000031)	
% Black	-0.0047*** (0.0016)	-0.0047*** (0.0015)	-0.000048 (0.000055)	-0.0047*** (0.0016)	-	
% Hispanic	-0.0071*** (0.0017)	-0.0071*** (0.0017)	-3.3e-07 (0.000081)	-0.0071*** (0.0017)	-	
Avg. Hh. income (\$10k)	0.00024* (0.00013)	0.00024* (0.00013)	7.6e-07 (9.8e-06)	0.00024* (0.00013)	-	
House value (\$10K)	-0.00024*** (0.000060)	-0.00024*** (0.000060)	1.0e-06 (2.1e-06)	-0.00024*** (0.000060)	-	
% renter occupied housing	0.0037 (0.0055)	0.0037 (0.0055)	-0.00050 (0.00079)	0.0036 (0.0055)	-	
% mobile occupied housing	-0.0040*** (0.0016)	-0.0038** (0.0016)	-0.00016 (0.00016)	-0.0038** (0.0016)	-0.00021*** (0.000057)	
% farm, fish, or forest wrkers	0.00088 (0.0018)	0.00060 (0.0017)	0.00043* (0.00022)	0.00069 (0.0017)	0.00028** (0.00013)	
% wrkers w/ residence	-0.0028*** (0.00088)	-0.0028*** (0.00088)	9.3e-06 (0.000036)	-0.0028*** (0.00088)	-	
Constant	-4.79*** (0.51)	-4.82*** (0.50)	-	-4.81*** (0.51)	-	
Log pop. & Log unempl.	Yes	Yes	-	Yes	-	
Controls for disaster decls.	Yes	Yes	-	Yes	-	
County fixed-effects	Yes	Yes	-	Yes	-	
State-year fixed-effects	Yes	Yes	-	Yes	-	
Observations	41,089	41,089		4,1089		
Number of counties	2,978	2,978		2,978		

**Table H5.1 (continued):**

R-squared	0.589	0.589	0.589
-----------	-------	-------	-------

Notes: Dependent variable is log of total out-migrant tax returns for a given county-year. Coefficients represent the percentage change in migrant household tax returns for a one unit change in the independent variable. The explanatory variables describe the county of origin for out-migration flows and the destination county for in-migration flows. Standard errors clustered by county are reported in parentheses; \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . All models include county-specific and state-year-specific fixed effects. Sample of observations includes all counties with a population less than 1 million residents. Observations are weighted by average county population over the sample time period from 1992 to 2005.

**Table H5.2—Demographic Response Heterogeneity for In-migration**

	Model 1	Model 2		Model 3	
	Baseline	Baseline w/	Interactions	Baseline w/	Interactions
# of tornado occurrences	0.000073 (0.00060)	0.00052 (0.00081)	X	0.00031 (0.00077)	X
% age 17-	0.0055 (0.0040)	0.0058 (0.0040)	-0.00022 (0.00028)	0.0055 (0.0040)	-
% age 18-24	0.0043 (0.0054)	0.0042 (0.0054)	0.00033 (0.00021)	0.0043 (0.0054)	-
% age 65+	0.0051 (0.0055)	0.0051 (0.0055)	0.00018 (0.00025)	0.0051 (0.0055)	-
% w/ no h.s.	-0.0073*** (0.0026)	-0.0075*** (0.0026)	-0.000079 (0.00029)	-0.0074*** (0.0026)	-
% w/ some h.s.	-0.018*** (0.0030)	-0.017*** (0.0029)	-0.00088** (0.00035)	-0.018*** (0.0029)	-
% w/ some college	0.0019 (0.0030)	0.0013 (0.0030)	0.00018 (0.00022)	0.0018 (0.0030)	-
% w/ college degree	0.0079*** (0.0031)	0.0081*** (0.0030)	-0.00057*** (0.00019)	0.0079** (0.0031)	0.000050 (0.000035)
% Black	-0.00044 (0.0019)	-0.00047 (0.0018)	0.00016*** (0.000062)	-0.00042 (0.0018)	-
% Hispanic	-0.0015 (0.0022)	-0.0017 (0.0022)	0.000017 (0.000099)	-0.0015 (0.0022)	-
Avg. Hh. income (\$10k)	-0.00053*** (0.00014)	-0.00055*** (0.00014)	0.000031** (0.000015)	-0.00053*** (0.00014)	-
House value (\$10K)	-0.00012** (0.000062)	-0.00013** (0.000061)	3.8e-06 (2.5e-06)	-0.00012** (0.000062)	-
% renter occupied housing	-0.013* (0.0075)	-0.014* (0.0075)	0.0011 (0.0012)	-0.013* (0.0075)	-
% mobile occupied housing	-0.013*** (0.0020)	-0.012*** (0.0020)	-0.000097 (0.00023)	-0.012*** (0.0020)	-0.000079 (0.000068)
% farm, fish, or forest wrkers	0.010*** (0.0026)	0.010*** (0.0026)	-0.00023 (0.00026)	0.010*** (0.0026)	-0.00035** (0.00018)
% workers w/ residence	-0.0051*** (0.0013)	-0.0051*** (0.0013)	0.000016 (0.000046)	-0.0050*** (0.0013)	-
Constant	-4.50*** (0.70)	-4.52*** (0.70)	-	-4.50*** (0.69)	-
Log pop. & Log unempl.	Yes	Yes	-	Yes	-
Controls for disaster decls.	Yes	Yes	-	Yes	-
County fixed-effects	Yes	Yes	-	Yes	-
State-year fixed-effects	Yes	Yes	-	Yes	-
Observations	40,999	40,999		40,999	
Number of counties	2,978	2,978		2,978	
R-squared	0.482	0.482		0.482	

Notes: Dependent variable is log of total in-migrant tax returns for a given county-year. Coefficients represent the percentage change in migrant household tax returns for a one unit change in the independent variable. The explanatory variables describe the county of origin for out-migration flows and the destination county for in-migration flows. Standard errors clustered by county are reported in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. All models include county-specific and state-year-specific fixed effects. Sample of observations includes all counties with a population less than 1 million residents. Observations are weighted by average county population over the sample time period from 1992 to 2005.

## REFERENCES

- National Center for Atmospheric Research (2001) Environmental and Socetal Impacts Group, and the Atmospheric Policy Program of the American Meterological Society. *Extreme weather sourcebook 2001: Economic and other soceital impacts related to hurricanes, floods, tornadoes, lighening, and other U.S. Weather phenomena*. National Center for Atmospheric Research Boulder, CO.
- Adamowicz, W., J. R. DeShazo. (2006). "Frontiers in stated preferences methods: An introduction," *Environmental & Resource Economics* 34 (1), 1-6.
- Ashley, Walker S. (2007). "Spatial and temporal analysis of tornado fatalities in the United States: 1880-2005," *Weather & Forecasting* 22 (6), 1214-1228.
- Baker, Justin, W. Douglass Shaw, David Bell, Sam Brody, Mary Riddel, Richard T. Woodward, William Neilson. (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.
- Banzhaf, H. Spencer, Randall P. Walsh. (2008). "Do people vote with their feet? An empirical test of tiebout's mechanism," *American Economic Review* 98 (3), 843-863.
- Bernheim, B. D., A. Rangel. (2009). "Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics," *Quarterly Journal of Economics* 124 (1), 51-104.
- Beshears, J., J. Choi, D. Laibson, B. Madrian. (2008). "How are preferences revealed?" *Journal of Public Economics* 92 (8-9), 1787-1794.
- Botvinick, Matthew M., Jonathan D. Cohen, Cameron S. Carter. (2004). "Conflict monitoring and anterior cingulate cortex: An update," *Trends in Cognitive Sciences* 8 (12), 539-546.
- Camerer, Colin, George Loewenstein, Drazen Prelec. (2005). "Neuroeconomics: How neuroscience can inform economics," *Journal of Economic Literature* 43 9-64.
- Cameron, Trudy Ann, J.R. DeShazo. (2009). "Demand for health risk reductions," Department of Economics, University of Oregon Working Paper.
- Cameron, Trudy Ann, Ian T. McConnaha. (2006). "Evidence of environmental migration," *Land Economics* 82 (2), 273-290.

- Carbone, Jared, Daniel Hallstrom, V. Smith. (2006). "Can natural experiments measure behavioral responses to environmental risks?," *Environmental and Resource Economics* 33 (3), 273-297.
- Carter, Cameron S., Vincent Van Veen. (2007). "Anterior cingulate cortex and conflict detection: An update of theory and data," 7 (4), 367-379.
- Chib, V. S., A. Rangel, S. Shimojo, J. P. O'Doherty. (2009). "Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex," *Journal of Neuroscience* 29 (39), 12315-12320.
- Crosson, P. L., M. E. Walton, J. X. O'Reilly, T. E. J. Behrens, M. F. S. Rushworth. (2009). "Effort-based cost-benefit valuation and the human brain," *Journal of Neuroscience* 29 (14), 4531-4541.
- Davies, Paul S., Michael J. Greenwood, Haizheng Li. (2001). "A conditional logit approach to U.S. state-to-state migration," *Journal of Regional Science* 41 (2), 337-360.
- Deschenes, Olivier, Michael Greenstone. (2007). "The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather," *The American Economic Review* 97 (1), 354-385.
- Deschenes, Olivier, Michael Greenstone, Jonathan Guryan. (2009). "Climate change and birth weight," *American Economic Review* 99 (2), 211-217.
- Deschenes, Olivier, Enrico Moretti. (2009). "Extreme weather events, mortality, and migration," 91 (4), 659-681.
- DeShazo, J. R., German Fermo. (2002). "Designing choice sets for stated preference methods: The effects of complexity on choice consistency," *Journal of Environmental Economics and Management* 44 (1), 123-143.
- Doswell III, Charles A., Harold E. Brooks, Nikolai Dotzek. (2009). "On the implementation of the enhanced Fujita scale in the USA," *Atmospheric Research* 93 (1-3), 554-563.
- Ebi, K. L. (2009). "Public health responses to the risks of climate variability and change in the United States," *Journal of Occupational and Environmental Medicine* 51 (1), 4-12.

- Ebi, K. L., D. M. Mills, J. B. Smith, A. Grambsch. (2006). "Climate change and human health impacts in the United States: An update on the results of the US national assessment," *Environmental Health Perspectives* 114 (9), 1318-1324.
- Evans, M. F., N. E. Flores, K. J. Boyle. (2003). "Multiple-bounded uncertainty choice data as probabilistic intentions," *Land Economics* 79 (4), 549-560.
- Fischer, G. W., M. F. Luce, J. M. Jia. (2000). "Attribute conflict and preference uncertainty: Effects on judgment time and error," *Management Science* 46 (1), 88-103.
- FitzGerald, Thomas H. B., Ben Seymour, Raymond J. Dolan. (2009). "The role of human orbitofrontal cortex in value comparison for incommensurable objects," 29 (26), 8388-8395.
- Greene, W. H., D. A. Hensher. (2007). "Heteroscedastic control for random coefficients and error components in mixed logit," *Transportation Research Part E-Logistics and Transportation Review* 43 (5), 610-623.
- Greene, William H. (2008). *Econometric analysis*. Upper Saddle River, N.J.: Prentice Hall.
- Grinband, Jack, Joy Hirsch, Vincent P. Ferrera. (2006). "A neural representation of categorization uncertainty in the human brain," 49 (5), 757-763.
- Haaijer, R., W. Kamakura, M. Wedel. (2000). "Response latencies in the analysis of conjoint choice experiments," *Journal of Marketing Research* 37 (3), 376-382.
- Harbaugh, William T., Ulrich Mayr, Daniel R. Burghart. (2007). "Neural responses to taxation and voluntary giving reveal motives for charitable donations," 316 (5831), 1622-1625.
- Hare, T. A., C. F. Camerer, A. Rangel. (2009). "Self-control in decision-making involves modulation of the vmPFC valuation system," *Science* 324 (5927), 646-648.
- Hare, Todd A., John O'Doherty, Colin F. Camerer, Wolfram Schultz, Antonio Rangel. (2008). "Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors," 28 (22), 5623-5630.
- Hensher, D. A. (2004). "Identifying the influence of stated choice design dimensionality on willingness to pay for travel time savings," *Journal of Transport Economics and Policy* 38 425-446.



- Hensher, D. A. (2006a). "How do respondents process stated choice experiments? Attribute consideration under varying information load," *Journal of Applied Econometrics* 21 (6), 861-878.
- Hensher, D. A. (2006b). "Revealing differences in willingness to pay due to the dimensionality of stated choice designs: An initial assessment," *Environmental & Resource Economics* 34 (1), 7-44.
- Hensher, D. A., J. Rose, T. Bertoia. (2007). "The implications on willingness to pay of a stochastic treatment of attribute processing in stated choice studies," *Transportation Research Part E-Logistics and Transportation Review* 43 (2), 73-89.
- Hensher, D. A., J. Rose, W. H. Greene. (2005). "The implications on willingness to pay of respondents ignoring specific attributes," *Transportation* 32 (3), 203-222.
- Hunter, L. M. (2005). "Migration and environmental hazards," *Population and Environment* 26 (4), 273-302.
- Jenkinson, Mark, Peter Bannister, Michael Brady, Stephen Smith. (2002). "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage* 17 (2), 825-841.
- Johnson, F. R. (2006). "Comment on "Revealing differences in willingness to pay due to the dimensionality of stated choice designs: An initial assessment", "*Environmental & Resource Economics* 34 (1), 45-50.
- Lancsar, E., J. Louviere, T. Flynn. (2007). "Several methods to investigate relative attribute impact in stated preference experiments," *Social Science & Medicine* 64 (8), 1738-1753.
- Landry, Craig E., Bin Okmyung, Paul Hindsley, John C. Whitehead, Kenneth Wilson. (2007). "Going home: Evacuation-migration decisions of Hurricane Katrina survivors," *Southern Economic Journal* 74 (2), 326-343.
- LeSage, James P., R. Kelley Pace. (2008). "Spatial econometric modeling of origin-destination flows," *Journal of Regional Science* 48 (5), 941-967.
- Li, C. Z., L. Mattsson. (1995). "Discrete-choice under preference uncertainty - an improved structural model for contingent valuation," *Journal of Environmental Economics and Management* 28 (2), 256-269.

- Louviere, J. J. (2001). "What if consumer experiments impact variances as well as means? Response variability as a behavioral phenomenon," *Journal of Consumer Research* 28 (3), 506-511.
- Louviere, J., D. Street, R. Carson, A. Ainslie, J. R. Deshazo, T. Cameron, D. Hensher, R. Kohn, T. Marley. (2002). "Dissecting the random component of utility," *Marketing Letters* 13 (3), 177-193.
- Louviere, J., K. Train, M. Ben-Akiva, C. Bhat, D. Brownstone, T. A. Cameron, R. T. Carson, J. R. Deshazo, D. Fiebig, W. Greene, D. Hensher, D. Waldman. (2005). "Recent progress on endogeneity in choice modeling," *Marketing Letters* 16 (3-4), 255-265.
- Luce, M. F., J. M. Jia, G. W. Fischer. (2003). "How much do you like it? Within-alternative conflict and subjective confidence in consumer judgments," *Journal of Consumer Research* 30 (3), 464-472.
- Malhotra, N. (2009). "Order effects in complex and simple tasks," *Public Opinion Quarterly* 73 (1), 180-198.
- Mas-Colell, Andreu, Michael D. Whinston, Jerry R. Green. (1995). *Microeconomic theory*. New York, NY: Oxford University Press.
- Merrell, David, Kevin M. Simmons, Daniel Sutter. (2005). "The determinants of tornado casualties and the benefits of tornado shelters," *Land Economics* 81 (1), 87-99.
- Mohr, Peter N. C., Shu-Chen Li, Hauke R. Heekeren. (2010). "Neuroeconomics and aging: Neuromodulation of economic decision making in old age," *Neuroscience & Biobehavioral Reviews* 34 (5), 678-688.
- Moll, Jorge, Frank Krueger, Roland Zahn, Matteo Pardini, Ricardo de Oliveira-Souza, Jordan Grafman. (2006). "Human fronto-mesolimbic networks guide decisions about charitable donation," *103* (42), 15623-15628.
- Myers, C. A., T. Slack, J. Singelmann. (2008). Social vulnerability and migration in the wake of disaster: The case of Hurricanes Katrina and Rita. 271-291.
- Paul, Bimal Kanti. (2005). "Evidence against disaster-induced migration: The 2004 tornado in north-central bangladesh," *Disasters* 29 (4), 370-385.
- Payne, J. W., Bettman, J. R., Johnson, E.J. (1993). *The adaptive decision maker*. New York: Cambridge Univ. Press.

- Pine, Alex, Ben Seymour, Jonathan P. Roiser, Peter Bossaerts, Karl J. Friston, H. Valerie Curran, Raymond J. Dolan. (2009). "Encoding of marginal utility across time in the human brain," 29 (30), 9575-9581.
- Plassmann, Hilke, John O'Doherty, Antonio Rangel. (2007). "Orbitofrontal cortex encodes willingness to pay in everyday economic transactions," 27 (37), 9984-9988.
- Pochon, Jean-Baptiste, Jason Riis, Alan G. Sanfey, Leigh E. Nystrom, Jonathan D. Cohen. (2008). "Functional imaging of decision conflict," 28 (13), 3468-3473.
- Rose, J. M., I. R. Black. (2006). "Means matter, but variance matter too: Decomposing response latency influences on variance heterogeneity in stated preference experiments," Marketing Letters 17 (4), 295-310.
- Rushworth, Matthew F. S., Timothy E. J. Behrens. (2008). "Choice, uncertainty and value in prefrontal and cingulate cortex," Nat Neurosci 11 (4), 389-397.
- Shannon, Claude E. (1948). "A mathematical theory of communication," Bell System Technical Journal 27 379-423 and 623-656.
- Sieg, H., V. K. Smith, H. S. Banzhaf, R. Walsh. (2004). "Estimating the general equilibrium benefits of large changes in spatially delineated public goods," International Economic Review 45 (4), 1047-1077.
- Simmons, Kevin M., Daniel Sutter. (2005). "WSR-88D radar, tornado warnings, and tornado casualties," Weather & Forecasting 20 (3), 301-310.
- Natural Hazards Center, University of Colorado at Boulder (2007) *The groundhog day Florida tornadoes: A case study of high-vulnerability tornadoes*. Natural Hazards Center, University of Colorado at Boulder,
- Smith, D. V., B. Y. Hayden, T. K. Truong, A. W. Song, M. L. Platt, S. A. Huettel. (2010). "Distinct value signals in anterior and posterior ventromedial prefrontal cortex," Journal of Neuroscience 30 (7), 2490-2495.
- Smith, Stephen M., Mark Jenkinson, Mark W. Woolrich, Christian F. Beckmann, Timothy E. J. Behrens, Heidi Johansen-Berg, Peter R. Bannister, Marilena De Luca, Ivana Drobnjak, David E. Flitney, Rami K. Niazy, James Saunders, John Vickers, Yongyue Zhang, Nicola De Stefano, J. Michael Brady, Paul M. Matthews. (2004a). "Advances in functional and structural MR image analysis and implementation in FSL," Neuroimage 23 (Supplement 1), S208-S219.

- Smith, V. K., H. Sieg, H. S. Banzhaf, R. P. Walsh. (2004b). "General equilibrium benefits for environmental improvements: Projected ozone reductions under EPA's prospective analysis for the Los Angeles air basin," *Journal of Environmental Economics and Management* 47 (3), 559-584.
- Smith, V. Kerry. (2008). "Risk perceptions, optimism, and natural hazards," *Risk Analysis: An International Journal* 28 (6), 1763-1767.
- Smith, V. Kerry, Jared C. Carbone, Jaren C. Pope, Daniel G. Hallstrom, Michael E. Darden. (2006). "Adjusting to natural disasters," *Journal of Risk and Uncertainty* 33 (1-2), 37-54.
- Vossler, C. A., R. G. Ethier, G. L. Poe, M. P. Welsh. (2003). "Payment certainty in discrete choice contingent valuation responses: Results from a field validity test," *Southern Economic Journal* 69 (4), 886-902.
- Welsh, M. P., G. L. Poe. (1998). "Elicitation effects in contingent valuation: Comparisons to a multiple bounded discrete choice approach," *Journal of Environmental Economics and Management* 36 (2), 170-185.
- Swait, J., W. Adamowicz. (2001a). "Choice environment, market complexity, and consumer behavior: A theoretical and empirical approach for incorporating decision complexity into models of consumer choice," *Organizational Behavior and Human Decision Processes* 86 (2), 141-167.
- Swait, J., W. Adamowicz. (2001b). "The influence of task complexity on consumer choice: A latent class model of decision strategy switching," *Journal of Consumer Research* 28 (1), 135-148.