UNIVERSITY OF CALIFORNIA,
IRVINE

Cross-Disciplinary Analyses Using Multi-Attribute Utility Theory

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Management

by

Jay Simon

Dissertation Committee:
Professor L. Robin Keller, Chair
Professor Donald Saari
Professor Shuya Yin

2009

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

ACKNOWLEDGEMENTS

I would like to enthusiastically acknowledge my committee chair and advisor, Professor L. Robin Keller, who provided me with invaluable guidance and support. She has shown me both by instruction and by example what it means to be a successful researcher, an effective teacher, and a collegial person. She has been an exceptional mentor, and without her helpful advice and diligence, it would not have been possible to write this dissertation. She has my utmost gratitude and appreciation!

I would also like to thank Professor Donald Saari, Professor Shuya Yin, Professor Rick So, and Professor Duncan Luce, for challenging me to improve and expand various areas of my research, and for always being willing and able to provide help and guidance whenever needed. They have been very supportive, and have contributed enormously to the very positive experience I have had in my doctoral studies.

I am also very grateful to my fellow doctoral students in Operations & Decision Technologies over the course of my years in the program: James Cao, Tianjun Feng, Yuhong He, Shaoxuan Liu, Wenting Pan, Sean Salleh, and Yitong Wang. It was a pleasure working with (and occasionally socializing with) all of them!

I am greatly indebted to Alladi Venkatesh for all of his work with the doctoral program at The Paul Merage School of Business, and to Noel Negrete, for her support in helping doctoral students navigate through a wide range of activities.

In addition, I would like to thank the Institute for Mathematical Behavioral Sciences (IMBS) at UCI, and all of the faculty and students involved with it. The IMBS seminars, colloquia, and conferences were extremely valuable in helping to shape and advance my research. IMBS also provided funding opportunities that were very much appreciated.

Additional financial support was provided by the University of California, Irvine, and through a doctoral dissertation grant from the National Science Foundation in Decision, Risk and Management Science (Award no. 0823458).

CURRICULUM VITAE

Jay Simon

Education

Ph.D.: Paul Merage School of Business, University of California, Irvine, Management. June 2009.

M.S.: Stanford University, Management Science and Engineering. June 2004.
B.S.: Stanford University, Mathematical and Computational Science. June 2003.

Publications

- Jay Simon, "Decision Making with Prostate Cancer: A Multiple-Objective Model with Uncertainty," forthcoming in *Interfaces*.

Manuscripts Under Review

- Jeffery Guyse and Jay Simon, "Consistency Among Elicitation Techniques for Intertemporal Choice: A Within-Subjects Investigation of the Anomalies"
- Jay Simon, Yitong Wang, and L. Robin Keller, "Paradoxes and Empirical Violations of Traditional Utility Theory"

Working Papers

- Jay Simon and L. Robin Keller, "Life Decisions with Health Outcomes"
- L. Robin Keller, Craig Kirkwood, and Jay Simon, "Decision Analysis Using Geographic Information Systems"
- Jay Simon and L. Robin Keller, "Modeling Altruistic Preferences"

Teaching Interests

- Decision Analysis
- Decision Making in Health Care
- Game Theory
- Management Science
- Statistics
- Operations Management

Teaching Experience

- Instructor:
  "*Business Statistics*," Spring 2008; 3.6/4.0 Instructor Rating
  "*Management Science*," Summer 2008; 3.7/4.0 Instructor Rating
- Teaching Assistant:
  "*Statistics for Management*," "*Management Science*," "*Quantitative Methods for Management*," "*Make Smart Choices*," "*Managing in Contemporary Organizations*"
- Experienced with Executive MBA, Health Care Executive MBA, Fully Employed MBA, Full-time MBA and Undergraduate students

Honors and Awards

- National Science Foundation Dissertation Grant, Social and Economic Sciences: Decision, Risk, and Management Sciences, 2008-2009
- Outstanding Teaching Assistant Award, upon vote of students in Full-Time MBA Class of 2008 (Statistics for Management)
- Sanli Pastore and Hill Excellence in Economics Writing Award, Third Place, 2009
- INFORMS Future Academician Colloquium participant, 2008
- Decision Sciences Institute Doctoral Student Consortium participant, 2008
- UCI Institute for Mathematical Behavioral Sciences Fellowship, 2006, 2007, 2008
- European Cooperation in the field of Scientific and Technical Research – International Doctoral School participant, Han-sur-Lesse, Belgium, Algorithmic Decision Theory, 2007
- UC Regents Fellowship, 2005

Manuscripts in Preparation

- Jay Simon and Frank L. Meyskens, "Analyzing Health Benefits of Chemoprevention Beyond the Treatment Period"
- Jay Simon and Nathan Fiala, "An Analysis of Nutrition Decisions Based on Multi-Attribute Utility"
- Jay Simon, "A Game-Theoretic Algorithm for Allocating Terrorism Defense Spending"

Technical Report

- Tianjun Feng, Nancy S. Jones, Lowell Kessel, Jay Simon, L. Robin Keller and Craig Kirkwood, "Survey of Key Stakeholder Evaluation Concerns and Tradeoffs for Central Arizona Water Resource Planning: Methods and Results," Technical Report, June 19, 2007, Decision Center for a Desert City, Global Institute of Sustainability, Arizona State University, P.O. Box 878209, Tempe, AZ 85287-8209.

Ongoing Research

- Jay Simon, "Choosing a Pricing Strategy for a Specialized Market"
- Jay Simon, "Health Insurance Premiums and Their Effects on Screening Decisions"
- Jay Simon, "Coalition Formation Games With Uncertainty"

Professional Experience

Quantitative Analyst, Prostate Cancer Decision Analyst (PCaDA), 6/2004-9/2005. Developed mathematical models to provide users with individualized probabilities of relevant events over time and to incorporate this information into a structured decision process.

Conference Presentations

- "Life Decisions with Health Outcomes[1]," INFORMS Western Regional Conference, Tempe, AZ, 2009.
- "Life Decisions with Health Outcomes[1]," INFORMS Annual Meeting, Washington, DC, 2008.
- "Decision Analysis Using Geographic Information Systems[2]," Society for Judgment and Decision Making Annual Meeting, Chicago, IL, 2008 (poster).
- "Analyzing Health Benefits of Chemoprevention Beyond the Treatment Period[3]," Frontiers in Cancer Prevention Research Annual Meeting, Washington, DC, 2008 (poster).
- "Decision Analysis Using Geographic Information Systems[2]," Southern California OR/OM Day, USC, 2008.
- "Life Decisions with Health Outcomes[1]," Institute for Mathematical Behavioral Sciences Graduate Conference, UCI, 2008.
- "Cross-Disciplinary Analyses Using Multi-Attribute Utility Theory[2]," The Paul Merage School of Business PhD Research Fest, UCI, 2008 (poster).
- "Multi-Attribute Utility in Classical Game Theory Models[1]," Behavioral Decision Research in Management Conference, UC San Diego, 2008 (poster).
- "Multi-Attribute Utility in Medical Decision Making[1]," Behavioral Decision Research in Management Conference, UC San Diego, 2008 (poster).
- "Preference Functions for Environmental Decisions with Spatially-Varying Attributes[2]," Focus the Nation Sustainability Showcase, UCI, 2008 (poster).
- "A Multiple-objective, Multiple-stakeholder Decision Analysis Approach for Water Resources Planning[4]," Focus the Nation Sustainability Showcase, UCI, 2008 (poster).

- "Decision Making Using Geographic Information Systems[2]," Institute for Mathematical Behavioral Sciences: Summer Fellowship Presentations, UCI, 2007.
- "Decision Making with Prostate Cancer," Society for Judgment and Decision Making Annual Meeting, Long Beach, CA, 2007 (poster).
- "Modeling Decision Situations with Spatially-Varying Attributes[2]," Society for Judgment and Decision Making Annual Meeting, Long Beach, CA, 2007 (poster).
- "Decision Analysis Using Geographic Information Systems[2]," Fifth Annual Institute for Mathematical Behavioral Sciences Graduate Conference, UCI, 2007.
- "Decision Making with Prostate Cancer," INFORMS Annual Meeting, Pittsburgh, PA, 2006.
- "Decision Making with Prostate Cancer," Institute for Mathematical Behavioral Sciences, UCI: Summer Fellowship Presentations, 2006.
- "Prostate Cancer Decision Analyst," INFORMS International, Hong Kong, 2006.
- "Prostate Cancer Decision Analyst," Fourth Annual Graduate Student Conference, UCI: Social Choice and Behavioral Sciences, 2006.

[1]Based on joint work with Keller
[2]Based on joint work with Keller and Kirkwood
[3]Based on joint work with Meyskens
[4]Based on joint work with Feng, Jones, Kessel, Keller, and Kirkwood

Service

- Session Leader: Creating Vision, Mesa Court Keystone Conference, UC Irvine, 2009
- Reviewer: *Decision Analysis*, *Risk Analysis*, Hong Kong Research Grants Council, Analytical Methods for Risk Management: A Systems Engineering Perspective, Paul Garvey (forthcoming)
- Founder, Merage Doctoral Student Summer Research Group
- Session Chair, Fifth Annual Institute for Mathematical Behavioral Sciences Graduate Conference, 2007

Professional References

- Dr. L. Robin Keller
  Professor, Operations and Decision Technologies
  Merage School of Business, University of California, Irvine
  Irvine, CA 92697-3125, LRKeller@uci.edu

- Dr. Donald G. Saari,
  Distinguished Professor: Mathematics and Economics
  Director: Institute for Mathematical Behavioral Sciences

University of California, Irvine
Irvine, CA 92697, DSaari@uci.edu
http://www.math.uci.edu/~dsaari

- Dr. Rick K. C. So
  Professor, Operations and Decision Technologies
  Merage School of Business, University of California, Irvine
  Irvine, CA 92697-3125, rso@uci.edu

- Dr. Shuya Yin
  Assistant Professor, Operations and Decision Technologies
  Merage School of Business, University of California, Irvine
  Irvine, CA 92697-3125, syin@merage.uci.edu

Professional Memberships

- Decision Analysis Society
- INFORMS
- Society for Judgment and Decision Making
- Society for Risk Analysis
- Decision Sciences Institute
- American Economic Association

ABSTRACT OF THE DISSERTATION

Cross-Disciplinary Analyses Using Multi-Attribute Utility Theory

By

Jay Simon

Doctor of Philosophy in Management

University of California, Irvine, 2009

Professor L. Robin Keller, Chair

This dissertation consists of three projects that apply multi-attribute utility theory to contexts in which there are unique structural relationships between the attributes. In the first project, we develop an initial model for maximizing expected utility in health decisions, based on the principles of quality-adjusted life years (QALY). We then extend this model to allow for adaptation to health states. We do this by maintaining a reference level for each possible health condition being considered. A particular formulation of the utility function and reference levels allows us to compute the a priori reference-dependent expected utility for any given alternative.

In the second project, we develop preference conditions for decisions made using Geographic Information Systems (GIS). These conditions allow us to simplify the elicitation of preferences over spatially-defined outcomes, and are based on standard results in multi-attribute utility theory. We then apply these tools to several specific decisions, in the contexts of urban development, irrigation, nuclear accident planning, and fire coverage.

In the third project, we formulate and develop utility structures that can effectively incorporate altruistic preferences. Specifically, we examine the utility implications of multiple individuals having altruistic tendencies toward one another. We explore utility structures with two altruistic individuals, and then expand our results to a more general model. The more general model can be greatly simplified by dividing the altruistic individuals up into groups, wherein the altruistic interactions are expressed at the group level. Our results can be expanded beyond altruism; they provide insight for any situation in which each individual outcome includes externalities that affect the outcomes experienced by others.

Chapter 1

Modeling Decisions with Multiple Health Conditions with Reference-Dependent Multi-Attribute Expected Utility

1.1 Introduction

In decision problems involving health outcomes, the decision maker is usually concerned with multiple attributes. These are a special class of multi-attribute decisions, however, due to the unusual structure of the outcomes that follow. The outcomes of these decisions occur over time rather than instantaneously, and in fact, the length of time over which they are experienced is itself one of the attributes. The decision maker's length of life is often a relevant issue, particularly for more serious decisions with health outcomes. In addition, the decision maker usually considers quality of life, which may be affected by several different factors pertaining to the decision.

A primary goal that a decision analyst would have when studying these situations is developing a useful and appropriate utility model. An associated dilemma is determining whether or not the required assumptions of the model are valid. Similar types of questions have been explored in the medical decision making literature; the primary model discussed in this paper is related to the concept of quality-adjusted life years (QALYs), as discussed by Pliskin, Shepard, and Weinstein (1980), Hazen (2004), Miyamoto (1999), and Torrance and Feeny (1989). We will discuss QALYs in detail later, as our initial model builds upon this idea.

When analyzing decisions with health outcomes, the most obvious application is in the medical domain. The first part of this work is, in some ways, a generalization and extension of the prostate cancer decision model discussed by Simon (2009). However, the models in this paper are appropriate for many types of personal life style decisions as well, such as choice of diet, exercise regimens, or consumption of alcohol. All of these decisions have outcomes which occur over time, with length of life as one of the relevant attributes.

In addition to our initial model based on the QALY concept, we also develop and discuss a reference-dependent expected utility model for decisions with health outcomes. This builds on the work of Baucells and Sarin (2007, 2008). We extend their work to a model with expectation, and apply it in a somewhat different context. The reference-dependent expected utility model will allow the decision maker to maximize the expected level of experienced utility, incorporating the idea that a person can adapt over time to changes in health state.

1.2 Literature Review

There are several streams of literature related to this paper. First, the literature dealing with multi-attribute utility is extremely important. Without many of the concepts from these papers, there would be little basis for the model developed here. The second stream deals with QALYs, and associated concepts in medical decision making. The third stream discusses issues with preference structures relevant to decisions with health outcomes. There is some overlap here, as there are many papers which analyze the preference structures implied by the use of QALYs. Finally, there is also a body of related work on reference-dependent utility models.

The development of multi-attribute utility theory can be attributed largely to a few influential works from the 1970's. Keeney and Raiffa (1976) provide utility independence conditions for multiple-attribute utility (for decisions under risk). Dyer and Sarin (1979) discuss several conditions under certainty, most notably mutual preferential independence, which permit the use of additive multi-attribute measurable value functions (for decisions under certainty). Keeney (1974) discusses the use of a multiplicative model. Cho, Keller, and Cooper (1999) and Eriksen and Keller (1993) apply these decision analysis approaches to situations involving health risks.

Currently, the most common method of expressing preferences for medical decisions is the quality-adjusted life years (QALY)[1] method. It is discussed in detail by Pliskin, Shepard, and Weinstein (1980), Hazen (2004), Miyamoto (1999), Torrance and Feeny (1989), and others. Hazen uses a straightforward multiplicative model that is very attractive from a normative perspective. It implies that the achieved proportion of a

[1] A related concept to measure adverse health conditions is the disability-adjusted life years (DALY) method. This method is discussed in detail by Murray (1994).

person's well-being is the product of the proportions achieved in each attribute. Miyamoto introduces several other possible utility functions using the QALY concept. Most of these are prescriptive rather than purely normative models, and are developed to fit various types of preferences observed in individuals. Torrance and Feeny expand upon the basic QALY models by considering the aggregation of utilities across a group. Simon (2009) uses a QALY-based decision model to help prostate cancer patients choose between various treatment alternatives.

The QALY concept requires some assumptions on the structure of the patient's preferences. Bleichrodt, Wakker, and Johannesson (1997) examine a case which requires only a very weak risk-neutrality condition. Bremner et al. (2007) directly assess utilities of various combinations of conditions, thus avoiding some potential violations of these independence assumptions. Feeny et al. (1995) explain that a large number of broadly-defined attributes may lead to violation of the assumptions. Bleichrodt and Johannesson (1997) examine which of the QALY assumption are likely to hold descriptively.

Reference-dependent utility is based on the idea popularized by Kahneman and Tversky (1979) that people view outcomes not in terms of their raw value, but as gains or losses relative to a reference point. Littman and Ackley (1991) provide evolutionary justification for the existence of this type adaptive utility in humans. Kahneman and Thaler (1991) discuss adaptation to income level. Recent applications of reference-dependent utility can be seen in work by Bleichrodt (2007) and Munro and Sugden

(2003). A general model using these concepts is discussed by Wathieu (2004). Our model, however, is most closely based on the reference-dependent models used by Baucells and Sarin (2007, 2008) in their work on utility of consumption.

1.3 Model Development

There are many different types of utility functions. Since health decisions have outcomes which are often described differently than those resulting from decisions in other contexts, it is important to make sure the decision maker is using an appropriate preference structure. In this section, we discuss the possible use of an additive utility function, and a utility function based on QALYs.

1.3.1 Additive Utility

The use of an additive utility function is preferred if possible for these decisions, since it allows for assessment of preferences on individual attributes without considering the values of the others. However, it turns out that preferences over outcomes in this context are generally not additive independent, and therefore cannot be represented by an additive utility function. To understand intuitively the reason for this, notice that an extra year of life will be valued less if the person is afflicted with a condition that seriously reduces quality of life. More formally, the person will care about the joint distribution of attribute values, not just the marginal distributions.[2] Pliskin, Shepard, and Weinstein (1980) discuss this concept in greater detail. Since the simplest type of multi-attribute function cannot be used for these decisions, a more complex function must be developed and justified.

1.3.2 Quality-Adjusted Life Years Model Extension

One useful concept that has been widely discussed in the literature is quality-adjusted life years (QALYs). QALYs are a helpful tool for analyzing the tradeoff between length and quality of life. They involve expressing the value of a t-length experience with a particular condition as a proportion of the value of a t-length experience without the condition. This yields the proportion of utility retained when the condition occurs. Our work extends the QALY concept by incorporating the possible occurrence of various conditions as probability curves over time.

The basic QALY model used by Hazen (2004) takes the form $U\left(y_1,...,y_n,t\right)=t\prod_{i=1}^{n}y_i$.

This expression indicates the utility achieved during an interval of length t with n health conditions present, having "levels" $y_1,...,y_n$, where y_i is the proportion of utility retained with condition i. However, when considering a medical alternative, conditions often do

[2] It is important to distinguish between real-world correlation of attribute values, and the preferences over the joint distribution of attribute values. The latter can be defined in the absence of any decision situation.

not occur with certainty. In addition, the length of the interval is uncertain; the patient does not know the precise length of remaining life. The basic QALY model as described in Hazen (2004) can be extended accordingly, resulting in the following expression for the patient's expected utility:

$$EU = \int_0^T s(t)\left(1 - \prod_i p_i(t)(1 - y_i)\right)dt, \tag{1.1}$$

where *T* is the longest possible life span of the patient. *s(t)* represents the probability that the patient will be alive at time *t*, $p_i(t)$ represents the probability that condition *i* will be present at time *t*, and has the same meaning as in the basic model. This expression uses 1 as the utility of the ideal condition at any instant in time. Each condition with non-zero probability contributes to the total reduction in expected utility at that instant.

(1.1) allows the decision maker to rigorously account for preferences over the various health conditions, for the effects of the decision on likelihood of each condition occurring over time, and for the effects of the decision (if any) on the survival probability over time. This normative model asserts that when faced with a life decision with health outcomes, the decision maker should choose the alternative which would result in the highest expected utility in (1.1).

1.3.3 Required Conditions for Extended QALY Model

Using and extending the QALY concept in this manner also gives rise to a number of independence conditions and requirements that must be verified. Any normative decision model will impose certain restrictions on the structure of the decision maker's preferences. These conditions and requirements will provide a guideline to help determine and understand the types of situations in which the decision model is valid.

The first requirement of note is covered by Hazen (2004). The intuitive goal is that preferences over gambles on length of life and a single health/lifestyle effect (such as lack of appetite) must not depend on the presence or absence of other effects (such as need for eye glasses). Hazen expresses this requirement using two equivalent conditions: standard-gamble independence and time-tradeoff independence. Both are methods of constructing gambles over y_i and t in the basic QALY formulation, and asserting that they must not depend on the other attribute values. Intuitively, these conditions should be met provided that the effects on well-being from the attributes do not overlap. For example, there is no apparent reason that the effects of decreased vision quality and reduced appetite should depend on one another. In a much more complex model, as discussed by Feeny et al. (1995), a large number of broadly-defined attributes might prevent the decision maker's preferences from satisfying this type of independence. However, if we assume that the problem is small and/or specific enough such that these independence conditions are satisfied, then preferences can be assessed on individual attributes without considering the values of the others. It is also important to consider the ranges of attribute values over the possible outcomes. If the feasible ranges are limited, these independence conditions might be reasonable even for broader decisions.

One caution when relying on this assumption is that it is unlikely to hold for very serious health states. Stalmeier, Wakker, and Bezembinder (1997) and Dolan and Stalmeier (2003) discuss alternate ways of modeling such states using a concept called "maximum endurable time" (MET). MET implies that after a certain length of time in a health state (having a migraine, for example), that state is actually less desirable than death. In scenarios approaching that length of time in this state, tradeoff questions with regard to length of life and other attributes become unreliable. In general, the first assumption of our model is likely to hold if the potential health conditions do not have significant overlap, and are not unusually detrimental.

The second condition is that preferences must be linear in probabilities. That is, a .2 probability of a side effect must result in a utility decrease which is twice that of a .1 probability of the side effect. This seems trivial, but actual human behavior often violates this assumption, as analyzed by Kahneman and Tversky (1979). Specifically, people often deviate from it greatly for very small probabilities. Whether or not to account for this type of human interpretation of probability is a philosophical dilemma for an analyst aiding a health-related decision. In normative models, it is standard to assume that the condition holds.

The third condition is time additivity. Time additivity states that:

$$\text{if the intervals } \left[t_a^1, t_b^1\right] \text{ and } \left[t_a^2, t_b^2\right] \text{ are disjoint, then}$$

$$\text{the utility achieved in } \left[t_a^2, t_b^2\right] \text{ does not depend on the utility achieved in } \left[t_a^1, t_b^1\right].$$

Simply put, this means that the patient's utility in year 10 can be assessed purely using the outcome over year 10, and ignoring anything that occurred in years zero through nine. This precludes any type of adaptation considerations. Adaptation is not widely applied in the medical decision making literature, but may be an important factor in some decision problems with health outcomes, as discussed in the next section.

1.4 Adding Adaptation to a Utility Model

Adaptation is an intriguing concept in preference modeling: one way of depicting adaptation would be to allow the weights that the decision maker places on attributes in an additive multi-attribute model to change over time. A recent paper by Keeney and Vernik (2007) presented the idea of a woman's weights changing over time on multiple objectives contributing to professional life, family life, and social life. The notion that a person's weights on objectives may (and often will) change over time is novel and has not been examined, in general, for multi-attribute preference models for decisions involving health outcomes. Cho, Keller, and Cooper (1993) suggest that such changes over time must be incorporated to fully represent preferences in decisions under risk. Generally speaking, the anticipated utility effect of a health condition is not necessarily equivalent to the effect on utility when the condition actually occurs. Chestnut et al. (1988) mention this phenomenon in the context of angina attacks, but in general, it has not yet been rigorously studied in the decision analysis literature.

1.4.1 Change the Proportion of Utility Retained Over Time

One potential way to incorporate adaptation into the initial utility model is to allow y_i (which is the proportion of utility retained when in condition *i*) to increase over time. That is, we can interpret the decision maker's preference assessment for condition *i* as being the proportion of utility retained immediately after the condition occurs, and then assume it will increase as (s)he "adapts" to the condition. Essentially, this method asserts that the decision maker will experience the condition as being less severe over time, in terms of disutility. The challenging step in this method is determining exactly how much and how quickly y_i increases. Since we are attempting to correct for an implicit projection bias of the decision maker, it is impossible to directly elicit this as an individual preference. Instead, we would have to rely on empirical data from typical preference changes over time for that specific health state. This is a practical problem, as such data are difficult to obtain and not normally available.

1.4.2 Reference-Dependent Utility Model

A more elegant method of incorporating adaption is the use of a reference-dependent utility model. This is fundamentally different from our initial model, in that instead of the decision maker directly experiencing the utility described in Equation (1), (s)he will experience the difference between this utility and some reference utility level R_t. That is, if we let U_t be the utility at time t as used in the initial model, the reference-dependent utility at time t is $(U_t - R_t)$. The origin of such models can be traced back to early work on prospect theory (Kahneman and Tversky, 1979), and recent applications can be seen in work by Bleichrodt (2007) and Munro and Sugden (2003). This method is fundamentally different from simply allowing y_i to increase over time, in that it asserts that the individual becomes accustomed to a utility level rather than to a particular condition.

1.4.2.1 Utility at Time t with a Single (Scalar) Reference Level

By expanding U_t, we can write the reference-dependent utility at time t as

$$\left(\prod_i \left(1 - X_i(t) * (1 - y_i)\right)\right) - R_t, \tag{1.2}$$

where y_i has the same meaning as it did in the initial model, and $X_i(t)$ is an indicator variable equal to 1 if condition i is present at time t, and equal to 0 otherwise. This expression gives us the difference between the QALY measurement of utility at time t and the reference utility level at time t.

The difficult question, of course, is: how do we determine the value of R_t? This is an unresolved question not only in decisions with health outcomes, but in any applications of reference-dependent utility models. The simplest method is to just let $R_t = U_{t-1}$. This implies that the utility actually experienced is the difference in utility from last period to this one. That is, it implies that adaptation is complete and immediate. This is probably not realistic, particularly for serious health states that involve a major change in lifestyle. Our initial model is the opposite of this method; considering only the absolute utility level is equivalent to using a reference-dependent model in which R_t is fixed. It is difficult to force R_t to follow a specific pattern, since it is likely to progress differently for various health states.

The most general method would be to simply assert that R_t is between R_{t-1} and U_{t-1}. This idea is discussed by Wathieu (2004). While descriptively accurate, this is not a very powerful assertion; it simply claims that some adaptation is occurring, but that we have no idea how much. For use in practical applications, it will be very difficult to apply a reference-dependent utility model without specifying some sort of function or pattern for R_t. However, it is intuitive that any specific model used should certainly conform to this general requirement.

One intuitively appealing method for determining reference levels is via exponential smoothing. Baucells and Sarin (2008) apply this exponential smoothing method for reference levels to happiness models in consumption. This method asserts that

$$R_t = \alpha R_{t-1} + (1-\alpha) U_{t-1}, \tag{1.3}$$

where $0 \le \alpha \le 1$. The parameter α in (1.3) represents the speed of adaptation to changes in utility level. For α close to 1, the reference level changes very slowly, meaning that utility is compared not only to recent experiences, but also to utility levels from farther in the past. For α close to 0, the reference level is determined largely by utility levels from the very recent past. In the context of this paper, the higher the value of α, the slower the individual will adapt to changes in health state.

1.4.2.2 Utility at Time t with Reference Levels for Each Health State

The use of a reference-dependent utility model does not circumvent all of the conditions required by the initial model. However, it does provide a method of incorporating adaptation to health states, which is a widely recognized characteristic that humans possess. The main drawback of this model above is that the reference utility level is scalar; it depends on the overall utility level, but not (directly) on which of the possible health states were actually present. This runs somewhat contrary to the original motivation for using reference-dependent utility in the first place. Fortunately, we can combine the two previously discussed methods into one that is tractable and also captures the original intent of adapting to specific health states. The new model will also allow us to construct a more sensible expected utility function.

In this new model, instead of using a single reference utility level R_t, we consider reference levels R_{it} for all individual health states. Accordingly, we also consider smoothing parameters α_i for all individual health states. It will be convenient to define a new variable Y_{it} as follows:

$$Y_{it} = \begin{cases} 1, & \text{if condition } i \text{ is absent} \\ y_i, & \text{if condition } i \text{ is present} \end{cases} \quad \text{at time } t. \tag{1.4}$$

Y_{it} represents the utility multiplier applied for condition i at time t in the initial model. Equivalently, we could define Y_{it} as $Y_{it} = 1-(X_i(t)(1-y_i))$. Using this new variable, we can express R_{it} as:

$$R_{it} = \alpha_i R_{it-1} + (1-\alpha_i) Y_{it-1}. \tag{1.5}$$

The reference levels defined in (1.5) tell us exactly the proportions of retained utility associated with each health condition to which the person is accustomed at any given time. Note that the α_i may differ; a person may adapt to different conditions more quickly than others.

We can now write a reference-dependent utility multiplier for condition *i* as:

$$1 + (Y_{it}\text{-}R_{it}) \tag{1.6}$$

(1.6) represents the perceived level of condition *i* at time *t*, expressed as a utility multiplier in the same way we used y_i earlier. For example, if the condition is present, y_i=.85, and the reference level is .95, then we will multiply the person's base utility by .9. If the condition is absent, and the reference level is .95, then considering this condition implies that the person's base utility should be multiplied by 1.05 at time *t*. Notice that (1.6) will often be greater than 1 if the condition is absent. This is consistent with our idea of adaptation; when a negative health condition goes away, a person will be at least temporarily happier than (s)he was before the negative condition occurred. Using (1.6), we can write the experienced utility at time *t* as $\prod_i \left(1+\left(Y_{it}-R_{it}\right)\right)$.

1.4.2.3 Required Conditions for the Reference-Dependent Expected Utility Model

Our primary purpose for using the exponential smoothing model and the multiplier format in (1.6) is that they will allow us to construct a reference-dependent expected utility model for health decisions. Before developing this model, we assume two things. First (as before), we assume that the conditions occur independently of one another. That is, $P(X_i(t)=1 \mid X_j(0)=x_0, X_j(1)=x_1,\ldots, X_j(t)=x_t) = P(X_i(t))=1$, for all i, j, t, x_k. Second, we assume that the survival probability $s(t)$ is independent of Y_{it} for all i, t. These assumptions are necessary in order to make the model tractable; without them we would have to incorporate the joint probability distributions, which are difficult to manage and rarely available in practice. This expected utility model should not be used if there is significant dependence between the health conditions, or between any included health condition and survival.

1.4.2.4 Development of the Reference-Dependent Expected Utility Model

First, consider only a single health condition at time t, with $s(t)$=1. We would like to be able to express the reference-dependent expected utility as a function of the probabilities that $X_1(t) = 1$. If we can do this, then we do not need to worry about joint distributions. That is, we would like the expected utility computation not to be "path-dependent." In fact, using the properties of exponential smoothing and the definition of Y_{it}, we see that this is indeed possible (see Appendix). This means that knowing the probabilities of the health condition being present at each point in time is sufficient for computing reference-dependent expected utility (after assessing the preferences of the decision-maker, of course).

Now we would like to extend this to the case of multiple health conditions. Consider E[U$_t$]. From our reference-dependent model, $\mathrm{U}_t = s(t)\prod_i \left(1+\left(Y_{it}-R_{it}\right)\right)$. As shown in the Appendix, E[R_{it}] is a function of E[Y_{i0}], E[Y_{i1}], …, E[$Y_{it\text{-}1}$]. Specifically,

$$\mathrm{E}[R_{it}] = \sum_{k=0}^{t-1} \alpha_i^{t-1-k} (1-\alpha_i) \mathrm{E}[Y_{ik}]. \tag{1.7}$$

Since we have assumed that $s(t)$ and all Y_{it} are independent, it must be true that E[U$_t$] is also a function of the marginal probabilities of each condition occurring in each given time period.. As in the single-condition case, we have $\mathrm{E}[\mathrm{U}] = \sum_t \mathrm{E}[\mathrm{U}_t]$. This means we can write E[U] as:

$$\sum_t \left(s(t) \prod_i \left(1 + E[Y_{it}] - \sum_{k=0}^{t-1} \alpha_i^{t-1-k} (1-\alpha_i) \mathrm{E}[Y_{ik}] \right) \right). \tag{1.8}$$

(1.8) is a closed-form reference-dependent expected utility formulation, and it depends only on the marginal probabilities of each health condition over time, the survival curve, and the preferences of the decision maker (the y_i and α_i).

1.5 Example of Reference-Dependent Multi-Attribute Expected Utility Applied to Pregnancy Planning

As an example, we consider the decision analyzed by Keeney and Vernik (2007) regarding the age at which a woman attempts to become pregnant. They consider the woman's utility as consisting of professional life utility (P), social life utility (S), and family life utility (F). We will use the stylized data presented in their paper. However, we will frame these three utility components as multiplicative factors varying between 0 and 1, consistent with the structure we have used thus far. They allow the relative weights placed on the three components to vary over time. For simplicity, we omit that in our example, but as we will discuss later, adaptation can still produce the desired qualitative effect.

Consider a 22-year-old woman at her first job choosing an age at which to begin attempting to have a child. She is considering beginning at any age between 22 and 35. As in Keeney and Vernik's model, she will continue trying for up to four years. Figure 1.1 shows an example of the component utility curves if she begins working at age 22, enters an MBA program at age 26, and either becomes pregnant at age 31, or not at all. These curves are estimated using the examples presented by Keeney and Vernik. The model is discrete (year-by-year), and overall utility is determined by taking the summation of the utilities achieved through age 75.

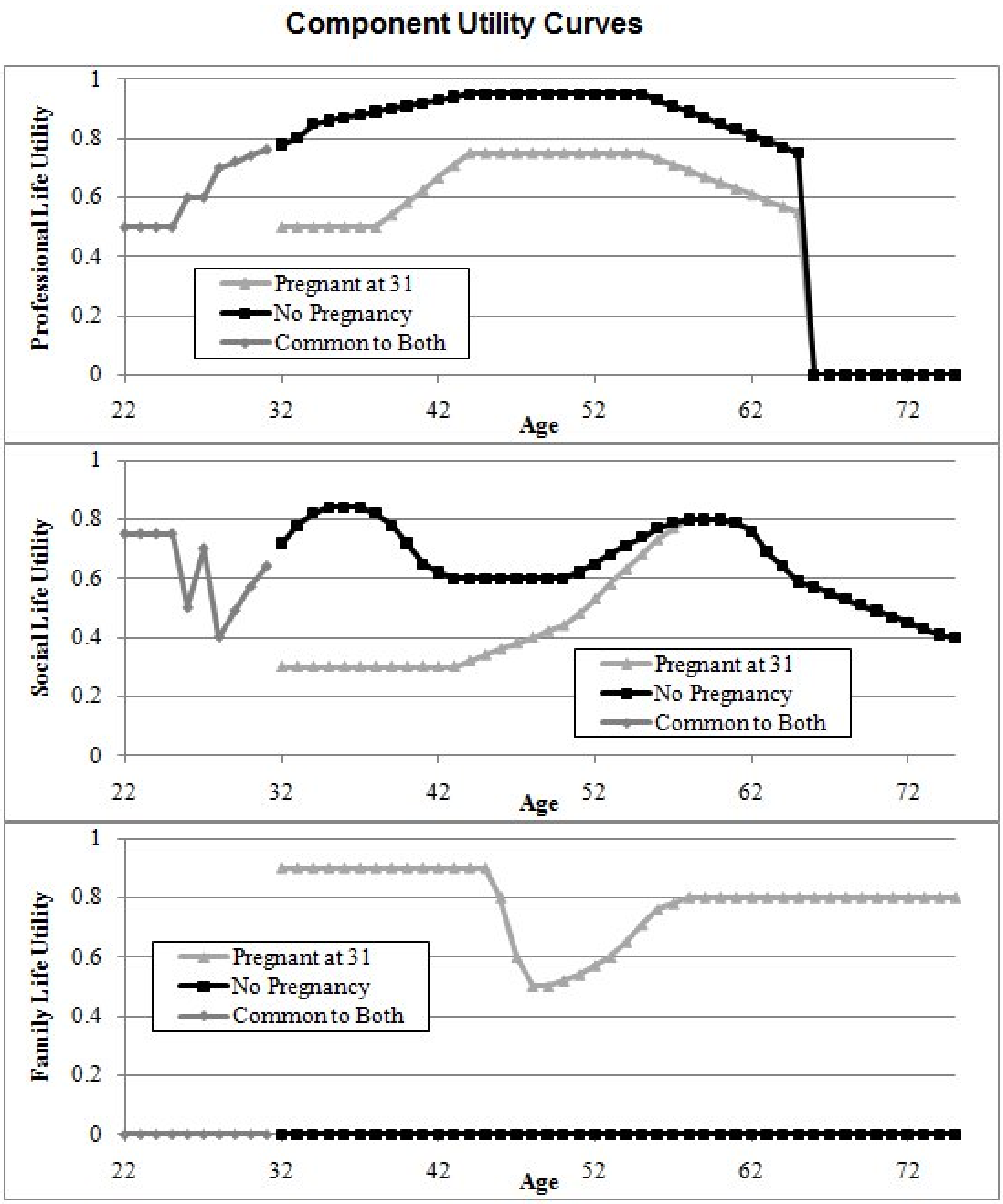

Figure 1.1. An example of component utility curves with and without a child, based on data from Keeney and Vernik (2007).

We use six parameters in our model: relative importance weights on professional life, social life, and family life (which sum to one), and exponential smoothing parameters (ranging from zero to one) for computing the reference level of each of the three utility components. We use the importance weights to scale the utility multipliers for each of the three components. Intuitively, the weights determine how strong the effect of a change in that area will be, and the smoothing parameter determines how quickly the utility effect of the change dissipates.

As a base case, we weight all three components equally, and use α=.7 (modest adaptation) for all three components. In the base case, the highest expected utility achieved using the reference-dependent model occurs when this woman begins trying to become pregnant at age 28. The expected utilities are shown in Figure 1.2. (The drop at age 27 occurs because of the erratic social life utilities assigned during the MBA program, as can be seen in Figure 1.1.) Figure 1.3 and Figure 1.4 show the results of one-way sensitivity analysis on each of the six parameters. This allows us to see how each parameter affects the optimal age.

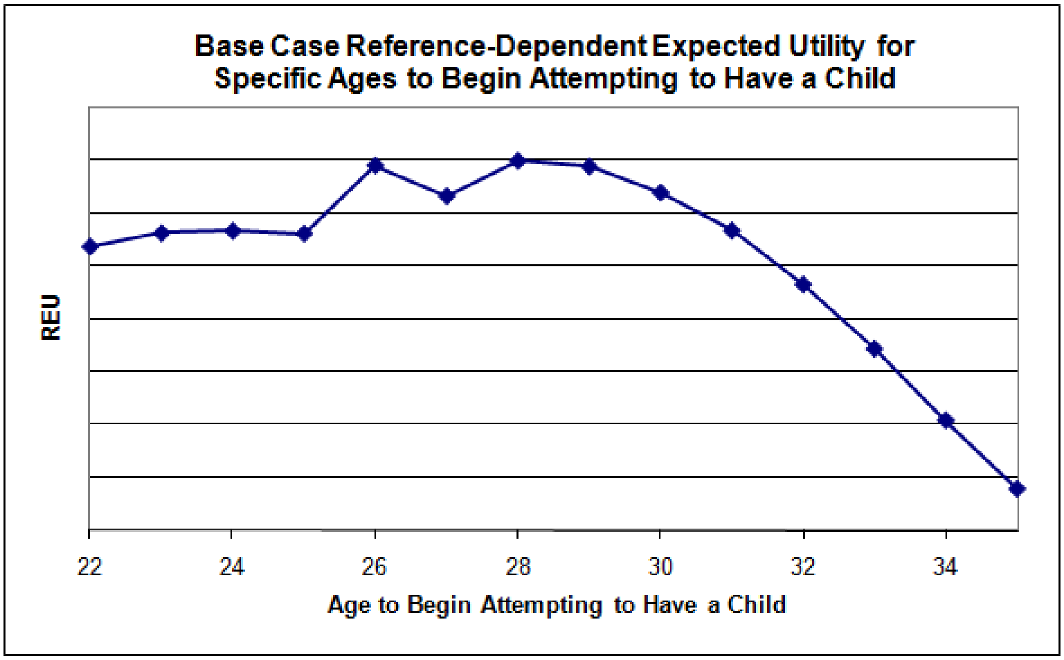

Figure 1.2. Reference-dependent expected utility for various ages at which a particular woman begins attempting to have a child.

Figure 1.3 gives us some interesting insights. First, as a disclaimer, the right side of the graph (which has extreme results) is omitted. In reality, if a woman is even considering this decision problem, it is very unlikely that she places nearly all of the importance on one component. The base case occurs at .33, where the three curves intersect. It is most informative to look at how each curve behaves as we move away from .33. The weight on social life has the effect we would guess: the more the woman cares about her social life, the longer she should wait to have a child. Trying to have a child at around age 30 is optimal if she cares mostly about her social life.

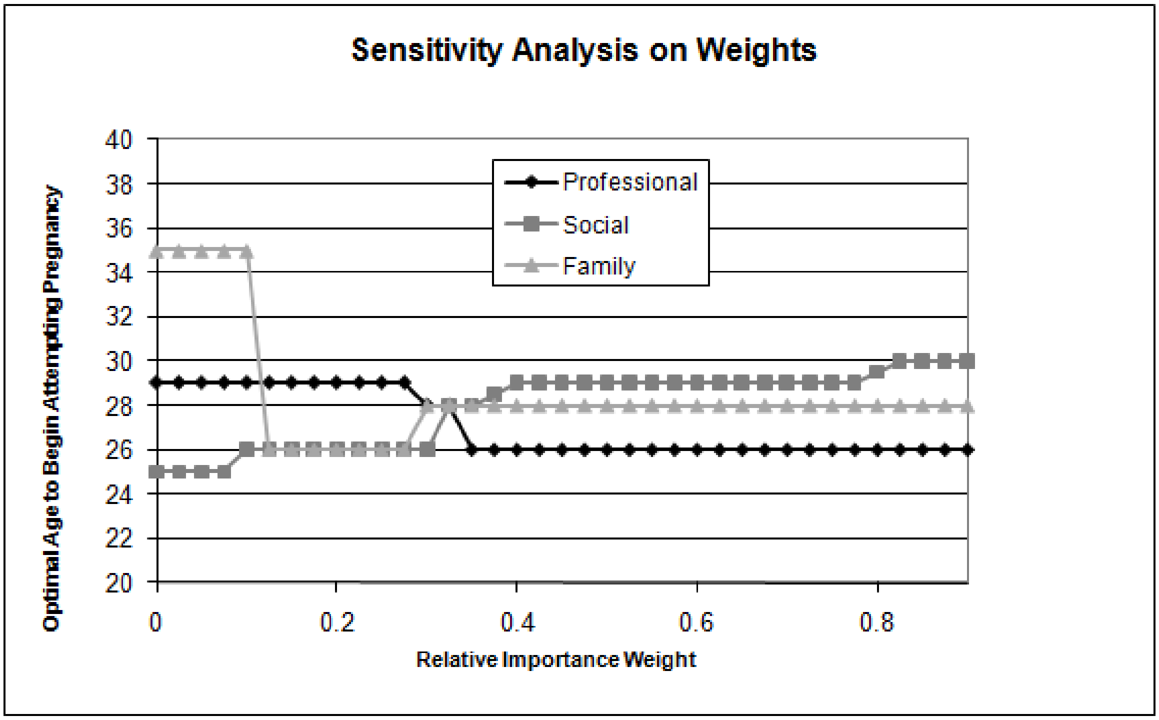

Figure 1.3. One-way sensitivity analysis for the importance weights that a particular woman places on the three utility components.

The other two parameters do NOT have the effects we would expect; this is a result of incorporating adaptation. The total utility achieved from family life is nearly identical when any of the younger ages is chosen. This is because for any of these ages, the woman will experience her children growing up and moving away, and then fully adapt to life without them. The difference in raw utility levels achieved before having children versus after they have moved away becomes irrelevant. As for professional life, provided the woman is able to enter the desired MBA program (and thus improve her career), the degree of raw improvement is less relevant than one might think. This is because she will adapt over time to whatever level she achieves. Thus, because of the pattern with which each of the utility components develops over time, introducing adaptation makes one (social life) generally more relevant than the other two.

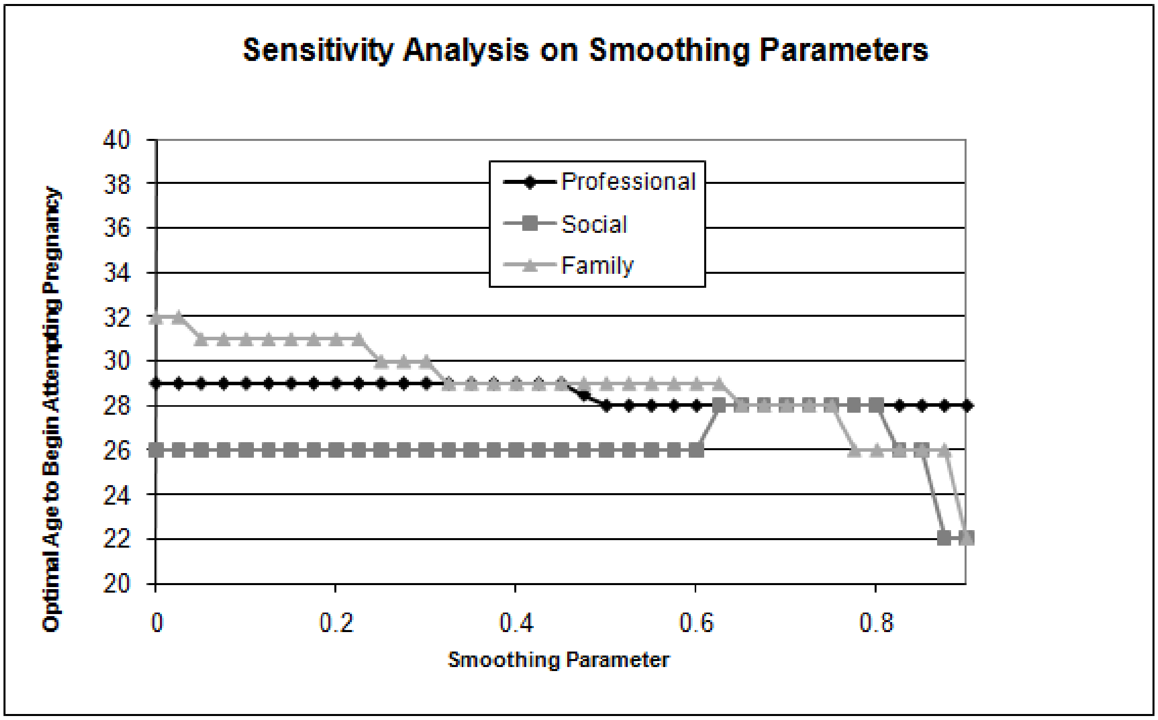

Figure 1.4. One-way sensitivity analysis for the smoothing parameters that a particular woman assigns to the reference levels of the three utility components.

The base case for the smoothing parameters is where all three curves in Figure 1.4 intersect, at α=.7. As in Figure 1.3, the far right end of the graph is omitted. The far right end represents the cases in which the woman does not adapt at all to a single component, and this again leads to extreme results. The strongest effect in Figure 1.4 is observed in the speed at which this woman adapts to changes in family life. The more slowly she adapts, the stronger her incentive will be to become pregnant at a young age, since the raw level of family life utility is much higher after having a child. When we include adaptation in the analysis, the overall degree to which the woman cares about family life is less relevant than the degree to which she becomes accustomed to it. This effect is similar to what we might observe if we allowed the woman's importance weights to change over time. Allowing her weight on family life to increase over time will have the same material effect we would observe from her smoothing parameter for family life having a high value.

1.6 Conclusion

We have developed and discussed a few different methods for analyzing decisions with health outcomes. We can extend the QALY method to include probability curves over time for each possible health condition, which allows the computation of a subjective expected utility for each alternative. This model follows the axioms of traditional utility theory, and can be very useful as a guide for decision makers. We also introduced a reference-dependent utility model, which allows for the fact that a person will adapt to health states over time. Using an exponential smoothing formulation for the reference points, we can compute reference-dependent expected utility. This represents the expected level of actual experienced utility over the time horizon being considered. Given a few basic independence assumptions, we can compute this reference-dependent expected utility using only the probabilities of the health conditions over time, the preferences of the decision maker, and (if appropriate) a survival curve. This is an extremely valuable tool for a person facing a health-related decision. Finally, we applied this model to Keeney and Vernik's analysis of when a woman should begin trying to have a child, and discussed the ways in which adaptation can affect that particular decision.

1.7 Appendix

Proof that single-condition reference-dependent expected utility depends only on the marginal probabilities that $X_1(t) = 1$:

Recall that $Y_{it} = 1-(X_i(t)(1-y_i))$. Since Y_{it} is a linear transformation of $X_i(t)$, it will suffice to show that expected utility is a function of the $E[Y_{it}]$.
Index the single health condition as i=1. We can write $E[U_t] = E[1 + (Y_{1t}-R_{1t})] = E[1] + E[Y_{1t}] + E[R_{1t}]$.
Thus, we need to show that $E[R_{1t}]$ is a function of $E[Y_{10}], E[Y_{11}], \ldots, E[Y_{1t-1}]$.
Since we are using exponential smoothing to determine the reference level, R_{1t} is defined recursively using (5). Solving the recursive equations yields $R_{1t} = \sum_{k=0}^{t-1} \alpha_1^{t-1-k}(1-\alpha_1)Y_{1k}$. Thus,

$E[R_{1t}] = E\left[\sum_{k=0}^{t-1} \alpha_1^{t-1-k}(1-\alpha_1)Y_{1k}\right] = \sum_{k=0}^{t-1} \alpha_1^{t-1-k}(1-\alpha_1)E[Y_{1k}]$, and we see that $E[R_{1t}]$ is a linear function of $E[Y_{10}], E[Y_{11}], \ldots, E[Y_{1t-1}]$. In addition, since we are assuming $s(t)$=1 over the time horizon being considered, observe that $E[U] = \sum_t E[U_t]$, which means E[U] is also linear in $E[Y_{10}], E[Y_{11}], \ldots, E[Y_{1t-1}]$. It does not even need to be linear; the fact that it is a function only of the $E[Y_{it}]$ is sufficient.

1.8 References

Baucells, M., R. Sarin. 2007. Satiation in Discounted Utility. *Operations Research* **55**(1) 170-181.
Baucells, M., R. Sarin. 2008. Does more money buy you more happiness? *Decision modeling and behavior in complex and uncertain environments* 199-226. Springer.
Bleichrodt, H. 2007. Reference-dependent utility with shifting reference points and incomplete preferences. *Journal of Mathematical Psychology* **51**(4) 266-276.
Bleichrodt, H., M. Johannesson. 1997. The Validity of QALYs: An Experimental Test of Constant Proportional Tradeoff and Utility Independence. *Medical Decision Making* **17**(1) 21-32.
Bleichrodt, H., P. Wakker, M. Johannesson. 1997. Characterizing QALYs by risk neutrality. *Journal of Risk and Uncertainty* **15** 107-114.
Bremner, K. E., C. A. K. Y. Chong, G. Tomlinson, S. M. H. Alibhai, M. D. Krahn. 2007. A Review and Meta-Analysis of Prostate Cancer Utilities. *Medical Decision Making* **27**(3) 288-298.
Chestnut, L.G., Colome, S.D., Keller, L.R., Lambert, W., Ostro, B., Rowe, R.D. and Wojciechowski, S.L. 1988. Heart disease patients' averting behavior, costs of illness, and willingness to pay to avoid angina episodes. In: Document No. EPA-230-10-88-042, final report prepared for US Environmental Protection Agency.

Cho, Y. H., L. R. Keller, and M. L. Cooper. 1999. Applying Decision-Making Approaches to Health Risk-Taking Behaviors: Progress and Remaining Challenges. *Journal of Mathematical Psychology* **43**(2) 261-285.

Dolan, P., P. Stalmeier. 2003. The validity of time trade-off values in calculating QALYs: constant proportional time trade-off versus the proportional heuristic. *Journal of Health Economics* **22**(3) 445-458.

Dyer, J. S., R. K. Sarin. 1979. Measurable multiattribute value functions. *Operations Research* **27**(4) 810-822.

Eriksen, S., L. R. Keller. 1993. A Multiattribute-utility-function Approach to Weighing the Risks and Benefits of Pharmaceutical Agents. *Medical Decision Making* **13**(2) 118-125.

Feeny, D., W. Furlong, M. Boyle, G. W. Torrance. 1995. Multi-attribute health status classification systems: health utilities index. *Pharmacoeconomics* **7**(6) 490-502.

Hazen, G. B. 2004. Multiattribute structure for QALYs. *Decision Analysis* **1**(4) 205-216.

Kahneman, D., R. Thaler. 1991. Economic Analysis and the Psychology of Utility. *The American Economic Review* **81**(2) 341-346.

Kahneman, D., A. Tversky. 1979. An analysis of decision under risk. *Econometrica* **47**(2) 263-292.

Keeney, R. 1974. Multiplicative utility functions. *Operations Research* **22** 22-34.

Keeney, R. L., H. Raiffa. 1976/1993. *Decisions with Multiple Objectives: Preferences and Value Trade-offs*. John Wiley and Sons, New York, and Cambridge University Press, New York.

Keeney, R. L., D. A. Vernik. 2007. Analysis of the Biological Clock Decision. *Decision Analysis* **4**(3) 114-135.

Littman, M. L., D. H. Ackley. 1991. Adaptation in constant utility non-stationary environments. In Belew, R. K., L. B. Booker (Eds.). *Proceedings of the Fourth International Conference on Genetic Algorithms* 136-142. Morgan Kaufmann, San Mateo, CA.

Miyamoto, J. M. 1999. Quality-adjusted life years (QALY) utility models under expected utility and rank dependent utility assumptions. *Journal of Mathematical Psychology* **43** 201-237.

Munro, A., R. Sugden. 2003. On the theory of reference-dependent preferences. *Journal of Economic Behavior and Organization* **50**(4) 407-428.

Murray, C. J. L. 1994. Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bulletin of the World Health Organization* **72** 495-509.

Pliskin, J. S., D. S. Shepard, M. C. Weinstein. 1980. Utility functions for life years and health status. *Operations Research* **28** 206-224.

Simon, J. 2009. Decision Making with Prostate Cancer: A Multiple-Objective Model with Uncertainty. Forthcoming in *Interfaces*.

Stalmeier, P. F. M., P. P. Wakker, T. G. G. Bezembinder. 1997. Preference reversals: Violations of unidimensional procedure invariance. *Journal of Experimental Psychology: Human Perception and Performance* **23**(4) 1196-1205.

Torrance, G. W., D. Feeny. 1989. Utilities and quality-adjusted life years. *International Journal of Technology Assessment in Health Care* **5** 559-575.

Wathieu, L. 2004. Consumer Habituation. *Management Science* **50**(5) 587-596.

Chapter 2

Decision Analysis Using Geographic Information Systems

2.1 Introduction

The use of geographic information systems (GIS) has become widespread over the past decade. Policy decisions made using GIS generally affect one or more environmental factors which vary geographically. That is, the outcome can be described in terms of one or more variables whose values are defined over a region. To judge the relative desirability of various outcomes, the decision maker must be able to (implicitly) express each map of attribute values as a single overall value that can be compared against other outcomes. This issue arises frequently, because the information is usually presented as one or more arrays of values defined over a region. In these models, it is extremely important to properly take the decision maker's preferences into account. Simply put, the main problem is this: looking at two possible maps, how can a decision maker properly consider the information presented by each one and conclude that one or the other would be more desirable?

It is likely that, in reality, decision makers employ certain heuristics to make these judgments possible. One possible example of such a heuristic is a lexicographic semiordering[1]. With this heuristic, the decision maker has in mind several objectives for the chosen map. He mentally ranks these objectives in order of importance. If he can discern a significant difference between the two possible maps regarding the first objective, he will choose the superior map. If he cannot, then he looks for a significant difference regarding the second objective. If he can find one, he will choose the superior map, and if not, he will move on to the third objective. This process is repeated until either a difference is found on an objective or the two maps are deemed to be equally desirable. For a more detailed treatment, see Tversky (1969) or Luce (1978).

Of course, there are many other types of heuristics and cognitive shortcuts that one could use to compare maps. It is almost certain that decision makers in these situations are using such shortcuts, because a map of GIS data presents an enormous quantity of information, which is likely beyond what can be incorporated into a simple mental model of the situation.

In this work, we analyze these types of policy questions using multi-attribute decision theory. The first step is a discussion of the basic quantitative methods for assessing the value of a geographically-defined outcome. It is also important to analyze the conditions required for their validity. We discuss several properties of spatial preference structures which will improve the ease of implementation and the real-world feasibility of our methods.

[1] We call this a heuristic, since it is unlikely that a decision maker would consider it optimal to purposefully follow a lexicographic semi-order process.

After the theoretical development has been completed, then the analysis will be applied to several real-world problems. We examine a decision related to urban development, an irrigation allocation decision, a nuclear accident planning decision, and a fire coverage decision. We will also discuss the applicability of spatial preference models, along with the significant potential benefits that can be gained from this approach.

2.2 Literature Review

The previous work relevant to this paper can be divided into two main categories. The first category deals with the fundamental concepts in multi-attribute preference theory, and the second involves specific applications of GIS. In addition, some work has been done in an effort to apply certain decision analysis concepts to spatial models.

The development of multi-attribute value/utility theory can be attributed largely to a few influential works from the 1970's. Keeney and Raiffa (1976) provide utility independence conditions for multiple-attribute utility (for decisions under risk). Dyer and Sarin (1979) discuss several conditions under certainty, most notably mutual preferential independence, which permit the use of additive multi-attribute measurable value functions (for decisions under certainty). Building upon this fundamental work, Kirkwood (1997) elaborates on the use of spreadsheets for modeling multi-attribute decision models.

GIS have become a standard tool for formulating and analyzing spatial problems. They underwent a sharp increase in popularity during the 1990's, and are now the center of a large stream of literature. There are many papers examining various applications of GIS. For example, Knox and Weatherfield (1999) discuss their use in the context of irrigation and water resource management, Pendleton et al. (1999) analyze the use of GIS in studying wildlife habitat selection in Alaska, and Kohlin and Parks (2001) look at a GIS model to analyze deforestation.

There are also many papers which provide helpful explanations and overviews of the GIS methods in general. A good overview is provided by Bond and Devine (1991). They illustrate GIS as incorporating techniques from classical statistics very effectively. Arbia (1993) also provides a good overview of GIS with some more detail. Sampling and modeling error are discussed, but not as uncertainty over outcomes in the way that they might be in the literature on decision making. There are no notions of preferences or utilities present, as they are not closely related to the goals of the paper. This is a recurring theme in many GIS papers. The goal of these authors is mainly to expound the fundamentals of the systems rather than analyze the ensuing decision process. A decision analysis approach could increase both the power and the applicability of GIS tools.

However, there is a small subset of the literature discussing the use of GIS in making decisions. Spatial decision support systems and GIS are discussed by de Silva and Eglese (2000) and Worrall and Bond (1997). Both of these papers are valuable in transitioning between GIS and decisions. Malczewski (1999) has a more detailed analysis of decision making using GIS from a multi-criteria decision making perspective. Beinat and Nijkamp (1998) use a multi-criteria approach to land-use decisions. Chan (2005) incorporates decision analysis techniques into GIS problems elegantly. However, these problems do not involve spatial aggregation of preferences. They are mostly "siting" types of problems, in which a location is chosen from a map, resulting in a single array of attribute values. Even in the GIS literature that explicitly deals with decision making, information is defined spatially, but preferences rarely are. One notable exception is a park planning problem discussed by Keisler and Sundell (1997), who do focus on integrating multi-attribute utility and spatial outcomes.

2.3 Models

In decision analysis terminology, a person's preferences over the outcome of a variable are characterized with a *value function*. For example, an individual places a value on every possible monetary outcome. In situations with uncertainty, these value functions are replaced by *utility functions*, which can be evaluated not only for specific outcomes, but also for gambles over possible outcomes.

In most real situations, an outcome is defined by more than one variable. For example, a person's preferences over the weather will consider both temperature and precipitation (and probably other factors as well). In this case, the person would need a *multi-attribute value function* which maps both of these outcomes to a single numeric value. If future weather, which is uncertain, were also a consideration, then a *multi-attribute utility function* (Keeney and Raiffa 1976) would be required, which would allow the decision maker to compute an expected utility over the possible outcomes.

In general, the simplest type of multi-attribute value/utility function is an additive function. An additive multi-attribute value function is one that can be written as

$$V(z) = \sum_i a_i v_i(z_i), \qquad (2.1)$$

where z_i is the outcome of attribute i, v_i is the single attribute value function, and a_i is the weight that the decision maker places on attribute i. That is, the overall value is a weighted sum of the values obtained from each individual attribute. These functions have a notion of separability; a change in one attribute does not affect the value contributed by the others. Notice that no restriction is placed on the v_i; a value function on an individual attribute can take on any form, as long as it is defined over all possible values of z_i.

The use of an additive multi-attribute value function assumes mutual *preferential independence*. The simplest test for preferential independence is to consider two alternatives, and assume that all but two of the attributes are identical between these two alternatives. If comparison of the two alternatives (using the values on the two differing attributes) can be done without knowing the values of the identical attributes, then preferential independence is satisfied, and it is valid to assume an additive multi-attribute value function. Essentially, this condition means that the tradeoff between any two attributes does not depend on the values of the others.

For decisions with geographically-defined outcomes, it is desirable to use additive multi-attribute value functions if at all possible. The separability of attributes provides an enormous practical benefit; it allows for assessment of a value function on an attribute in one region without having to consider any other regions. This requires a variation on preferential independence which we call *spatial preferential independence*. If we consider a single attribute z, regions $R_1, R_2, \ldots, R_m$ exhibit spatial preferential independence with respect to z if the rank-ordering of alternatives with common values of z for some subset of $R_1, R_2, \ldots, R_m$ does not depend on those common values. That is, the decision maker is able to simply ignore this subset of regions, and look only at the regions over which z differs between alternatives. The resulting spatial additive value function has the form:

$$V(R) = \sum_{j=1}^{m} a_j v_j(z_j). \tag{2.2}$$

Another desirable condition on the decision maker's preferences is what we call *spatial homogeneity*. If $R_1, R_2, \ldots, R_m$ are spatially preferentially independent with respect to z, they are spatially homogeneous if the midvalue z_j^* of any interval $[z_j', z_j'']$ does not depend on j. That is, spatial homogeneity means that the value function for an attribute has the same shape for every region. A spatially additive and homogeneous function has the form:

$$V(R) = \sum_{j=1}^{m} a_j v(z_j). \tag{2.3}$$

Using this preference structure avoids the necessity of having to elicit a value function for each individual region. Notice that there is no subscript for location (or anything else) on the value function v. Spatial homogeneity makes the assessment of the decision maker's preferences much more tractable.

It is also possible to extend these conditions to continuous value functions. If mathematical expressions for the attribute values z and the weights a are defined continuously over the region, then the value function can be written instead as:

$$V(R) = \iint_R a(x, y) v[z(x, y)]\, dxdy, \tag{2.4}$$

where x and y are simply coordinates within region R. Note that spatial homogeneity is extremely important when using continuous value functions. It would be very difficult to assess a mathematical expression for a value function whose shape varies throughout the region.

These conditions can also be extended to decisions with uncertainty. Instead of preferential independence, however, the use of an additive multi-attribute utility function requires a slightly stronger condition called *additive independence*. Additive independence essentially requires that the decision maker could compute a separate expected utility component for each attribute, and then take the sum to obtain the overall expected utility. In other words, if attributes x and y can take values $\{x_h, x_l\}$ and $\{y_h, y_l\}$ respectively, then the decision maker must be indifferent between the following two gambles:

Gamble 1: (x_h, y_h) with probability .5, or (x_l, y_l) with probability .5

Gamble 2: (x_h, y_l) with probability .5, or (x_l, y_h) with probability .5

We say that the partition $R_1, R_2, \ldots, R_m$ displays *spatial additive independence* with respect to z if the rank-ordering for any set of alternatives depends only on the marginal probability distributions over the levels $z_1, z_2, \ldots, z_m$ of z in each of the subregions $R_1, R_2, \ldots, R_m$ for each alternative. This is analogous to the additive independence condition discussed above. In spatial decision problems with uncertainty, instead of maximizing a spatial value function, we maximize expected utility using a spatial utility function. If spatial additive independence is satisfied, this expected utility is given by:

$$EU = \sum_{k=1}^{K} p_k \left(\sum_{j=1}^{m} a_j u_j \left(z_{jk} \right) \right), \tag{2.5}$$

where p_k represents the probability of outcome k, $k = 1,\ldots,K$. If spatial homogeneity is satisfied as well, we can rewrite (2.5) as:

$$EU = \sum_{k=1}^{K} p_k \left(\sum_{j=1}^{m} a_j u\left(z_{jk}\right) \right). \tag{2.6}$$

As in the case under certainty, we can extend (2.6) to be defined continuously, resulting in expected utility given by:

$$EU = \sum_{k=1}^{K} p_k \left(\iint_R a\left(x,y\right) u\left(z_k\left(x,y\right)\right) \right). \tag{2.7}$$

Once again, it is also very important that spatial homogeneity holds in the continuous case. If it does not hold, then assessing the utility function over the region will likely be an intractable task.

Thus far, we have considered only cases with a single variable defined across a region. In many cases, the decision maker will consider multiple variables, one or more of which may vary geographically. Incorporating this is actually a very straightforward extension. We already have the structures in place to compute the scalar value achieved for each individual variable. We can either assume (for simplicity) that the overall multi-attribute value or utility function is additive in these scalar values, or assess a more complex form if tractable. We can also extend the model to allow weights to vary over time, or to incorporate notions of fairness or other interactions across subregions.

This set of functions and conditions on preference structures will serve as the foundation for the next section of this paper, in which we apply these spatial decision tools to several practical applications.

2.4 Applications

We now turn to four examples which demonstrate the practicality of these concepts as well as some of the insights that we can gain from them. The first example is based on data from the Decision Center for a Desert City, at Arizona State University. The other three are stylized examples based on data and models used in other analyses.

2.4.1 Urban Development

This example is based on data from a heat flux model developed by the Decision Center for a Desert City. The model is known as the "Local-Scale Urban Meteorological Parameterization Scheme" (LUMPS). Urban development has led to increased temperatures in Phoenix, mostly by reducing the amount of night cooling that occurs. As a result, there is a strong motivation to increase the quantity of vegetation, as "green" areas retain far less heat. However, this would also require more water, as these areas lose far more to evaporation than developed urban areas. Night cooling and evaporation rate are both important considerations when choosing development strategies.

Using the LUMPS model, evaporation rate and night cooling estimates were obtained for ten different tracts of land using each of three potential development strategies. The three strategies are: “compact” (denser urban development), “oasis” (more non-native vegetation), and “desert” (more native vegetation). The current levels of evaporation rate and night cooling are shown in Figure 2.1. The ten tracts for which data are available are not contiguous; we do not have information on evaporation rate or night cooling in the blank areas.

Base-Case LUMPS Data for Ten Tracts

Figure 2.1. Current evaporation rate and night cooling for each of the ten tracts.

Figure 2.2 shows the changes that would result from each strategy. The different shades of the tracts represent the current classifications: the darkest shade represents industrial tracts, the lightest shade xeric (desert) tracts, and the middle shade mesic (non-native vegetation) tracts.

LUMPS Data for Changes Resulting from Three Development Strategies

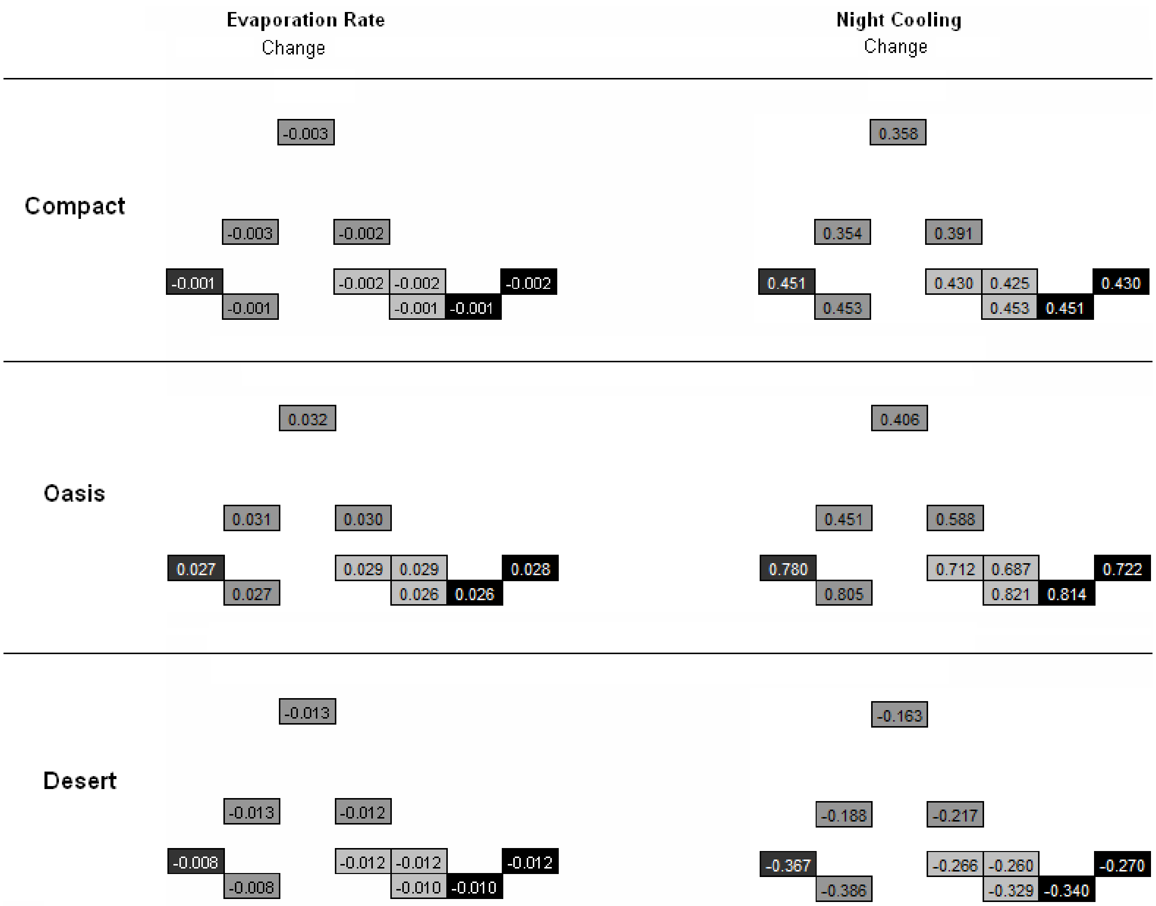

Figure 2.2. Changes in evaporation rate and night cooling that would result from implementing each of the three strategies in the ten tracts.

Given only the LUMPS data shown in Figures 2.1 and 2.2, it is not at all clear what should be done. To choose the optimal development strategy for each tract, we need to know the value functions for evaporation rate and night cooling, as well as the relative importance of each. We can then formulate the overall value using equation (2.3). Once these preferences are incorporated into the model, we can determine which strategy would result in the highest value for each tract. For example, if the decision maker has the single-attribute value functions for evaporation rate and night cooling shown in

Figure 2.3, and places a 40% weight on evaporation rate, then the resulting optimal development plan is the one shown in Figure 2.4, as shown below in detail. It is important to note that when determining the weight to place on each attribute, the decision maker must be aware of the possible ranges of each one. That is, the weights should reflect the relative benefit of moving from the worst achievable level of each attribute to the best achievable level, and the assessment questions asked of the decision maker would emphasize this very clearly.

Possible Value Functions on the Two Attributes

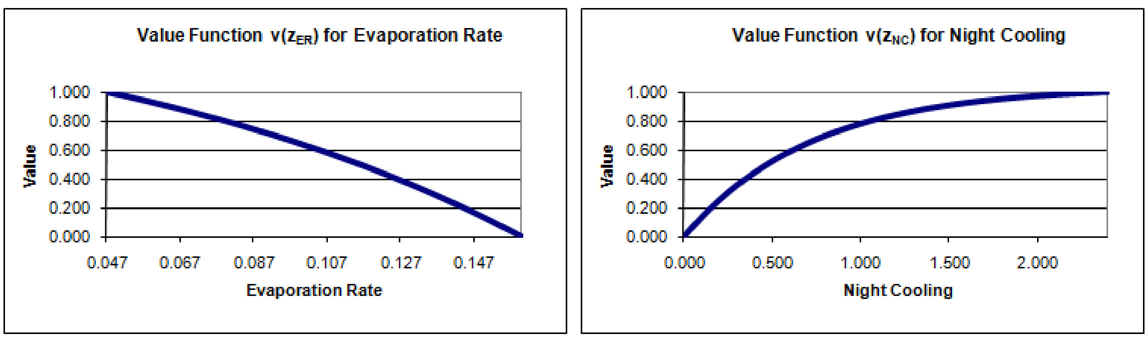

Figure 2.3. Example value functions for night cooling and evaporation rate.

Optimal Development Strategies

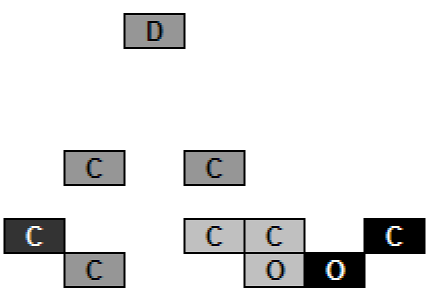

Figure 2.4. The optimal development plan with a particular set of decision maker preferences. C=Compact, O=Oasis, D=Desert

Once we frame this scenario using multi-attribute preferences, the decision problem becomes much more straightforward. For example, in the uppermost tract, we can compute the value achieved by each of the three development types, as shown in Table 2.1. Based on the overall values in the rightmost column of Table 2.1, we would recommend choosing the "desert" strategy for this tract. A similar calculation is done for each tract, resulting in the optimal development plan in Figure 2.4.

Values Achieved in One Tract by Three Development Plans

Dev. Type	Evap. Rate z_{ER}	Value $v(z_{ER})$	Night Cooling z_{NC}	Value $v(z_{NC})$	Overall Value $.4v(z_{ER}) + .6v(z_{NC})$
Compact	0.125 (0.128-0.003)	0.409	2.357 (1.999+0.358)	0.998	0.762
Oasis	0.160 (0.128+0.032)	0.000	2.405 (1.999+0.406)	1.000	0.600
Desert	0.115 (0.128-0.013)	0.508	1.836 (1.999-0.163)	0.957	Maximum = 0.777

Table 2.1. The value achieved from night cooling and evaporation rates given the three possible development plans in the uppermost tract.

We could even extend this model to incorporate other factors. For example, it would be possible to include restrictions on the combination of strategies implemented. There might be a minimum number of “compact” tracts required to meet industry needs, or a minimum number of “oasis” tracts to appease residents upset with increasing energy costs.

2.4.2 Irrigation Allocation

The second example deals with determining an optimal irrigation allocation strategy for the Anglian region in the United Kingdom. The data used are based on information presented by Knox and Weatherfield (1999). The question is: given the amount of irrigation requested by each subregion, the average amount of money generated per unit of irrigation in each region, and a total amount available, how can we best allocate the available resources? The demand and average amount of money per unit are shown in Figure 2.5.

Demand and Average Values for Irrigation

Demand (in cubic meters / hectare)

30	30	20			
40	40	30			
45	45	40			
40	45	55			
45	70	110	75	90	120
35	120	170	140	80	105
30	35	165	110	70	80
20	25	75	125	75	105
10	15	20	40	70	110
10	10	115	40	80	135

Avg Value of 1 cu. meter / hectare
(hundreds of pounds per hectare)

30	20	20			
25	20	20			
20	25	25			
15	25	30			
10	20	30	23	16	22
5	17	35	26	13	15
2	5	26	30	12	11
2	3	20	18	10	15
1	3	20	8	14	17
1	2	18	12	20	22

Figure 2.5. Demand and average values for irrigation in each subregion. In the left map, darker shading represents higher demand, and in the right map, darker shading represents higher average value.

The obvious next step is to run a linear optimization over the possible allocations to determine which one would generate the most money. In this optimization model, the decision variables are the amounts of irrigation provided to each subregion, the objective function coefficients are the average monetary values per unit of irrigation in the subregions, and the constraints are the total demands (as upper bounds) and the overall amount of irrigation available. The resulting optimal policy is shown in Figure 2.6.

This result is very lopsided; it suggests that all of the demand should be met in the most profitable subregions, while the less profitable ones should be completely ignored. This is most likely an undesirable strategy, both practically and politically. The primary concern here is that this “optimization” led to a result that was clearly not optimal. This happened because we did not adequately incorporate the preferences of the decision maker. We assumed that the profit each region obtained from a unit of irrigation was constant (which is extremely unlikely), and that the implied value achieved consisted purely of the profit. Rather than assuming such a linear value function, actually assessing the value function for a region's allocation will yield more detailed preferences for irrigation in a region as shown in Figure 2.7. This particular value function represents a modestly decreasing marginal value of irrigation within a subregion.

Initial Optimal Policy

Optimal Policy Using Average Values

(cubic meters per hectare)

30	0	0			
40	0	0			
0	45	40			
0	45	55			
0	0	110	50	0	0
0	0	170	140	0	0
0	0	165	110	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

Figure 2.6. The optimal irrigation policy with an implicit linear value function. Only 12 areas (shaded) receive non-zero allocations.

Value Function for Irrigation Level

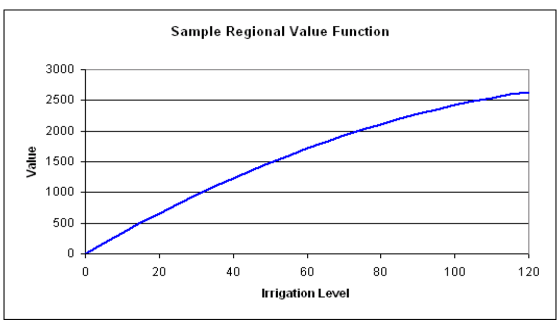

Figure 2.7. A nonlinear value function on the irrigation amount in one subregion.

If we assume spatial homogeneity holds, then we may apply this value function to every subregion. If we optimize using the resulting preferences (with (2.3) as the objective function, all $a_j = 1$) instead of the initial implicit linear value functions, we obtain the policy shown in Figure 2.8. This optimal policy is much more intuitively appealing; it still provides more irrigation to the more profitable regions, but it accounts for the fact that we also prefer to give modest irrigation to many areas rather than meet the full demand of a few areas. The changes in allocations from the initial model to the new model are shown in Figure 2.9. If the new allocations were plugged into the initial optimization model (i.e. without including a nonlinear value function), the profit would be approximately 90% of the profit achieved by the policy shown in Figure 2.6.

Optimal Policy with Nonlinear Values

Optimal Policy Using a Nonlinear Value Function

(cubic meters per hectare)

20.2	10.3	6.92			
21.7	13.8	10.3			
15.5	24.4	21.7			
0.64	24.4	37.1			
0	24.2	74.2	35.5	8.84	52.2
0	20.5	131	80.2	0	1.67
0	0	94.6	74.2	0	0
0	0	25.9	29.4	0	1.67
0	0	6.91	0	0	18.8
0	0	27.1	0	27.6	58.7

Figure 2.8. The optimal irrigation policy resulting from a nonlinear value function, with 29 areas receiving non-zero allocations.

Resulting Change in Allocations

Changes in Optimal Allocations

(cubic meters per hectare)

-9.8	10.3	6.92			
-18	13.8	10.3			
15.5	-20.6	-18.3			
0.64	-20.6	-17.9			
0	24.2	-35.8	-14	8.84	52.2
0	20.5	-39	-60	0	1.67
0	0	-70.4	-36	0	0
0	0	25.9	29.4	0	1.67
0	0	6.91	0	0	18.8
0	0	27.1	0	27.6	58.7

Figure 2.9. The change in optimal allocations from the original model to the model incorporating a nonlinear value function.

2.4.3 Nuclear Accident Countermeasures

The third example is based on the scenario discussed by French (1995) regarding countermeasures following a nuclear accident. We assume that a disaster has occurred at a nuclear plant, and try to determine how best to protect the people in the surrounding region. There are three possible countermeasures that can be employed in each subregion. First, officials may encourage “sheltering,” which means ensuring that people are shielded from radiation as much as possible. Second, they may provide iodine supplements, which offer some degree of protection from negative effects of radiation. Third, they may encourage or even force residents to evacuate.

In this particular example, the wind is blowing from the northwest, and thus the southeastern part of the region is at the greatest risk. In addition, the radioactive plume does not touch down immediately, meaning that areas farther southwest are actually at more risk than the areas immediately adjacent to the plant. This is a fact often not understood or accepted by the general public.

Figure 2.10 shows the increase in value (as would be determined by the decision maker) achieved by implementing each of the countermeasures in each subregion of the map. There is a budgetary constraint; they cannot provide the maximum possible protection to everyone. Each countermeasure has an associated per-subregion cost. In addition, there is a logical constraint; they cannot both shelter and evacuate a subregion. The shading in Figure 2.10 is simply to illustrate physical proximity; the darker subregions are closer to the plant.

Additional Value Provided by Each of the Three Countermeasures

Region Benefit

Sheltering

0	0	0.1	0.1	0.1	0.1	0.1	0.1	0.2
0	0	0.1	0.1	0.1	0.1	0.1	0.2	0.2
0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2
0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2
0.1	0.1	0.1	0.1		0.2	0.3	0.3	0.3
0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.3
0.1	0.1	0.2	0.2	0.3	0.3	0.3	0.3	0.3
0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.3	0.3
0.2	0.2	0.2	0.3	0.3	0.3	0.3	0.3	0.3

Iodine

0	0	0	0.1	0.1	0.1	0.1	0.1	0.1
0	0	0	0.1	0.1	0.1	0.1	0.1	0.1
0	0	0	0.1	0.1	0.1	0.1	0.1	0.1
0	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2
0.1	0.1	0.1	0.1		0.2	0.2	0.2	0.2
0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.2
0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.2	0.2
0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.2
0.1	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.2

Evacuation

0	0	0.1	0.1	0.1	0.1	0.1	0.2	0.2
0	0	0.1	0.1	0.1	0.1	0.1	0.2	0.2
0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.2
0.1	0.1	0.1	0.2	0.2	0.3	0.3	0.4	0.4
0.1	0.1	0.2	0.2		0.3	0.5	0.5	0.5
0.1	0.1	0.2	0.3	0.3	0.5	0.7	0.6	0.5
0.1	0.1	0.2	0.3	0.5	0.7	0.7	0.7	0.6
0.2	0.2	0.3	0.4	0.5	0.6	0.7	0.6	0.5
0.2	0.2	0.2	0.4	0.5	0.5	0.6	0.5	0.5

Figure 2.10. Value provided by the three countermeasures in each subregion.

The general form of the optimization problem is given by (2.8). The binary decision variables represent whether or not a strategy is implemented in each individual subregion, and the data for the value functions are shown in Figure 2.10. The first constraint is the budgetary constraint, and the second precludes both sheltering and evacuating the same area:

$$
\begin{aligned}
\underset{S_j, I_j, E_j (j=1,\dots,n)}{Max} \quad & \sum_{j=1}^{n} \left(S_j v_j(S) + I_j v_j(I) + E_j v_j(E) \right) \\
s.t. \quad & \sum_{j=1}^{n} \left(S_j c_s + I_j c_s + E_j c_s \right) \le B \\
& S_j + E_j \le 1 \qquad (j = 1,\dots,n) \\
& S_j, I_j, E_j = 0 \text{ or } 1 \; (j = 1,\dots,n)
\end{aligned}
\tag{2.8}
$$

Consider the two potential strategies shown in Figure 2.11. The highlighted cells represent the implementation of the corresponding countermeasure in those subregions. Strategy 1 (the "risk-based strategy") is an example of what an expert might recommend;

the subregions with the greatest risk are being evacuated, the subregions with moderate risk are being sheltered, and iodine is provided for anyone who is likely to get at least some minimal level of exposure. Strategy 2 (the "public's perceived risk-based strategy") is an example of what a politician might recommend, as it also considers public reaction. It is not drastically different from the expert recommendation, but it focuses also on protecting the subregions closest to the plant.   Based on the model in (2.8), strategy 1 would be preferred, with an objective function value of 14.29. Strategy 2 would have an objective function value of 13.35. In both strategies, the entire budget is spent. A practical course of action might be to solve for an optimal solution to this model, and then to adjust it according to politics or public perception.

Two Potential Strategies For Distributing the Countermeasures

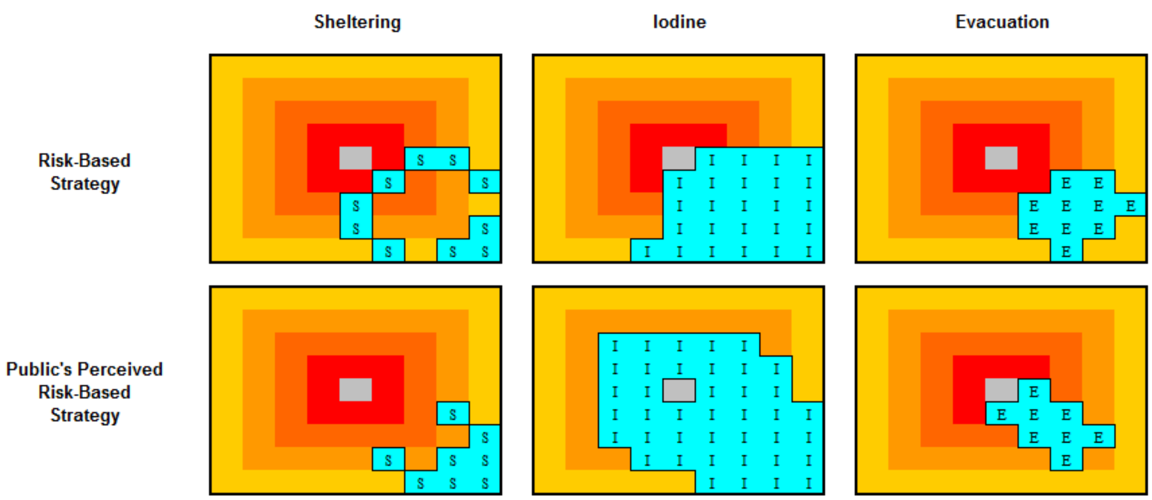

Figure 2.11. Two possible strategies being considered for distributing the three countermeasures. (Highlighted cells indicate that this action is used in this subregion.)

An interesting extension, however, would be to include the public reaction to a strategy as an attribute[2]. In the model shown in (2.8), Strategy 1 will likely be deemed a better choice than Strategy 2. However, Strategy 2 will likely be met with a more favorable reaction, since it provides more protection to the subregions closer to the plant.

2.4.4 Location of Fire Stations to Provide Fire Coverage

The fourth example is a stylized fire coverage problem. It is motivated by the location problem discussion in Church and Roberts (1983). In our example, the decision is where to locate three fire stations within a city to minimize response times. The basic optimization model is shown in (2.9):

$$\begin{aligned} \min_{X_k(k=1,\ldots,n)} \quad & \frac{1}{n}\sum_{i=1}^{n}\max\left(m, 1-\sum_{j=1}^{n}A(i,j)\,X_k\right) \\ s.t. \quad & \sum_{i=1}^{n}X_i \leq 3 \\ & X_i = 0 \text{ or } 1 \quad (i=1,\ldots,n) \end{aligned} \tag{2.9}$$

The objective is to minimize the average response time over all of the subregions of the city. The lengths of time are scaled such that the maximum possible average response time for a subregion is 1. In this stylized model, X_k is a binary variable equal to 1 if a fire station is located in subregion k and 0 otherwise, m represents the minimum achievable

[2] See also Feng and Keller (2006) for a multi-attribute decision analysis approach to choosing among options for distributing radioactive iodine, where public reactions are considered.

average response time for a region, and $A(i,j)$ is a measure of accessibility of subregion i from subregion j.

Solving (2.9) results in the optimal fire station locations shown in the left side of Figure 2.12. Surprisingly, two stations are located in adjacent subregions.

Optimal Fire Station Locations Using the Initial Optimization Model

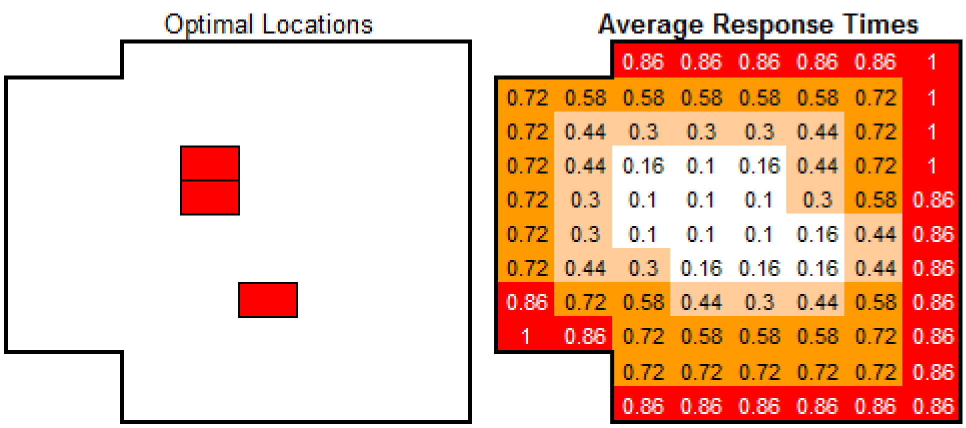

Figure 2.12. Optimal locations (shaded cells on left) and average response times (on right) when minimizing overall average response time.

However, this assumes that the value function of average response times within a subregion is linear. It is likely that there are diminishing returns on decreases in response time from the perspective of a policy maker. If we allow for a nonlinear value function that incorporates this, as illustrated in Figure 2.13, then the optimal locations will be the ones shown in Figure 2.14.

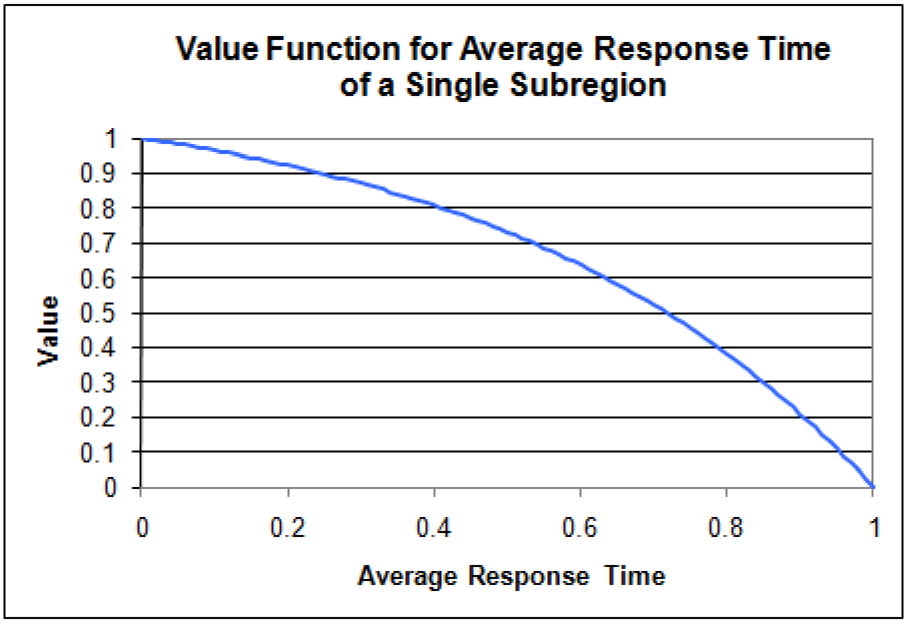

Figure 2.13. A nonlinear value function showing diminishing returns on decreasing average response times for a single subregion.

Optimal Fire Station Locations with a Nonlinear Value Function

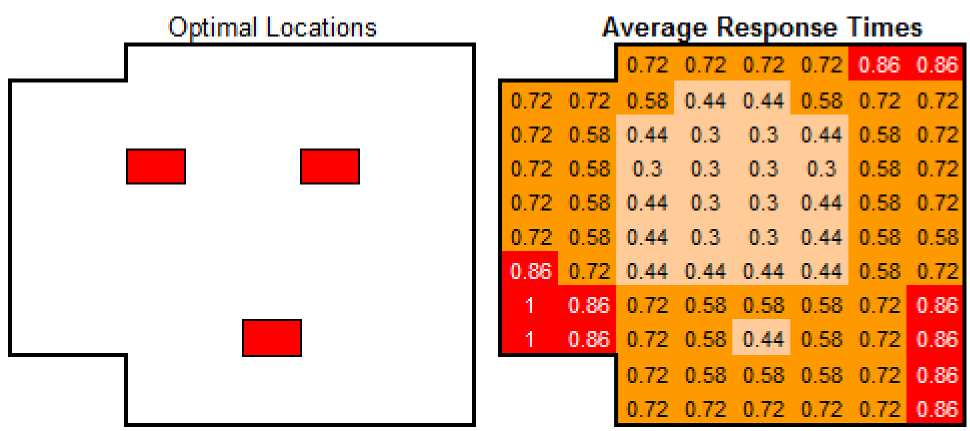

Figure 2.14. Optimal locations (shaded cells on left) and average response times (on right) when using a nonlinear value function.

Notice that, just as it did in the irrigation example, assessing the value function leads to a more equitable distribution of resources. The three fire stations are more spread out, and the average response times vary less across the city than they do with the initial optimization. When making these types of decisions, a policy maker may find that an optimization which does not explicitly incorporate a value function will lead to results which are very clearly non-optimal. Ideally, an interface could be set up to allow the policy maker to impose additional constraints (such as a maximum acceptable average response time in particular subregions), and to graphically see the effects of changing station locations on the average response times.

2.5 Conclusion

We have introduced a framework for analyzing decisions made using GIS. These types of decisions are becoming increasingly important for policy makers, and it is therefore vital to develop theoretical foundations for studying them. When faced with a decision that has spatially-defined outcomes, formulating specific structures and conditions on the decision maker's preferences will allow an analyst to elicit an appropriate value function and correctly gauge the desirability of any given outcome. To demonstrate the applicability of these concepts, we used them to solve four particular examples. They were helpful in analyzing urban development, irrigation, nuclear disaster planning, and fire coverage. It is our hope that these spatial decision tools will be applied to a wide range of real-world policy decisions.

2.6 References

Arbia, G. 1993. The Use of GIS in Spatial Statistical Surveys. *International Statistical Review* Aug 339-359.

Beinat, E., P. Nijkamp. 1998. *Multicriteria analysis for land-use management.* Kluwer Academic Publishers. Boston, MA.

Bond, D., P. Devine. 1991. The Role of Geographic Information Systems in Survey Analysis. *The Statistician* **40**(2) 209-216.

Chan, Y. 2005. *Location, transport and land-use : modelling spatial-temporal information*. Springer. Berlin; New York.

Church, R. L., K. L. Roberts. 1983. Generalized Coverage Models and Public Facility Location. *Papers in Regional Science* **53**(1) 117-135.

de Silva, F. N., R. W. Eglese. 2000. Integrating Simulation Modelling and GIS: Spatial Decision Support Systems for Evacuation Planning. *The Journal of the Operational Research Society* Apr 423-430.

Dyer, J. S., R. K. Sarin. 1979. Measurable multiattribute value functions. *Operations Research* **27**(4) 810-822.

Feng, T., L. R. Keller. 2006. A Multiple-Objective Decision Analysis for Terrorism Protection: Potassium Iodide Distribution in Nuclear Incidents. *Decision Analysis* **3**(2) 76-93.

French, S. 1996. Multi-attribute decision support in the event of a nuclear accident. *Journal of Multi-Criteria Decision Analysis* **5**(1) 39-57.

Hirschfield, A., P. J. B. Brown, P. Bundred. 1995. The Spatial Analysis of Community Health Services on Wirral Using Geographic Information Systems. *The Journal of the Operational Research Society* Feb 147-159.

Keeney, R. L., H. Raiffa. 1976/1993. *Decisions with Multiple Objectives: Preferences and Value Trade-offs*. John Wiley and Sons, New York, and Cambridge University Press, New York.

Keisler, J. M., R. C. Sundell. 1997. Combining Multi-Attribute Utility and Geographic Information for Boundary Decisions: An Application to Park Planning. *Journal of Geographic Information and Decision Analysis* **1**(2) 100-119.

Kirkwood, C. W. 1997. Strategic Decision Making Multiobjective Decision Analysis with Spreadsheets. Wadsworth Publishing, Belmont, CA.

Knox, J. W., E. K. Weatherfield. 1999. The Application of GIS to Irrigation Water Resource Management in England and Wales. *The Geographical Journal* Mar 90-98.

Knox, J. W., E. K. Weatherhead, R. I. Bradley. 1996. Mapping the Spatial Distribution of Volumetric Irrigation Water Requirements for Maincrop Potatoes in England and Wales. *Agricultural Water Management* Jun 1-15.

Kohlin, G., P. J. Parks. 2001. Spatial Variability and Disincentives to Harvest: Deforestation and Fuelwood Collection in South Asia. *Land Economics* May 206-218.

Luce, R. D. 1978. "Lexicographic Tradeoff Structures." *Theory and Decision* **9**(2) 187-193.

Malczewski, J. 1999. *GIS and Multicriteria Decision Analysis*. Wiley. Hoboken, NJ.

Opsomer, J. D., S. M. Nusser. 1999. Sample Designs for Watershed Assessment. *Journal of Agricultural, Biological, and Environmental Statistics* Dec 429-442.

Pendleton, G. W., K. Titus, E. DeGayner, C. J. Flatten, R. E. Lowell. 1998. Compositional Analysis and GIS for Study of Habitat Selection by Goshawks in Southeast Alaska. *Journal of Agricultural, Biological, and Environmental Statistics* Sep 280-295.

Recker, W., Y. Chung, T. Golob. 2005. A Tool for the Incorporation of Non-Recurrent Congestion Costs of Freeway Accidents in Performance Management. *California PATH Research Report*. Nov.

Tversky, A. 1969. "The Intransitivity of Preferences." *Psychological Review* **76** 31-48.

Worrall, L., D. Bond. 1997. Geographical Information Systems, Spatial Analysis and Public Policy: The British Experience. *International Statistical Review* Dec 365-379.

Chapter 3

Modeling Altruistic Preferences

3.1 Introduction

The economic definition of altruism is the concept of a person making a decision to increase the well-being of others despite incurring some individual cost. The idea of altruism challenges the traditional definition of utility over wealth or over a bundle of goods as a descriptive model, since an altruistic individual will blatantly violate the axioms of single-attribute utility maximization described by Savage (1954). Altruism, however, is not incompatible with utility theory; it simply requires a multiple-attribute utility function that can incorporate altruistic preferences. To this end, we will distinguish between a traditional single-attribute utility function, which we will write as *u*(), and a multi-attribute utility function as discussed by Keeney and Raiffa (1976), which we will write as *U*().

The concept of distinguishing the single-attribute *u*() and the multi-attribute *U*() in the context of altruism can be illustrated very well by considering an individual suffering from autism. The National Institute of Neurological Disorders and Stroke (NINDS) explains:

> "Autism is characterized by impaired social interaction, problems with verbal and nonverbal communication, and unusual, repetitive, or severely limited activities and interests."

Autism is a complex set of disorders with a vast array of symptoms, but one common observation, as NINDS points out, is that autistic individuals usually lack any sort of empathy. This is the trait that sparked our practical interest in the concept of altruistic utility.

This autistic person frequently decides to take actions that are perceived by others as mean-spirited. However, those who know him well realize that he does not make decisions with the intent to harm others; he simply does not take into account the effects of his decisions on other people. Society deems this to be a serious social disorder. So why are most economists and decision analysts not incorporating the effects of a decision on others into an individual's preferences?

In this paper, we develop altruistic utility models beginning with a simple two-person additive case, and then extend the results to more general structures, to more than two individuals, and to single and multiple group models. For each model, we analyze solvability conditions, and we provide several illustrative examples.

3.2 Literature Review

In general, the economic literature has focused on developing descriptive theories of altruism which explain phenomena that have been empirically observed. There is no widespread agreement on the exact definition of altruism (Monroe 1998), but we are concerned primarily with the economic interpretations. This concept of altruism has been discussed extensively in two main contexts. The first examines altruism as observed in charitable giving, and the second studies altruistic behavior from an evolutionary perspective. There are also many papers discussing the psychological aspects of altruism, which we will touch on briefly. Finally, we will discuss some previous work on multiple-individual outcomes that are relevant to the group utility models we will explore toward the end of this paper.

One of the most obvious examples of altruism in society is the donation of money, time, or other resources to charity. Titmuss (1971) discusses the rationale behind giving blood, and cannot explain it as "complete, disinterested, spontaneous altruism." Arrow (1975) discusses Titmuss' work in the context of utility theory, and suggests different types of preferences for which giving blood might be desirable. The one which is most relevant to this paper is the individual whose well-being depends "both on his own satisfaction and on the satisfactions obtained by others." Sugden (1982) argues that charitable giving cannot be explained simply by including improvement of the public good (charity) in the philanthropist's utility function. Andreoni and Scholz (1998) acknowledge that there may be some interdependencies, as an individual may compare his/her own charitable donations to a reference level[1]. However, they find that this is not a major factor in the underlying motivation for altruistic behavior. Andreoni (1990) shows that altruism in society can be explained using "impurely altruistic" utility functions, where individuals receive utility both from the improvement in the public good (e.g. charity) and from the act of giving itself.

In this paper, we frame altruistic preferences as characteristics of individual decision makers. However, given the elusiveness and controversy surrounding this topic, it is important to discuss the research streams that explain and justify the existence of altruism in rational decision makers in the first place. There is a large stream of literature dedicated to the development and spread of altruistic behavior in evolutionary settings.

Hamilton (1963) discusses an important point regarding the evolution of an altruistic gene. He explains that for the altruistic gene to spread, it must lead NOT to behavior which helps the individual survive, but to behavior that helps the gene survive, regardless of who contains it. So, for example, if we divide a population up into 100 tribes, and three of the tribes contain many individuals who are altruistic (toward their own tribe members), it may be that these three tribes are more likely to survive, and thus the altruistic gene can spread. Simon (1990) proposes that it is often in the best interest of

[1] See McCardle, Rajaram, and Tang (2009) for an example of utility functions incorporating tiered levels of charitable giving.

docile individuals to engage in altruistic behavior, since they are given a social "tax" for doing otherwise.

In recent years, more mathematical evolutionary behavior models have been developed, some of which provide insights into altruism. Trivers (1971) discusses reciprocal altruism: the idea that two individuals engaging in altruistic acts toward one another can both reach better outcomes than if neither had done so. Alexander (1987) and Nowak and Sigmund (2005) discuss the evolution of "indirect reciprocity." Indirect reciprocity is the concept that a person may engage in altruistic behavior not directed at the individual(s) being altruistic toward them (or expected to be altruistic toward them in the future). The expectation is that, in the long run, someone will reciprocate. Bergstrom and Stark (1993) show that even in the absence of reciprocity, there are simple evolutionary models that explain altruism. They develop several models using a standard one-shot prisoner's dilemma, and demonstrate that it is very possible for cooperation to become the norm in the population.

Another interesting concept is altruistic punishment: the idea that individuals may incur some direct cost to punish another individual who acts in an undesirable (usually selfish) fashion. Fehr and Gächter (2002) show that altruistic punishment in an evolutionary setting is likely to increase cooperation. Boyd et al. (2003) expand this concept to explain altruistic punishment and cooperation even in large populations where interactions are anonymous and are not repeated.

Any altruistic preference model depends, whether implicitly or explicitly, on some form of interpersonal comparison of utility. The issue of interpersonal utility comparison is discussed by Robbins (1935), and Sen (1979) provides an informative summary of many of the issues involved. Narens and Luce (1983) show that individuals may often achieve ordinal agreement of outcomes, but that this is dependent on social factors, and may not represent true intercomparability. This paper relies on interpersonal comparison of utility only within the individual preferences; that is, we do not attempt to determine optimal social outcomes.

Chen and Plott (2002) discuss the aggregation of individual beliefs into a single group belief. For example, if individuals are permitted to buy and sell shares of an election candidate, then the group outcome (market price) is a function of the individual values placed on the candidate. This idea is also analyzed by Forsythe, Rietz, and Ross (1999), and by Forsythe et al. (1995). These papers are similar to our group altruistic model, in that they examine an individual's interaction with the overall group rather than with other individuals.

3.3 Modeling Approach

Contrary to the approach taken by many researchers in the past, we would like to analyze altruistic preference models in a normative fashion. In particular, we are interested in preference structures over outcomes affecting multiple altruistic individuals.

Consider a decision made by person 1 for which the outcome is also experienced by person 2. If person 1 is not altruistic, then $U_1(x) = u_1(x)$ (where x is some physical outcome). Person 1 is neither hateful nor overly generous. He is, in a loose sense, suffering from autism. If person 1 is altruistic, then the question is: what should $U_1(x)$ look like? The simplest approach is to let $U_1(x) = f(u_1(x), u_2(x))$. This implies that person 1's utility function depends on the traditional utility experienced by both himself and by person 2. However, this does not fully capture what we mean by altruism. We would also like to ensure that a better outcome for person 2 will have a positive effect on person 1's utility. We can account for this simply by asserting that

$$\frac{\partial U_1}{\partial u_2(x)} > 0, \tag{3.1}$$

provided that U_1 is differentiable in u_2. This idea seems reasonable; at the very least, it is consistent with our basic notion of altruism. It is tempting to assume that (3.1) should hold when this partial is not equal to zero. However, there are many decision models and psychological concepts that claim this is not the case. For example, any happiness model that includes social comparison implies that (3.1) does not hold, since an individual would be made better off when others suffer (although these utility models generally are not framed that way).

A much more interesting scenario arises if person 2 is altruistic as well. In this case, we need to decide whether it would be more appropriate to model person 1's utility as

$$U_1(x) = f\left(u_1(x), u_2(x)\right) \tag{3.2a}$$

or as

$$U_1(x) = f\left(u_1(x), U_2(x)\right). \tag{3.2b}$$

This is really a philosophical question: does an altruistic person care inherently about the physical outcome experienced by others, or the overall well-being of others?

We claim that (3.2b) is a more accurate representation of altruistic utility. Altruistic individuals do tend to prefer that others achieve higher levels of traditional utility. However, preferences over others' traditional utility can also be viewed simply as a proxy for preferences over the well-being of others, without relying on such a rigid restriction on how that well-being is derived.

The difficulty in developing models based on the idea in (3.2b) is that they will involve complex interdependencies between the utilities of the individuals. The major concern is that it might not be possible to uniquely solve for each individual's utility. In this paper, we prove some general existence results which show that it is indeed possible if a few basic conditions are met. We will begin with a very simple specific scenario, and then expand to more general structures.

It is also important to note that the results obtained here extend beyond altruistic utility. In fact, the analysis in this paper can be applied to any situation where individual outcomes have externalities, whether positive or negative. Altruistic utility is a practical and concrete example, but these models will be applicable and informative with only minor modifications for any such setting.

3.4 Bell and Keeney's Framework

Bell and Keeney (2009) describe a scenario with two people and a set of physical outcomes $A = \{a_1, \ldots, a_n\}$ (we will refer to the outcomes as $\{x_1, \ldots, x_n\}$ for consistency with the rest of this paper). The specific example they use involves two people going out to eat a meal together. They have utility functions $u_1(x_i)$ and $u_2(x_i)$ over the physical outcomes, which in their scenario are the possible restaurants. They also have altruistic (additive) utility functions, which are written as

$$U_1(x_i) = \alpha_1 u_1(x_i) + (1-\alpha_1) u_2(x_i) \tag{3.3}$$

and

$$U_2(x_i) = \alpha_2 u_2(x_i) + (1-\alpha_2) u_1(x_i), \tag{3.4}$$

where $0 \leq \alpha_1, \alpha_2 \leq 1$. This allows each person to express preferences over how much the other likes the physical outcome. Notice, however, that person 1's preferences are unaffected by the altruistic nature of person 2. In person 1's altruism, she does not consider the idea that her own utility has a positive effect on person 2's utility. She displays what we will refer to in this paper as "specific altruism."

Bell and Keeney do mention the possibility of including $U_2(x_i)$ instead of $u_2(x_i)$ in person 1's utility function (and $U_1(x_i)$ in person 2's utility function). They point out, however, that this leads to "double counting" of utility. That is, person 1 is made happier not only by a physical outcome that she enjoys, but also by the positive effect that her resulting happiness has on person 2. They then conclude that this invalidates the model. We contend that not only is this *not* a problem, it is precisely what a model of true altruism should do!

3.5 Results on Two-Person General Altruistic Utility with Additive (Altruistic Utility) Functions

We view the "specific altruism" utility expression as simply a proxy for what we call "general altruistic utility" (GAU). GAU expresses the idea that altruism is nothing more or less than including a positive effect from another person's overall utility when forming one's own utility function. If the other person's overall utility includes an altruistic component, then so be it. Anything that makes the other person better off will make the person with GAU better off.

Let's consider, as Bell and Keeney do, two altruistic people going out to dinner. Both of them care about each other very much. Is Person 1's altruistic benefit a direct result of Person 2 enjoying her meal? Or is his altruistic benefit a direct result of Person 2's high level of well-being? Of course, it may be difficult to separate the two, since Person 2 is likely to be happy largely as a consequence of enjoying her meal. However, in terms of Person 1's preferences, we contend that Person 2's happiness is most likely what will directly improve his utility.

If we include $U_2(x_i)$ in Person 1's utility function and $U_1(x_i)$ in Person 2's utility function, the functions are of the form:

$$U_1(x_i) = \alpha_1 u_1(x_i) + (1-\alpha_1) U_2(x_i) \qquad (3.5)$$

and

$$U_2(x_i) = \alpha_2 u_2(x_i) + (1-\alpha_2) U_1(x_i). \tag{3.6}$$

This creates an interesting dynamic, as each person's altruistic utility depends on the altruistic utility of the other. The first natural question to ask here is: can we actually solve this? That is, can we express $U_1(x_i)$ and $U_2(x_i)$ uniquely in terms of $u_1(x_i)$ and $u_2(x_i)$?

It turns out that we can, provided that α_1 and α_2 are not both equal to 0. The result is the following pair of GAU functions:

$$U_1(x_i) = \frac{\alpha_1 u_1(x_i) + (1-\alpha_1)\alpha_2 u_2(x_i)}{1-(1-\alpha_1)(1-\alpha_2)} \tag{3.7}$$

$$U_2(x_i) = \frac{\alpha_2 u_2(x_i) + (1-\alpha_2)\alpha_1 u_1(x_i)}{1-(1-\alpha_1)(1-\alpha_2)}. \tag{3.8}$$

See Appendix for details. The reason that we cannot have $\alpha_1=\alpha_2=0$ is that it reduces (3.5) and (3.6) to $U_1(x_i)=U_2(x_i)$ and $U_2(x_i)=U_1(x_i)$, which gives us no information about the levels of utility actually achieved.

If Person 1 and Person 2 both display GAU, then both can be made better off as a result of each other's altruism. If Person 2 enjoys her meal, this makes Person 1 happier, which in turn makes Person 2 happier, which then makes Person 1 happier, ad infinitum. One common reaction to this realization is that perhaps this means that two people satisfying GAU should converge to identical utility functions. In fact, this is not the case.

Let's assume that Person 1 and Person 2 both have additive utility functions as described by (3.3) and (3.4), where $0<\alpha_1, \alpha_2<1$, and therefore they also have GAU functions as described by (3.7) and (3.8). We would like to know what is required to have $U_1(x_i) = U_2(x_i)$, that is, for Person 1 and Person 2 to have identical utility functions. Using (3.7) and (3.8), it turns out that in order to have $U_1(x_i) = U_2(x_i)$, it must be true that $u_1(x_i) = u_2(x_i)$ for all i (see Appendix for details). That is, the ONLY WAY that Person 1 and Person 2 will have equivalent GAU preferences over A is if their physical preferences over A are identical. This seems somewhat counterintuitive to the idea of altruism; it implies that even strongly altruistic individuals will sometimes prefer alternatives that decrease the utility of others. However, this is undeniably true for the pairwise additive model. If Person 1 and Person 2 factor their own physical preferences into their utilities at all, then they will never have identical utility functions. Proponents of successful marriages might do well to consider this idea.

3.6 Results on Two-Person General Altruistic Utility for More General Utility Functions

We have shown that functions of the form in (3.7) and (3.8) are achievable for two people with additive altruistic utility functions. The next question is: can we expand this result to pairwise GAU preferences without relying on a specific structure? That is, if we express these general utility functions as:

$$U_1(x_i) = f_1\left(u_1(x_i), U_2(x_i)\right) \tag{3.9}$$

and

$$U_2(x_i) = f_2\left(u_2(x_i), U_1(x_i)\right), \tag{3.10}$$

can we solve for $U_1(x_i)$ and $U_2(x_i)$ purely as functions of $u_1(x_i)$ and $u_2(x_i)$? It turns out that we can, although the required assumption will be slightly more involved. To do this, we must use the implicit function theorem. The implicit function theorem states that if we have a function f: $R^{n+m} \rightarrow R^m$ (with variables $x_1,\ldots,x_m, x_{m+1},\ldots,x_{m+n}$) then under certain conditions on the partial derivatives of f, it is possible to express $x_1,\ldots,x_m$ purely as a function of $x_{m+1},\ldots,x_{m+n}$. For more details on this theorem, see Krantz and Parks (2002).

In this case, we would like to express $U_1(x_i)$ and $U_2(x_i)$ as functions only of $u_1(x_i)$ and $u_2(x_i)$. It turns out that this is possible, provided that

$$\left(\frac{\partial f_1}{\partial U_2}\right)\left(\frac{\partial f_2}{\partial U_1}\right) \neq 1, \tag{3.11}$$

and, of course, that these partial derivatives exist. See the appendix for the derivation of (3.11). It should be apparent that (3.11) is not a very restrictive condition; it simply rules out degenerative utility expressions, such as $\alpha_1=\alpha_2=0$ in the additive structure. This means that for virtually ANY pairwise GAU structure, it is possible to solve for U_1 and U_2! Thus, altruistic utility clearly is mathematically tractable for two individuals.

Bell and Keeney's objection to GAU is a legitimate concern. It illustrates a common reaction to interdependent results, which can be described as follows: *person 2's utility is reflected in person 1's utility, which is then reflected in person 2's utility, which is then reflected in person 1's utility, etc.* This is an infinite process, and seems to be rather intuitively challenging. However, the implicit function theorem approach completely avoids this process by solving for the specific point at which the interdependencies are satisfied. While not a perfect analogy, this is similar to Nash's idea of directly solving for an equilibrium in game theory models (1951), thus sidestepping a more complex analysis of a sequence of best responses.

As an example, consider the following preference structures for the two altruistic individuals going out to dinner, where

$$U_1(x) = \left(\sqrt{x}\right) U_2(x) \tag{3.12}$$

and

$$U_2(x) = .5(1-x) + .5U_1(x). \tag{3.13}$$

In this example, x is a measure of how spicy the food at the restaurant is ($0 \le x \le 1$). Person 1 prefers spicier food (with diminishing returns), but cares more about Person 2's well-being. Person 2 prefers less spicy food, and places equal weight on her physical utility and Person 1's altruistic utility. To be precise, $u_1(x)=\sqrt{x}$, and $u_2(x)=1-x$. Person 1 has a purely multiplicative altruistic utility function, and Person 2 has an additive altruistic utility function. Using the implicit function theorem as described earlier, we require that $\left(\frac{\partial f_1}{\partial U_2}\right)\left(\frac{\partial f_2}{\partial U_1}\right)\neq 1$ to solve for U_1 and U_2. However, this is equal to $-.5\left(-\sqrt{x}\right)=.5\sqrt{x}$. Since $0\leq x\leq 1$, this is never equal to 1, and thus our example is solvable for all feasible values of x. If we solve (3.12) and (3.13) for U_1 and U_2, we find that:

$$U_1(x)=\frac{(1-x)\sqrt{x}}{\left(2-\sqrt{x}\right)} \tag{3.14}$$

and

$$U_2(x)=\frac{(1-x)}{\left(2-\sqrt{x}\right)}. \tag{3.15}$$

Figure 3.1a shows the resulting altruistic utility curves, along with the physical utility curves. Figure 3.1b shows the set of Pareto efficient values of x. That is, for any value outside of this set, there exists another alternative that is at least as desirable to both individuals, and strictly more desirable to at least one of them. These two individuals should undoubtedly choose a restaurant for which x is between 0.07 and 0.43. The most desirable outcome within that range would depend on either their relative bargaining powers, or some type of interpersonal utility comparison outside the scope of this paper.

Example Altruistic Preference Results for Two Individuals

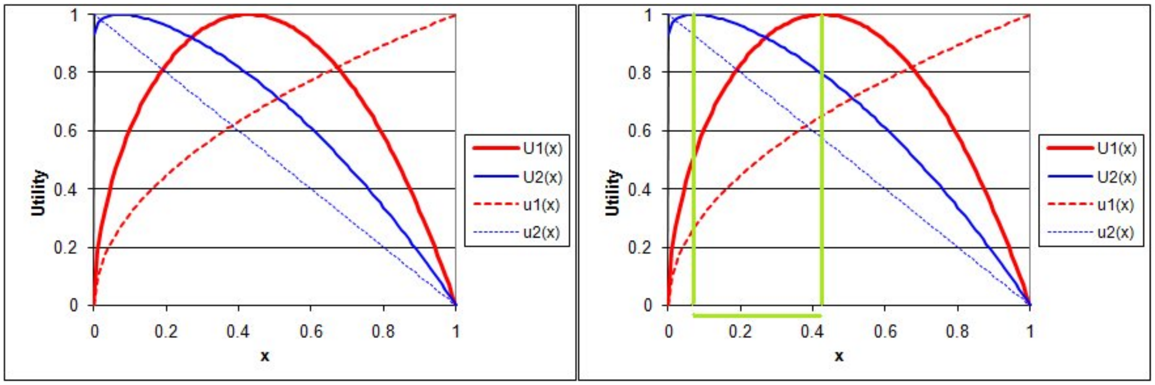

Figure 1a (left). The physical and altruistic utilities of two individuals whose preferences are described by (12) and (13).
Figure 1b (right). The pareto efficient, or non-dominated, alternatives for the two individuals.

3.7 Results for Multiple Individuals

The next step is to expand the model to more than two altruistic utility functions. Consider *n* individuals, each of whom has an altruistic utility function that incorporates his own physical utility, and the altruistic utilities of the other *n*-1 individuals. That is, we can write person *i*'s utility as:

$$U_i(x) = f_i\left(u_i(x), U_1(x), \ldots, U_{i-1}(x), U_{i+1}(x), \ldots, U_n(x)\right). \tag{3.16}$$

Our goal is now to write each $U_i(x)$ as a function of $u_1(x), \ldots, u_n(x)$. As in the pairwise case, the implicit function theorem can be used to determine the required conditions. With *n* individuals, the main condition is somewhat more abstract. We require that the determinant below is non-zero:

$$\det \begin{bmatrix} 1 & -\frac{\partial f_1}{\partial U_2} & \cdots & -\frac{\partial f_1}{\partial U_{n-1}} & -\frac{\partial f_1}{\partial U_n} \\ -\frac{\partial f_2}{\partial U_1} & 1 & & & \\ \cdots & & \ddots & & \\ -\frac{\partial f_{n-1}}{\partial U_1} & & & 1 & -\frac{\partial f_{n-1}}{\partial U_n} \\ -\frac{\partial f_n}{\partial U_1} & & & -\frac{\partial f_n}{\partial U_{n-1}} & 1 \end{bmatrix} \neq 0, \qquad (3.17)$$

(or equivalently that this matrix has full rank or is invertible), and again, that these partial derivatives all exist. See Appendix for details (under "Derivation of (3.17)"). Though more difficult to understand in terms of specific utility structures, this condition has the same general purpose as the main condition in the pairwise model. If this matrix does not have full rank, then we are in a higher-dimensional version of the "$U_1=U_2$, $U_2=U_1$" situation, where there is a degenerative interdependency among the altruistic utilities. Provided no such degenerative structure exists, it is possible to solve the *n*-person GAU model in terms of $U_1,\ldots, U_n$. This is a promising, but still somewhat vague result. We will now discuss an alternative formulation that fits very well into the GAU framework and has a simpler required condition.

Consider a group of *n* individuals. We create an artificial "group entity" as the *n*+1th person, and let each of the *n* individuals display altruism ONLY toward this group entity. That is, the individuals are concerned with the success of the group, but not directly with the utilities of the other individuals. However, we assert that the success or well-being of the group is permitted to depend (partially) on the well-being of each individual. We define the altruistic utility function of individual *i* as

$$U_i(x) = f_i(u_i(x), G(x)), \tag{3.18}$$

where G(*x*) is some measure of group success or well-being, defined as

$$G(x) = f_g(g(x), U_1(x), ..., U_n(x)), \tag{3.19}$$

with *g*(*x*) representing the basic effectiveness of outcome *x* toward achieving the goals of the group. We can think of *g*(*x*) as being analogous to u(*x*). (3.18) and (3.19) describe two intertwined relationships. (3.18) expresses the idea that each individual has preferences over how well the group is doing. It is similar in motive to the combination of self and group utilities used by Margolis (1982). (3.19) shows that the group's state is a function that includes the utilities of all of the individuals within it. For example, consider a volunteer social action group deciding on their next activity. Each individual will have a higher level of utility if the activity goes well. However, the success of the group depends partly on the individuals being happy and motivated to dedicate their time and energy to it. Thus, we have the same type of interrelated utility dynamic that we saw in the previous models. This could be interesting when considering individualistic vs. collectivistic cultures. For example, U.S. and Chinese cultures tend to differ significantly in terms how much they care about a group.

Notice that the individual utility functions now have a domain that is only two-dimensional, as opposed to the *n*-dimensional function shown in (3.16). As a consequence, we will see that the condition required to express $U_i(x)$ and $G(x)$ as a function of $u_1(x),\ldots, u_n(x)$ and $g(x)$ is much simpler.

Once again, we will make use of the implicit function theorem. However, the matrix that must now have full rank is much sparser, and the required condition reduces to

$$\sum_i \frac{\partial f_i}{\partial G}\frac{\partial f_g}{\partial U_i} \neq 1 . \qquad (3.20)$$

Details can be found in the appendix. Notice that this is very similar to the condition needed for the general pairwise model. In fact, the two-person model is mathematically equivalent to the group model with one individual (plus the artificial group individual). As long as (3.20) holds, it is possible to express each individual's utility (and the group utility) purely as a function of $u_1(x),\ldots, u_n(x)$ and $g(x)$.

As an example, we will model a specific structure for the volunteer group mentioned above. In our example, there are eight individuals, each of whom has the utility function:

$$U_i(x) = .75(1-x) + .25G(x) . \qquad (3.21)$$

The group utility function is given by:

$$G(x) = .6\sqrt{x} + \sum_{i=1}^{8} .05U_i(x) . \qquad (3.22)$$

In this particular context, *x* represents the amount of time each individual spends promoting the group's next activity ($0 \le x \le 1$). (For simplicity, we assume that all individuals spend the same amount of time promoting the activity.) In terms of direct physical outcomes, each individual prefers to spend as little time as possible doing this, but this is counteracted somewhat by the positive effect that it has on the group's success. The individuals only consider "fair" outcomes in which each person devotes the same amount of time. The group's success is determined mostly by the amount of time the individuals spend doing promotion (with diminishing returns), but also by the individuals' levels of utility.

To be certain that the model is solvable, it must satisfy (3.20). It is straightforward to compute $\frac{\partial f_i}{\partial G}\frac{\partial f_g}{\partial U_i} = (-.25)(-.05) = .0125$. Since there are eight individuals, it is clear that $\sum_i \frac{\partial f_i}{\partial G}\frac{\partial f_g}{\partial U_i} = 8(.0125) = .1 \ne 1$, and therefore it is possible to express $U_i(x)$ and $G(x)$ as a function of $u_1(x), \ldots, u_n(x)$ and $g(x)$. If we solve (3.21) and (3.22) for $U_i(x)$ and $G(x)$, we obtain:

$$U_i(x) = \frac{5}{6} + \frac{1}{6}\sqrt{x} - \frac{5}{6}x \qquad (3.23)$$

and

$$G(x) = \frac{1}{3} + \frac{2}{3}\sqrt{x} - \frac{1}{3}x. \qquad (3.24)$$

These utility curves are shown in Figure 3.2. From the perspective of the individuals in the group, the most desirable choice of x is 0.01. If they were not altruistic, their most desirable choice would be 0. The effect of including altruism is that a very small amount of promotion by the individuals is Pareto superior to none; it is preferred by all individuals, and improves the well-being of the group.

Example Altruistic Preference Results for a Group of Individuals

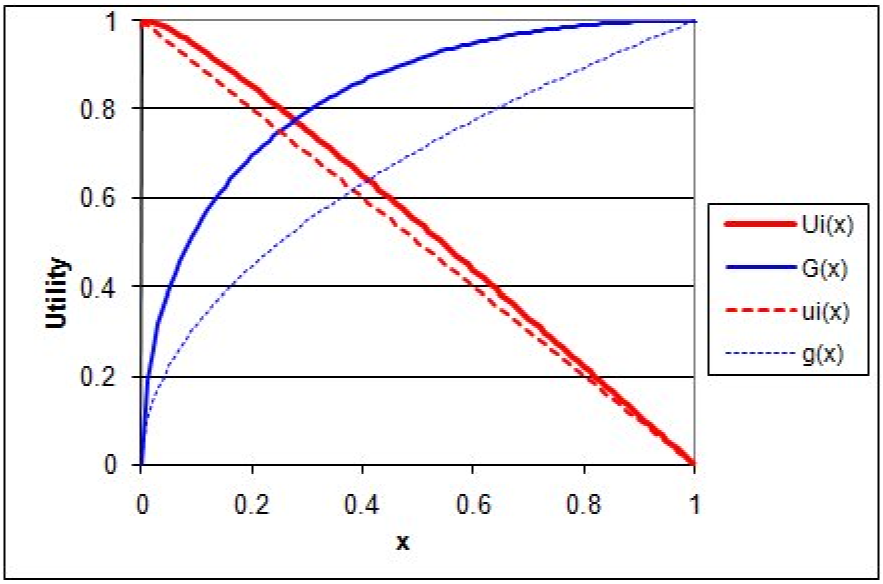

Figure 3.2. The physical and altruistic utilities of an individual and the group whose preferences are described by (3.21) and (3.22).

3.8 Results for Multiple Groups

The last step we will take in this analysis is an extension of the previous model to multiple groups. We allow each individual's altruistic utility function to consist only of his/her own physical utility and the well-being of the group to which the individual belongs. However, the utility of each group is a function of not only its members' utilities, but also the utilities of other groups. For example, consider multiple neighboring tribes: on an individual level, members interact only with their own tribe, but the tribes themselves may interact on a macro level through trade, competition for resources, warfare, etc.

We will restrict our analysis here to a two-group model. The utilities are shown in (3.25). In this structure, *k* of the *n* total individuals belong to group 1, and *n-k* belong to group 2:

$$\begin{aligned}
&U_1(x) = f_1(u_1(x), G_1(x)) \\
&\vdots \\
&U_k(x) = f_k(u_k(x), G_1(x)) \\
&U_{k+1}(x) = f_{k+1}(u_{k+1}(x), G_2(x)) \\
&\vdots \\
&U_n(x) = f_n(u_n(x), G_2(x)) \\
&G_1(x) = f_{G_1}(g_1(x), U_1(x), \ldots, U_k(x), G_2(x)) \\
&G_2(x) = f_{G_2}(g_2(x), U_{k+1}(x), \ldots, U_n(x), G_1(x))
\end{aligned} \tag{3.25}$$

Once again, the goal is to be able to express each individual's altruistic utility as a function of the physical utilities. Using the implicit function theorem, the required condition reduces to:

$$\left(1 - \sum_{i=1}^{k} \frac{\partial f_i}{\partial G_1} \frac{\partial G_1}{\partial U_i}\right)\left(1 - \sum_{i=k+1}^{n} \frac{\partial f_i}{\partial G_2} \frac{\partial G_2}{\partial U_i}\right) \neq \frac{\partial G_1}{\partial G_2} \frac{\partial G_2}{\partial G_1} \tag{3.26}$$

The derivation of (3.26) is in the appendix. As in the earlier models, this condition rules out degenerative situations. With increasingly complex models, it becomes even more important to check the existence condition, since violations of it are likely to be much less transparent.

Extension to a more general model with *m* groups is possible. For the sake of brevity, we express the model simply by showing the utility functions for an arbitrary individual *i* and group *j*:

$$\begin{aligned} U_i(x) &= f_i\left(u_i(x), G_{B(i)}(x)\right) \\ G_j(x) &= f_{G_j}\left(g_j(x), U_{j1}(x), \ldots, U_{jk_j}(x), G_1(x), \ldots, G_{j-1}(x), G_{j+1}(x), \ldots, G_m(x)\right) \end{aligned} \tag{3.27}$$

where $B(i)$ is the group to which individual *i* belongs, U_{jk} is the *k*th individual in group *j*, and k_j is the number of individuals in group *j*. Unfortunately, there is no simple existence condition for the *m*-group model. The interactions among the *m* groups create the same difficulty as we observed earlier in the *n*-individual model (without groups): the determinant of the matrix of partial derivatives does not reduce to a simple expression. In particular *m*-group examples with sparse interaction (i.e. most group utilities do not directly affect one another), it may still be possible to determine reasonably simple solvability conditions.

3.9 Conclusion

Most individuals clearly incorporate the well-being of others into their own decision making, whether implicitly or explicitly. We make no judgments here about whether this is being done for selfless reasons or not. We simply set out to accurately model the preference structures that underlie many people's decisions. Altruism is generally not incorporated into utility models, largely due to analytical and philosophical complexity. A major obstacle in modeling and implementing altruistic utility functions is the often complicated dynamic of interdependent utilities that can occur.

In this paper, we have examined the general altruistic utility concept in several different models. We began with a specific (additive) pairwise model, for which we were able to compute a specific result. We then extended this to a general pairwise model, which had a simple condition for solvability. We expanded this to a general n-person GAU model, for which we also determined a solvability condition. We then developed an alternative n-person GAU model using the concept of group utility, which simplified the required condition, and made it clear that it was not overly restrictive. Finally, we expanded this to a multiple-group model, and showed that we can also determine the solvability conditions when two different groups interact with one another.

We have shown that, in general, incorporating altruism into preference models is not an analytically insurmountable task. It is nearly always possible to resolve the utility interdependencies. We hope that, given these possibility results, further effort can be made in the future to develop effective altruistic decision models. It is also possible to extend these results to other settings in which individual outcomes have positive or negative externalities, and we would certainly recommend this as an avenue for further research.

3.10 Appendix

Derivation of (3.7) and (3.8):

Substituting (3.6) into (3.5) yields:

$$U_1(a_i) = \alpha_1 u_1(a_i) + (1-\alpha_1)\left[\alpha_2 u_2(a_i) + (1-\alpha_2)U_1(a_i)\right]$$

$$U_1(a_i) - (1-\alpha_1)(1-\alpha_2)U_1(a_i) = \alpha_1 u_1(a_i) + (1-\alpha_1)\alpha_2 u_2(a_i)$$

$$U_1(a_i) = \frac{\alpha_1 u_1(a_i) + (1-\alpha_1)\alpha_2 u_2(a_i)}{1-(1-\alpha_1)(1-\alpha_2)}.$$

The derivation of (3.8) is similar.

Proof that $U_1(x_i) = U_2(x_i)$ requires $u_1(x_i) = u_2(x_i)$ in the pairwise additive model:

Setting the expressions for $U_1(x_i)$ and $U_2(x_i)$ from (3.7) and (3.8) equal to one another yields:

$$\frac{\alpha_1 u_1(x_i) + (1-\alpha_1)\alpha_2 u_2(x_i)}{1-(1-\alpha_1)(1-\alpha_2)} = \frac{\alpha_2 u_2(x_i) + (1-\alpha_2)\alpha_1 u_1(x_i)}{1-(1-\alpha_1)(1-\alpha_2)}$$

This can be simplified as follows:

$$\alpha_1 u_1(x_i) + (1-\alpha_1)\alpha_2 u_2(x_i) = \alpha_2 u_2(x_i) + (1-\alpha_2)\alpha_1 u_1(x_i)$$
$$\alpha_1 u_1(x_i) - (1-\alpha_2)\alpha_1 u_1(x_i) = \alpha_2 u_2(x_i) - (1-\alpha_1)\alpha_2 u_2(x_i)$$
$$\alpha_1 \alpha_2 u_1(x_i) = \alpha_1 \alpha_2 u_2(x_i)$$
$$u_1(x_i) = u_2(x_i)$$

Thus, $U_1(x_i) = U_2(x_i)$ implies $u_1(x_i) = u_2(x_i)$.

Derivation of (3.11):

The implicit function theorem tells us that if

$$\hat{f}_1(x_1,\ldots,x_N) = 0$$
$$\ldots$$
$$\hat{f}_n(x_1,\ldots,x_N) = 0$$

where $n<N$, then we can write $x_1,\ldots,x_n$ as functions of $x_{n+1},\ldots,x_N$ if and only if

$$D = \begin{bmatrix} \frac{\partial \hat{f}_1}{\partial x_1} & \cdots & \frac{\partial \hat{f}_1}{\partial x_n} \\ \cdots & \ddots & \\ \frac{\partial \hat{f}_n}{\partial x_1} & & \frac{\partial \hat{f}_n}{\partial x_n} \end{bmatrix}$$

has full rank. (Having full rank is equivalent to being invertible, and also equivalent to having a non-zero determinant.) D is the matrix of partial derivatives with respect to the first n variables. If D has full rank, then we can write:

$$x_1 = g_1(x_{n+1},\ldots,x_N)$$
$$\ldots$$
$$x_n = g_n(x_{n+1},\ldots,x_N)$$

This is the main result of the implicit function theorem.

In this case, n=2 (U_1 and U_2), and N=4 (U_1, U_2, u_1, and u_2). We would like to write U_1 and U_2 as functions of u_1 and u_2. By moving all terms in (3.9) and (3.10) to the left side of the equations, we obtain:

$$U_1(a_i) - f_1(u_1(a_i), U_2(a_i)) = 0$$
$$U_2(a_i) - f_2(u_2(a_i), U_1(a_i)) = 0$$

The matrix of partial derivatives with respect to U_1 and U_2 is

$$\begin{bmatrix} 1 & -\frac{\partial f_1}{\partial U_2} \\ -\frac{\partial f_2}{\partial U_1} & 1 \end{bmatrix}.$$

We can express U_1 and U_2 as functions only of u_1 and u_2 if and only if this determinant is non-zero. The determinant is $(1)(1)-\left(-\frac{\partial f_1}{\partial U_2}\right)\left(-\frac{\partial f_2}{\partial U_1}\right)$, or $1-\left(\frac{\partial f_1}{\partial U_2}\right)\left(\frac{\partial f_2}{\partial U_1}\right)$. Thus, this determinant is non-zero provided that $\left(\frac{\partial f_1}{\partial U_2}\right)\left(\frac{\partial f_2}{\partial U_1}\right) \neq 1$.

Derivation of (3.17):

We have n equations of the form shown in (3.16). There are $2n$ total variables: $U_1,\ldots,U_n$, and $u_1,\ldots,u_n$. We would like to express $U_1,\ldots,U_n$ in terms of $u_1,\ldots,u_n$. Per the implicit function theorem, we need to compute the n x n matrix of the partial derivatives of the n equations with respect to $U_1,\ldots,U_n$. First, we move all terms to the left side of each equation, yielding

$$U_i(x) - f_i\left(u_i(x), U_1(x), \ldots, U_{i-1}(x), U_{i+1}(x), \ldots, U_n(x)\right) = 0.$$

The partial derivative of the ith equation with respect to U_i is 1, and the partial derivative with respect to U_j ($i \neq j$) is $\frac{\partial f_i}{\partial U_j}$. Therefore, this n x n matrix is:

$$\begin{bmatrix} 1 & -\frac{\partial f_1}{\partial U_2} & \cdots & -\frac{\partial f_1}{\partial U_{n-1}} & -\frac{\partial f_1}{\partial U_n} \\ -\frac{\partial f_2}{\partial U_1} & 1 & & & \\ \cdots & & \ddots & & \\ -\frac{\partial f_{n-1}}{\partial U_1} & & & 1 & -\frac{\partial f_{n-1}}{\partial U_n} \\ -\frac{\partial f_n}{\partial U_1} & & & -\frac{\partial f_n}{\partial U_{n-1}} & 1 \end{bmatrix}.$$

As in the pairwise case, the implicit function theorem tells us that if (and only if) this matrix has a non-zero determinant, we can solve for $U_1,\ldots,U_n$ in terms of $u_1,\ldots,u_n$.

Derivation of (3.20):

First, move all terms in (3.18) and (3.19) to the left side, yielding:

$$U_1(x) - f_1(u_1(x), G(x)) = 0$$
$$\vdots$$
$$U_n(x) - f_n(u_n(x), G(x)) = 0$$
$$G(x) - f_g(g(x), U_1(x), \ldots, U_n(x)) = 0$$

We would like to express $U_1, \ldots, U_n$, and G in terms of $u_1, \ldots, u_n$ and g. Per the implicit function theorem, we need to construct the *n*+1 x *n*+1 matrix of partial derivatives. The partial derivative of the *i*th equation ($1 \le i \le n$) with respect to *i* is 1. The partial derivative with respect to U_j ($i \ne j, 1 \le j \le n$) is 0, since U_j does not appear in any individual utility functions other than the *j*th one. The partial derivative with respect to G is $-\frac{\partial f_i}{\partial G}$. The partial derivative of the last equation with respect to U_i is $-\frac{\partial f_g}{\partial U_i}$, and the partial derivative with respect to G is 1. Therefore, the matrix of partial derivatives is:

$$\begin{bmatrix} 1 & 0 & \ldots & 0 & -\frac{\partial f_1}{\partial G} \\ 0 & 1 & & & \\ \vdots & & \ddots & & \\ 0 & & & 1 & -\frac{\partial f_n}{\partial G} \\ -\frac{\partial f_g}{\partial U_1} & & & -\frac{\partial f_g}{\partial U_n} & 1 \end{bmatrix}.$$

That is, all entries are zero, except for the main diagonal, the rightmost column, and the bottom row. Again, we require that the determinant be non-zero. However, given this structure, we can simplify the expression for the determinant. A formula for computing the determinant of an *m* x *m* matrix *D* is:

$$\sum_{\sigma} sgn(\sigma) \prod_i D_{i,\sigma(i)},$$

where the σ are all possible permutations of *m* elements, and $sgn(\sigma)$ is 1 for even permutations and -1 for odd permutations. (The identity permutation is considered even). This is often overly complex to compute for large matrices. However, notice that in the matrix above, nearly all of the permutations will result in $\prod_i D_{i,\sigma(i)} = 0$. There are only *n*+1 permutations for which this product is non-zero. The identity permutation yields a product of 1. The other permutations are those which involve only one switch, where one of the elements is the *n*+1th element (these are odd permutations). When the *n*+1th element and the *i*th element are switched, the resulting product is $\frac{\partial f_i}{\partial G}\frac{\partial f_g}{\partial U_i}$ (the x's are omitted for clarity). Thus, the determinant is equal to $1 - \sum_i \frac{\partial f_i}{\partial G}\frac{\partial f_g}{\partial U_i}$, so as long as $\sum_i \frac{\partial f_i}{\partial G}\frac{\partial f_g}{\partial U_i} \ne 1$, we can solve for $U_1, \ldots, U_n$, and G in terms of $u_1, \ldots, u_n$ and g.

Derivation of (3.26):

Using the implicit function theorem as in the previous models, we obtain the following matrix of partial derivatives:

$$\begin{bmatrix}
1 & 0 & 0 & & & & & & -\frac{\partial f_1}{\partial G_1} & 0 \\
0 & 1 & 0 & & & & & & -\frac{\partial f_2}{\partial G_1} & 0 \\
0 & 0 & 1 & & & & & & -\frac{\partial f_3}{\partial G_1} & 0 \\
 & & & \ddots & & & & & \vdots & \\
 & & & & & & & & -\frac{\partial f_k}{\partial G_1} & 0 \\
 & & & & & & & & 0 & -\frac{\partial f_{k+1}}{\partial G_2} \\
 & & & & & & & & & \vdots \\
 & & & & & & & & 0 & -\frac{\partial f_n}{\partial G_2} \\
-\frac{\partial G_1}{\partial U_1} & -\frac{\partial G_1}{\partial U_2} & -\frac{\partial G_1}{\partial U_3} & \cdots & -\frac{\partial G_1}{\partial U_k} & 0 & & 0 & 1 & -\frac{\partial G_1}{\partial G_2} \\
0 & 0 & 0 & & 0 & -\frac{\partial G_2}{\partial U_{k+1}} & \cdots & -\frac{\partial G_2}{\partial U_n} & -\frac{\partial G_2}{\partial G_1} & 1
\end{bmatrix}$$

As before, we require that this matrix has full rank, i.e. that its determinant is not equal to zero. The determinant can be written as:

$$\sum_{\sigma} sgn(\sigma) \prod_{i} D_{i,\sigma(i)}$$

Notice that the upper left *n* x *n* submatrix is simply the identity matrix. As in the derivation of (3.20), this matrix is sparse enough such that the products associated with most of these permutations are equal to zero. To be precise, there are *k**(*n*-*k*) + *n* + 1 non-zero products. The determinant can be written as:

$$\sum_{i=1}^{k} \sum_{j=k+1}^{n} \frac{\partial f_i}{\partial G_1} \frac{\partial G_1}{\partial U_i} \frac{\partial f_j}{\partial G_2} \frac{\partial G_2}{\partial U_j} - \sum_{i=1}^{k} \frac{\partial f_i}{\partial G_1} \frac{\partial G_1}{\partial U_i} - \sum_{i=k+1}^{n} \frac{\partial f_i}{\partial G_2} \frac{\partial G_2}{\partial U_i} - \frac{\partial G_1}{\partial G_2} \frac{\partial G_2}{\partial G_1} + 1,$$

When this expression is set not equal to zero, the inequality reduces to:

$$\left(1 - \sum_{i=1}^{k} \frac{\partial f_i}{\partial G_1} \frac{\partial G_1}{\partial U_i}\right)\left(1 - \sum_{i=k+1}^{n} \frac{\partial f_i}{\partial G_2} \frac{\partial G_2}{\partial U_i}\right) \neq \frac{\partial G_1}{\partial G_2} \frac{\partial G_2}{\partial G_1}.$$

3.11 References

Alexander, R. D. 1987. *The Biology of Moral Systems*. Aldine de Gruyter, New York.

Andreoni, J. 1990. Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal* **100**(401) 464-477.

Andreoni, J., J. K. Scholz. 1998. An Econometric Analysis of Charitable Giving with Interdependent Preferences. *Economic Inquiry* **36**(3) 410-428.

Arrow, K. J. 1975. Gifts and Exchanges. *Altruism, Morality and Economic Theory* (ed. E. S. Phelps). Russell Sage Foundation.

Bell, D. E., R. L. Keeney. 2009. Altruistic Utility Functions for Joint Decisions. In *The Mathematics of Preference, Choice, and Order: Essays in Honor of Peter C. Fishburn*. S. J. Brams, W. V. Gehrlein, and F. S. Roberts (eds.), Springer-Verlag, New York.

Bergstrom, T. C., O. Stark. 1993. How Altruism Can Prevail in an Evolutionary Environment. *The American Economic Review* **83**(2) 149-155.

Boyd, R., H. Gintis, S. Bowles, P. J. Richerson. 2003. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences of the United States of America* **100**(6) 3531-3535.

Chen, K.Y., C. R. Plott. 2002. Information Aggregation Mechanisms: Concept, Design, and Implementation for a Sales Forecasting Problem. California Institute of Technology Social Science Working Paper No. 1131, March.

Fehr, E., S. Gächter. 2002. Altruistic punishment in humans. *Nature* **415** 137-140.

Forsythe, R., M. Frank, V. Krishnamurthy, T. W. Ross. 1995. Using Market Prices to Predict Election Results: The 1993 UBC Election Stock Market. *The Canadian Journal of Economics* **28**(4a) 770-793.

Forsythe, R., T. A. Rietz, T. W. Ross. 1999. Wishes, expectations and actions: a survey on price formation in election stock markets. *Journal of Economic Behavior & Organization* **39**(1) 83-110.

Hamilton, W. D. 1963. The evolution of altruistic behavior. *The American Naturalist* **97**(896) 354-356.

Keeney, R. L., H. Raiffa. 1976. Decisions with Multiple Objectives. Wiley, New York (Reprinted in 1993, Cambridge University Press, New York).

Margolis, H. 1982. *Selfishness, Altruism and Rationality*. Cambridge University Press.

McCardle, K. F., K. Rajaram, C. S. Tang. 2009. A Decision Analysis Tool for Evaluating Fundraising Tiers. *Decision Analysis* **6**(1) 4-13.

Narens, L., R. D. Luce. 1983. How We May Have Been Misled Into Believing in the Interpersonal Comparability of Utility. *Theory of Decision* **15** 247-260.

Nash, J. 1951. Non-Cooperative Games. *The Annals of Mathematics* **54**(2) 286-295.

Nowak, M. A., K. Sigmund. 2005. Evolution of indirect reciprocity. *Nature* **437** 1291-1298.Robbins, L. 1938. Interpersonal Comparisons of Utility: A Comment. *The Economic Journal* **48**(192) 635-641.

Savage, L. 1954. *The Foundations of Statistics*. John Wiley, New York.

Sen, A. 1979. Interpersonal comparisons of welfare. *Economics and Human Welfare*. M. Boskin (ed.). Academic Press, New York. 183-201

Simon, H. A. 1990. A Mechanism for Social Selection and Successful Altruism. *Science* **250**(4988) 1665-1668.

Sugden, R. 1982. On the economics of philanthropy. *Economic Journal* **92** 341-350.

Titmuss, R. 1971. *The Gift Relationship: From Human Blood to Social Policy*. Pantheon Books, New York.

Trivers, R. L. 1971. The evolution of reciprocal altruism. *The Quarterly Review of Biology* **46**(1) 35-57.