INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

University of Alberta

Benefit Transfer with Multiple Sources of Heterogeneity in Non-Market Valuation Random Utility Models

By

Xiaosong Xu ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirement for the degree of Doctor of Philosophy

in

Agricultural Economics

Department of Rural Economy

Edmonton, Alberta

Fall 1997

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-23093-7

Canada

University of Alberta

Release Form

Name of the Author: Xiaosong Xu

Title of Thesis: Benefit Transfer with Multiple Sources of Heterogeneity in Non-market Valuation Random Utility Models

Degree: Doctor of Philosophy

Year this Degree Granted: 1997

Permission is hereby granted to the University of Alberta Library to produce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Xu Xiaosong

Xiaosong Xu
11131 85 Ave.
Edmonton, AB
Canada, T6G 0W7

July 16, 1997

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Benefit Transfer with Multiple Sources of Heterogeneity in Non-Market Valuation Random Utility Models** submitted by Xiaosong Xu in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Agricultural Economics.

(Dr. W. L. Adamowicz, Supervisor)

(Dr. S. R. Jeffrey, Committee Chair)

(Dr. M. M. Veeman)

(Dr. T. S. Veeman)

(Dr. M. K. Luckert)

(Dr. D. Young)

(Dr. G. C. van Kooten)

July 15, 1997

Dedicated to my brothers, parents and son

Abstract

This thesis employs two sets of stated preference data to investigate the transferability of nonmarket benefits across regions. It focuses on two aspects: benefit estimation with multiple sources of heterogeneity and statistical and economic significance tests for benefit transfers.

A heterogeneous multinomial logit model and a random coefficient multinomial probit model are developed and applied to the two data sets. It is found that accounting for the heterogeneity in choice data not only improves the model's goodness of fit but also affects benefit calculations. While the explicit formula for benefit calculation is derived in the heterogeneous model, the benefit measures are simulated in the random coefficient model. The random coefficient model also provides the best goodness of fit and benefit transferabilities.

Based on the benefits calculated from the random coefficient probit model, transfers are examined using two nonparametric procedures: the Mann-Whitney test and the convolutions approach. It is suggested that most of benefits generated from the "true" and transferred models are transferable statistically at a 10% level.

A test procedure for the economic significance of benefit transfer is also developed. Two indicators, the probability of making incorrect decisions and the expected benefit of benefit transfer, are used for the test of the economic significance of benefit transfer. For a specific case study, it is found that the possibility of making an incorrect decisions is reasonablely low when benefit transfer is applied.

Acknowledgements

Sincere gratitude is expressed to Dr. W. L. Adamowicz, my supervising professor, for his strong technical support and well-organized guidance throughout the present study. His encouragement and patience in helping me to improve this study and my writing are especially appreciated.

Gratitude is extended to Dr. S. R. Jeffrey, Dr. T. S. Veeman, Dr. M. M. Veeman and Dr. F. S. Novak for their understanding and timely support which make my study in the Department of Rural Economy possible. I also wish to acknowledge the financial support of the Network of Centres of Excellence on Sustainable Forest Management. I am grateful to Mr. P. C. Boxall and Dr. W.L Adamowicz for the data provided to me.

I am also grateful to my committee members, Dr. M.K. Luckert, Dr. S.R. Jeffrey, Dr. G.C. van Kooten, Dr. D. Young, Dr. M.M. Veeman and Dr. T.S. Veeman, for their valuable suggestions, comments and advises.

Special thanks are given to Jim, Wendy, Liz, Judy, Dawn and the graduate students in the Department of Rural Economy for their friendly companionship and willingness to help, making my years of studying in the department very rewarding.

Finally, I would thank my wife and son for their understanding and patience.

Table of Contents

Chapter 1	Introduction	1
1.1	Background on Benefit Transfer	2
1.2	Benefit Transfer in Practice	7
1.3	Issues in Benefit Transfer	8
1.4	Study Objectives	10
1.5	The Organization of the Study	11
Chapter 2	The Random Utility Model of Choice and Benefit Estimation	13
2.1	Introduction	13
2.2	Random Utility of Choice	15
2.3	Benefit Estimation in Random Utility Models	21
Chapter 3	The Stated Preference Approach and Survey Data	26
3.1	The Stated Preference Approach	26
3.2	Survey Data	31
Chapter 4	Benefit Transfer Using Heterogenous Multinomial Logit Models	37
4.1	Introduction	37
4.2	Heterogeneity in the Multinomial Logit Model	39
4.3	Benefit Calculation in the Heterogenous MNL Model	43
4.4	Model Estimation...	45
4.5	Model and Benefit Transferability Evaluations	50
4.6	Conclusions	55
Chapter 5	Benefit Transfer Using Random Coefficient Multinomial Probit Models	57
5.1	Introduction	57
5.2	The RCMP Model and Heterogeneity in Choice Data	59
5.3	RCMP Model Estimation and Identification	62
5.4	Benefit Estimation in the RCMP Model	67
5.5	RCMP Model Estimation Results	69
5.6	Benefit Transferability Evaluation	78
5.7	Conclusion	83
Chapter 6	Statistical Tests for Benefit Transferability Using Nonparametric Procedures	85
6.1	Introduction	85
6.2	Methods of Testing for Underlying Distributions	87
6.3	Convolutions Approach	96
6.4	Test Results and Discussion...	101
Chapter 7	Test for the Economic Significance of Benefit Transfers	105

7.1 **Introduction** 105
7.2 **The Policymaker's Problem in Benefit Transfer** 107
7.3 **Test for the Economic Significance of Benefit Transfer** 113
7.4 **Summary and Conclusions** 120

Chapter 8 **Conclusions and Discussions** 123

References 141

Table of Tables

Table 1. Attributes and Levels Used in the Stated Preference Experiments 34

Table 2. Maximum Likelihood Estimates of the HMNL Models....................................48

Table 3. Maximum Likelihood Estimates of the TMNL Models 49

Table 4. Objective Attribute Levels of Alberta WMUs............................ 53

Table 5. Benefit Estimates from True and Transferred Models................. 54

Table 6. Maximum Likelihood Estimates of the MNL and
IMNL Models, Alberta 72

Table 7. Maximum Likelihood Estimates of the MNL and
IMNL Models, Saskatchewan 73

Table 8. Maximum Likelihood Estimates of the RCMP
Models, Alberta and Saskatchewan 76

Table 9. Benefit Estimates from the True and Transferred RCMP Models 81

Table 10. Simulated Benefits of Eliminating a Site and
the Probability of Visiting the Site 82

Table 11. Wald Statistic of the Normality Test for the Simulated
Empirical True and Transferred Benefit Distributions..91

Table 12. Statistical Test Results for Benefit Transferability 102

Table 13. The Probability of Making an Incorrect Decision
at Different Per Trip Market Costs 117

Table 14. Policymaker's Benefits From Benefit Transfer
Under Different Market Costs 119

Table of Figures

Figure 1. Study Plan 128

Figure 2. Example of the instrument used to gather stated preference data 129

Figure 3. PDF of the True and Transferred Benefits for Policy Scenario 1 130

Figure 4. PDF of the True and Transferred Benefits for Policy Scenario 2 131

Figure 5. PDF of the True and Transferred Benefits for Policy Scenario 3 132

Figure 6. PDF of the True and Transferred Benefits for Policy Scenario 4 133

Figure 7. PDF of the True and Transferred Benefits for Policy Scenario 5 134

Figure 8. PDF of the True and Transferred Benefits for Policy Scenario 6 135

Figure 9. PDF of the True and Transferred Benefits for Policy Scenario 7 136

Figure 10. PDF of the True and Transferred Benefits for Policy Scenario 8 137

Figure 11. Probability of Making an Incorrect Decision under
Different Market Costs (Policy Scenario 8) 138

Figure 12. Probability of Making an Incorrect Decision under
Different Market Costs (Policy Scenario 5) 139

Chapter 1. Introduction

Benefits of environmental changes are measured by comparing an existing known level of environmental services with some specified hypothetical alternative where the environmental services have been changed. The concept of the benefit of environmental good and service changes is a natural extension of the compensating or equivalent variation of marketed goods and services. In recent years, demand models suitable for the estimation of non-market benefit values have been developed. This has enabled economists to use the concept of Hicksian surplus for the valuation of environmental goods and services. Monetarily quantified benefit measures provide an important tool for natural resource management and policy analysis.

Motivated by both intensive resource management and assorted legislative and juridical mandates, public and private agencies have expended considerable resources in order to quantify the economic consequences of altering service flows and stocks of non-marketed features of the natural environment. It is expected that the demand for non-market valuation will continue to increase.

The increasing demand for non-market valuation of environmental assets as well as financial constraints on researchers has brought about the idea of using benefit transfers as a time-saving and cost-effective alternative to environmental valuation. The term benefit transfer aptly captures its objective: transfer the estimated economic value of environmental quality changes from one site to a second site. In the case of natural resource and environmental policies and projects, benefit transfer involves transferring value estimates from a "study site" to a "policy site" where sites can vary across geographic space and/or time. For example, benefits of improving moose habitat in Newfoundland may be estimated by using a model from a study that estimated the benefits of improving moose habitat in Alberta.

The demand for benefit transfer applications by public policy makers and resource management agencies is expected to increase in the future for three interrelated reasons. First, primary data collection on a site-by-site basis is expensive. Second, agencies face considerable uncertainty regarding continued budget support for primary data collection and valuation model development. Third, primary data collection is time consuming, often taking one or more years to complete a study. Policy or management decision makers require inexpensive benefit estimates in a timely manner. Benefit transfer offers an opportunity to meet this need.

1.1 Background on Benefit Transfer

While the objective of benefit transfer is straightforward, conceptual discussions as to what is a benefit transfer application have been ongoing for several years. Boyle and Bergstrom (1992) described benefit transfer as the transfer of existing estimates of non-market values to a new study that is different from the study for which the values were originally estimated. In their opinion, benefit transfer is simply the application of secondary data to a new policy issue. McConnell (1992) defined benefit transfer as a process by which researchers take recreational demand models or other models that are estimated for one site or region and apply them to another site or region. Smith (1992) suggested the process of a benefit transfer involves focusing on measuring (in dollars) how much the people affected by some policy will gain from it. He states that benefit transfer is not a forecast, and usually does not attempt to predict other exogenous influences on people's behaviour. Instead, a predefined set of conditions is assumed to characterize the nonpolicy variables, and thus benefit estimates are derived by focusing on the effects of the conditions assumed to be changed by the policy. Atkinson et al. (1992) defined benefit transfer by confronting the policy maker's choice: in considering a demand for yet another site-specific environmental improvement, policy makers must decide whether to extrapolate the results of benefit assessments done elsewhere or to commission a new assessment study. Luken et al. (1992) view the process of benefit transfers as a limit-setting process: we use these existing studies only to suggest some likely limits on willingness to pay.

To summarize, benefit transfer is an application of a data set, and its estimated model, that were developed for one particular use to a quite distinct alternative application. Thus the process of a benefit transfer is recognized to be potentially less than ideal.

Implicit in all of the above is the notion that benefit transfers are valid under well-defined conditions. However, there is a so-called impossibility philosophy toward benefit transfer. This philosophy states that benefit transfer is impossible. Consider the transfer of a value estimate at site A to a similar issue at site B. Those who adhere to the impossibility philosophy have two major points (Boyle and Bergstrom, 1992): first, the levels of common technical attributes of site A and site B are different and some attributes might occur at one site and not at the other; and secondly, there may exist some unobserved social and economic heterogeneity across individuals and between site A and site B. The potential sources of error in benefit transfer could be commodity measurement error, population characteristic measurement error, welfare change measurement error, physio economic linkages measurement error, and estimation procedure and judgment error. Proponents of this philosophy would argue that these potential errors may result in different values at each of the sites, so benefit transfer is impossible. The retort of the pragmatist might be that we simply need to learn whether and how these differences actually affect values, and if they do, whether the differences are large and how might one statistically control for these differing effects.

Regardless of the conceptual arguments, researchers have proceeded to evaluate the viability and the methodologies of benefit transfers. In recent years, a number of studies have been conducted to test empirically the feasibility of benefit transfer. Parsons and Kealy (1994) used revealed preference data of lake recreation in the state of Wisconsin to test the viability of benefit transfer in a Random Utility Model (RUM). They found that the estimated models were significantly different in the statistical test, but the benefit transfer had considerable accuracy. These results are not surprising because some significantly different coefficients in their models have little impact on benefit calculation or work to offset one another. However, it has to be recognized that the base of benefit transfer is model transferability, and that a good benefit transfer from poorly transferable models is by chance.

Downing and Ozuna (1996) introduced an experiment designed to test the reliability of the benefit function transfer approach using contingent valuation methods. Using data collected from anglers surveyed across eight contiguous Texas Gulf Coast bay regions over three distinct time periods, they tested function transferability by using dummy variables to represent different regions and time periods. They tested benefit transferability by computing benefit confidence intervals. Their results are the opposite of Parson and Kealy's. In Downing and Ozuna's study, most of the contingent valuation functions were transferable because the dummy variables were insignificant, but the computed benefits were significantly different. However, the question in their study is "how large is large" economically in the determination of the significance of the coefficients of the dummies. The authors did not provide the estimates of the contingent valuation functions. It is possible that the magnitude of some estimated dummy coefficients are quite high but insignificant statistically. As a result, function transferability was accepted but benefit transferability was rejected.

Loomis (1992) introduced an approach that involves an application of travel cost demand equations and contingent valuation benefit functions from existing sites to the new site to test the assumptions of benefit transfer from recreation sites in one state to another state for the same recreation activity. In Loomis (1992) the equality of demand coefficients for ocean sport salmon fishing in Oregon versus Washington and for freshwater steelhead fishing in Oregon versus Idaho is rejected. Thus transfer of either demand equations or average benefits per trip is likely to be in error.

Based on a large number of previous non-market valuation studies, Desvouges, Naughton and Parsons (1992) investigated the problems encountered in using existing studies to measure the benefits of water quality improvements and proposed criteria for selecting transfer studies. They indicated that although benefit transfer may offer promise, the fact that existing studies were not designed for transfer places severe limitations on the current effectiveness of transfer. Their general recommendations for how benefit transfer might be improved are: (1) focusing on multisite models; (2) comparing the estimates of the same basic structure derived from different geographic areas; (3) using quality measures that are relevant to policy decisions; (4) selecting explanatory variables where possible that have measures readily available in current regional data bases.

Smith (1992) illustrated the need for guidelines in deciding when benefit transfer methods can be used to value changes in environmental resources. He recommends that four research directions would contribute to the knowledge base of benefit transfer: (1) continue to develop theory that links contingent valuation studies to consumer behaviour, (2) develop criteria that will standardize the description of key components of the valuation process such as the commodity and geographic description, (3) continue to explore the use of meta-analysis, and (4) begin a process of specifying a protocol for benefit transfers.

McConnell (1992) examined the methodology of a benefit transfer by forming an analytical model based on preferences. He suggested that one cannot separate the benefit estimation process from the benefit transfer process because potential errors arise in both demand estimation and benefit estimation. Brookshire and Neill (1992) shared the same idea with McConnell. They argued, that since the development of the methodologies and conceptual framework for non-market valuation is not complete, rigorous empirical studies and advancements in non-market valuation theory are needed in order to find a "home" for benefit transfer. The above literature suggests that benefit transfer is a new concept in environmental valuation and that studies on both benefit transfer theory and methods have just begun.

1.2 Benefit Transfer in Practice

Despite the recognized limitations of benefit transfer, the technique is widely used in practice. In the United States, benefit transfer has been used by government agencies to facilitate benefit-cost analysis of public policies and projects affecting natural resources. For water resource management, U.S. Army Corps of Engineers, U.S. Bureau of Reclamation, U.S. Natural Resources Conservation Service, and the Tennessee Valley Authority recommended benefit transfer techniques for measuring recreation benefits (U.S. Water Resources Council, 1983; Vincent et al,. 1986; Henderson and Allen, 1994). For forest and rangeland resources managed by the U.S. Forest Service, benefit transfer is used to estimate forest commodity benefits (including recreation benefits) in the National Forest program and planning process at national, regional and local levels (U.S. Forest Service, 1990). A well-known application of benefit transfer is the *U.S. Forest Services* [1987] development of RPA (Resource Planning Act) values for individual national forests to use in their long-range planning processes to meet the requirements of the National Forest Management Act. The U.S. Army Corps of Engineers also developed a regional demand model used to transfer recreational benefits from one Corps reservoir to another.

The U.S. Department of Commerce, National Oceanic and Atmospheric Administration (NOAA) recently issued its Final Rule for natural resource damage assessments covered under the U.S. Oil Pollution Act of 1990 (NOAA, 1996). This Final Rule allows for the use of benefit transfer techniques after the following three factors are carefully considered: "the comparability of the users and of the natural resource and/or service being valued in the initial studies and the transfer context; the comparability of the change in quality or quantity of natural resources and/or services in the initial study and in the transfer context (where relevant); and the quality of the studies being transferred" (Federal Register, January 5, 1996, p 499).

Benefit transfer techniques are also widely used in Canada to assess both minor and major policy decisions by Environment Canada. The relatively few primary studies undertaken in Canada require analysts to rely, in part, on U.S. studies for benefit transfer. Formal validity tests have not been undertaken to determine the transferability of values from U.S. sites to Canadian sites.

Recognizing the need for a "non-market valuation library", Environment Canada in collaboration with the U.S. Environmental Protection Agency and leading North American experts are developing a benefit transfer database: the Environmental Valuation Reference Inventory™ (EVRI™). The EVRI™ will provide a useful data base for future benefit transfers.

1.3 Issues in Benefit Transfer

Benefit transfer has been discussed extensively in the environmental evaluation literature and used widely in natural resource management and policy making practice. Yet, there is little research on the validity of benefit transfer or on the circumstances in which it is appropriate. Several issues that need to be addressed include the following.

First, the benefit transfer process and benefit estimation process cannot be separated from each other. A good benefit transfer application is based on high quality valuation studies. There are two general conditions, one theoretical and one practical, for benefit transfer. The theoretical condition is that the underlying behavioural process described by the demand model is the same in the study site as in the policy site. If this condition does not hold (if, for example, people in one context are utility maximizers and people in another context are satisfiers), benefits will not be transferable. However, if this condition does hold, a further practical condition for benefit transfer is that the model be well-specified and that the data used to estimate it are such that the model describes the underlying behavioural process. In benefit transferability tests, the practical condition is a maintained hypothesis, and the theoretical condition is the hypothesis to be tested. This requires that the study to be transferred be well done.

Second, dealing with individual heterogeneity is a significant problem in benefit estimation and transfer. In benefit estimation, unmeasured, household (individual) specific factors may influence household (individual) behaviour. Even with the specification of demographic and social variables, households (individuals) may differ in their responses to prices and environmental attributes. Failure to control for such heterogeneity is likely to yield biased and inconsistent model estimates, and more importantly, biased benefit estimates.

Third, benefit transferability is not satisfactorily described as a dichotomous property. Rather, it is appropriate to consider the degree of transferability of the estimated benefits. This has an important implication for transferability tests. The commonly used difference of means test not only has serious statistical problems but also only provides a Yes or No answer. Thus, measures that describe transferability in continuous terms should be developed.

Fourth, because benefit transfer is an application of a data set and its estimated model that were developed for one specific use to a distinct alternative application, it is inevitable that there may be an error (or a difference) between the "true" and transferred benefit. The questions then are: Is this error (difference) large enough to affect a policymaker's decision? How important is the error (difference) in a policymaker's decision? These are questions of the economic significance of benefit transfer.

1.4 Study Objectives

This study is designed to investigate the above issues by employing non-market random utility models and stated preference data. The general purpose of this study is to answer the following three questions: (1) Are benefit transfers valid and reliable? (2) How can the process of accomplishing benefit transfers be improved? (3) What is the economic significance of benefit transfers? Since benefit transferability and benefit estimation cannot be separated, the first two questions are considered jointly.

The specific objectives of this study are:

1. To develop and estimate the random utility models which can account for observable and unobservable multiple sources of heterogeneity.

2. To calculate/simulate consistent benefit estimates from the developed and estimated random utility models.

3. To evaluate/test benefit transferability using advanced statistical techniques.

4. To investigate the economic significance of benefit transfer in simulated policy settings.

This thesis mainly focuses on the empirical aspects of non-market valuation. It tries to provide several significant contributions to this academic field. First, heterogenous multinomial logit models and random coefficient probit models are developed and used to deal with multiple sources of heterogeneity in individual choice data, in order to obtain consistent benefit estimates. Secondly, the compensating variation of environmental changes are derived or simulated for the heterogenous and random coefficient models. Thirdly, several nonparametric procedures are used to compare two simulated benefit distributions. This provides an appropriate method of statistically testing the reliability of benefit transfers. Fourthly, the economic significance of benefit transfer is investigated by solving a policymaker's decision problem and examining a specific case study.

1.5 The Organization of the Study

This study could be seen as a scientific experiment in which two data sets are collected: one at a study side and the other at a policy site. Assume that the model estimated from policy site data is the "true" model of the policy site, and the model estimated from the study site data is the transferred model. For a given policy or project at the policy site, the benefits are calculated from both the "true" model and transferred model. Benefits calculated from the true model are referred to as the "true" benefits and benefits calculated from the transferred model are referred to as the transferred benefits. The statistical and economic significance of the difference between the "true" and transferred benefits are then evaluated.

The study plan of this thesis is demonstrated in Figure 1. This thesis contains two major parts: consistent benefit estimation with multiple sources of heterogeneity, and the statistical and economic significance of benefit transfers.

The remainder of this thesis is organized as follows. In Chapter 2, the basic random utility model and benefit measurement are presented. In Chapter 3, the experimental design, data collection procedures and survey design are discussed. Chapters 4 and 5 deal with models with multiple sources of heterogeneity. Various random utility models are developed, estimated and compared. The purpose is to obtain consistent benefit estimates. Chapter 4 discusses heterogenous multinomial logit models and benefit calculations. Chapter 5 develops the random coefficient probit models and the benefit simulation approach. Chapter 6 and Chapter 7 examine, respectively, the statistical and economic significance of benefit transfers. The final chapter contains a discussion of the models and transferability tests and discusses some extensions.

Chapter 2. The Random Utility Model of Choice and Benefit Estimation

2.1 Introduction

Since the late 1970s, multiple choice (discrete choice) models have been widely used in transportation, marketing and non-market valuation. Although they have different names in different disciplines, the underlying assumptions and theoretical setting are the same.

In these models, individual decision problems are modelled in terms of multiattribute choice systems that link objective measures of the attributes of alternatives to observed choices. Such systems assume that the process of choice can be described by three fundamental component relations (Meyer and Johnson, 1995, p180): (1) The *valuation rules* that map objective measures of alternative attributes to their perceived attractiveness; (2) The *integration rules* that map perceptions of the attractiveness of a site attributes to overall impressions of the attractiveness of the site; and (3) The *choice or behavioural rules* that map overall impressions to overt behaviours, most commonly choices. Meyer and Johnson (1995, p181) find support for three major generalizations about the form of consumer decision processes: (1) Attribute valuation are a non-linear, reference-point dependent function of objective product attributes; (2) The algebraic integration rule which best describes how valuations are integrated into overall valuations is multiplicative-multilinear; and (3) Overall valuations of an option are linked to choices by a function which recognizes the proximity or similarity of the option to others in the choice set.

In the transportation literature, multiple choice models are usually called discrete choice or random utility models and have been widely applied to analyse the travel mode of the urban commuter (see Ben-Akiva and Lerman 1985, Hensher 1986, and McFadden 1974). In the marketing literature, these models are usually used as Multinominal Logit models to represent choice among alternatives (see Punj and Staelin 1978, Flath and Leonard 1979, Gensch and Recker 1979, Jones and Zufryden 1980, 1981, 1982, Guadagni and Little 1983, Carpenter and Lehamann 1985, Lattin 1985, and Bucklin and Lattin 1986).

In environmental valuation this type of model is usually called a random utility model (RUM) and has been widely used to model recreational demand. Studies by Admowicz (1994), Adamowicz, Louviere and Williams (1994), Adamowicz, Jennings and Coyne (1990), Bockstael, Hanemann and Strand (1984), Coyne and Admowicz (1992), Hellerstien and Mendelsohn (1993), Luckert and Adamowicz (1993), Morey, Rowe and Watson (1993), Stynes and Peterson (1984) and Yen and Admowicz (1994) have illustrated the range of applications and types of random utility models. The prevalence of such studies suggest that the basic structure of RUM is well established.

Compared to conventional demand models, RUMs provide quite a different structure in which to model recreation demand, a structure that focuses attention on the choice among substitute sites for any given recreational trip. Thus it is especially suitable when substitution among quality differentiated sites characterizes the problem. Because site characteristics are instrumental in explaining how individuals allocate their trips across sites, RUMs have been used chiefly to value changes in the specific characteristics of the site, such as catch rates or water quality. Moreover, they are also capable of valuing the losses from eliminating a site, or the benefit of introducing a new site.

2.2 Random Utility of Choice

A major accomplishment of econometric research in recent years has been the development of statistical models suitable for the analysis of discrete choices. This has enabled economists to study behavioural relationships involving purely qualitative variables that are not amenable to conventional regression techniques.

The basic structure of RUM was developed by McFadden (1974). Suppose that an individual n faces a choice set C_J. Define the utility of choice (site) i as

$$U_{in}=V_{in}(y_n-p_{in}\ ,\ Q_i\ ,\ Z_n\ :\ \beta)\ +\ \epsilon_{in} \qquad i\ \in\ C_J \tag{2.1}$$

where V_{in} is the systematic (or explainable) portion of the utility function and ϵ_{in} is an error term associated with joint random variation across both individuals and alternatives. The indirect utility V_{in} is a function of the income available to the individual n, y_n , the price for individual n to access site i, p_{in} , a vector of characteristics of site i, Q_i , and a vector of individual specific variables Z_n . ϵ_{in} can have either an omitted variable or random utility interpretation. ϵ_{in} is a known number for individual n, but a random number for the econometric investigator. β is a set of coefficients to be estimated. With additive errors, individual n chooses site i if $V_{in}\ +\ \epsilon_{in}\ \geq\ V_{jn}\ +\ \epsilon_{jn}$. The utility U_{in} is ordinal utility. Individual consumer selects an alternative which has highest utility in the whole choice set.

Model (2.1) has different names, depending on the assumptions on the distribution of the errors. If the errors are assumed to be independent and identically distributed Gumbel variates, the model becomes the commonly used Multinomial Logit (MNL) model; When there is a general pattern of dependence among alternatives, model (2.1) is referred to as the Generalized Extreme Value (GEV) model; if the errors are assumed to have a mutivariate normal distribution, the model is called a Multinomial Probit (MNP) model. The more restrictive the assumption on the distribution of the errors, the easier to estimate the model becomes. Among these models the MNL model is the easiest to be estimated, and the estimation of MNP model is quite burdensome and thus rarely used.

In the MNL model, the cumulative distribution function (cdf) of an individual error term is

$$F(\epsilon_{in}<\epsilon)=\exp(-e^{-\mu\epsilon}) \tag{2.2}$$

where $\mu>0$ is a scalar. This *cdf* implies that $V(\epsilon_{in}) = \pi^2/(6\mu^2)$. Assume the vector of explanatory variables $X_{in} = (y_n-p_{in}\ ,\ Q_i\ ,\ Z_n)$. The probability that a randomly observed individual n chooses alternative i from choice set C_J is

$$P_{in} = \frac{\exp(\mu\beta X_{in})}{\sum\limits_{j\in C_J} \exp(\mu\beta X_{jn})} \tag{2.3}$$

Because μ in the above equation can not be identified, one must estimate the product $(\mu\beta)$.

The associated likelihood function for a sample of N individuals and J alternative sites is

$$L = \prod_{n=1}^{N} \prod_{i=1}^{J} [\ \exp(V_{in}) \ / \sum_{j=1}^{J} \exp(V_{jn})]^{t^{in}} \tag{2.4}$$

where $t^{in} = 1$ if individual n chooses site i and $t^{in} = 0$ otherwise. Maximizing L produces the maximum likelihood estimates of $\mu\beta$.

An important disadvantage of this MNL model is the well-known independence of irrelevant alternatives (IIA). Under IIA, introduction of a new choice leaves the relative odds of choosing among existing alternatives unaltered. This property requires that all choices be perceived as distinct and independent, i.e., that the errors in estimating the utility associated with each alternative be uncorrelated.

A partial solution to IIA is to use a GEV model. The most commonly used form of GEV model is the Nested Multinomial Logit (NMNL) model. The NMNL models allow alternatives to be grouped in a manner that allows the choice of alternatives to be correlated within, though not between, groups. Suppose individual hunters make the site choice in two stages: whether to hunt or not and where to hunt. Assuming that errors in model (2.1) follow a generalized extreme value distribution, the probability of individual n choosing site i is

$$p_n(i,\ m) = p_n(i/m) * p_n(m) = \frac{e^{V_{im}/\alpha_m}[\Sigma_{j\in D_m} e^{V_{im}/\alpha_m}]^{\alpha_m - 1}}{\Sigma_{k\in m}[\Sigma_{j\in D_m} e^{V_{jk}/\alpha_k}]^{\alpha_k}} \tag{2.5}$$

where D_m are the choices in subsets m (D_1 =not hunting, D_2 = site 1, site 2, ... site J); m are the subsets (m=1, 2); V_{im} is the utility associated with hunting at site j in subset (mode) m. α_m is a parameter that measures the degree of substitution between the various subsets (modes). The coefficient α_m is referred to variously as the "inclusive value coefficient" or the "dissimilarity parameter." When $\alpha_m = 1$ for all m, the probability expression (2.5) collapses to the standard multinomial logit probability, where the IIA property holds between all alternatives. Full information maximum likelihood estimation of the coefficients is accomplished by defining the log likelihood function as the product over a sample of individuals of the probability expressed in equation (2.5):

$$L = \prod_{n=1}^{N} P_{nim} \tag{2.6}$$

One appealing alternative to MNL and NMNL model is the multinomial probit model, in which the residuals in the random utility model have a multivariate normal distribution. Consider the case of three alternatives: $U_{1n} = V_{1n} + \epsilon_{1n}$; $U_{2n} = V_{2n} + \epsilon_{2n}$; and $U_{3n} = V_{3n} + \epsilon_{3n}$. Assume that the residuals (ϵ_{1n} ϵ_{2n} ϵ_{3n}) have a trivariate normal distribution with mean vector zero and covariance matrix Σ given by

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix} \tag{2.7}$$

Consider the probability that the first alternative will be chosen. This is

$$P(\,U_{1n} > U_{2n}\,,\;\; U_{1n} > U_{3n}\,) = P(\,\epsilon_{2n} - \epsilon_{1n} < V_{1n} - V_{2n}\,,\, \epsilon_{3n} - \epsilon_{1n} < V_{1n} - V_{3n}\,) \quad .$$

Write $\eta_{21} = \epsilon_{2n} - \epsilon_{1n}$, $\eta_{31} = \epsilon_{3n} - \epsilon_{1n}$, and $V_{13} = V_1 - V_3$, then η_{21} and η_{31} have a bivariate normal distribution with covariance matrix

$$\Omega_1 = \begin{bmatrix} \sigma_1^2+\sigma_2^2-2\sigma_{12} & \sigma_1^2-\sigma_{13}-\sigma_{12}+\sigma_{23} \\ \sigma_1^2-\sigma_{13}-\sigma_{12}+\sigma_{23} & \sigma_1^2+\sigma_3^2-\sigma_{13} \end{bmatrix} \tag{2.8}$$

Thus the probability that alternative *1* will be chosen is given by

$$P(1) = \int_{-\infty}^{v_{12}} \int_{-\infty}^{v_{13}} f(\eta_{21},\eta_{31})\,d\eta_{21}\,d\eta_{31} \tag{2.9}$$

where $f\,(\,\eta_{21}\,,\,\eta_{31}\,)$ has a bivariate normal distribution with covariance matrix Ω_1 and mean vector zero. The probabilities P(2) and P(3) can be similarly calculated.

With *J* alternatives, the multinomial probit model ends up with *J-1* integrals. For *J>5* the estimation could be very costly and almost impractical. Two alternative methods have been suggested for approximating MNP choice probabilities at moderate cost. The most popular method is the approximation based on the formulas developed by Clark (1961), and applied by Daganzo, et al (1977) to MNP estimation. This method approximates the distribution of the maximum of the normal variates with that of a normal variate. It is good for nonnegatively correlated variates of equal variances, but is poor for negative correlations or unequal variances.

The second method is a Monte Carlo method used by Lerman and Manski (1982) and McFadden (1989). It starts with given values of V_i and draws vectors $(\epsilon_1, \epsilon_2, \ldots, \epsilon_J)$ from a multivariate normal distribution, and the frequency with which utility is maximized at alternative i is recorded. What this procedure amounts to is a search procedure on both the V_i and the covariance matrix of the errors Σ. Even this can be computationally cumbersome.

More recently Geweke (1992), Hajivassiliou (1993), Hajivassiliou and Ruud (1993) and Keane (1994) (GHK) have made great advances in designing simulators, and applying them in MNP estimation. The GHK method simulates the choice probabilities by the smooth recursive sampling methods. Based on the root mean squared error criterion, Hajivassiliou, McFadden, and Ruud (1992) show that the GHK simulator is unambiguously the most reliable method for simulating normal probabilities, compared to twelve other simulators considered. Borsh-Supan and Hajivassiliou (1993) also compare the GHK simulator with the frequency simulator and Stern's simulator, and show that the GHK simulator generates substantially smaller variance than the others.

While various approximation and Monte Carlo methods provide practicable tools for MNP estimation, the best way to estimate MNP is to do the multiple integrations if the computation is tractable. It is found that even GHK is not an unbiased simulator of the log likelihood contribution $ln(P_j)$ and under certain conditions the simulated maximum likelihood for an MNP model is inconsistent (Hajivassiliou and Ruud, 1993).

Among the above models, MNL and NMNL have been commonly used in marketing, transportation and environmental valuation. The MNP model is rarely used due to its computational burden. All of these models assume homogenous preferences by individuals (observations). In Chapters 4 and 5 random utility models with multiple sources of heterogeneity will be developed and applied, based on the models described above.

2.3 Benefit Estimation in Random Utility Models

For environmental valuation, the final product is the estimate of benefit. Benefit estimation in random utility models parallels that of compensating variation (CV) in continuous demand functions. Once the indirect utility function V_i is estimated, the compensating variation of a change in any explanatory variable(s) can be obtained following Small and Rosen (1981) and Hanemann (1982).

First, assume that an individual's utility function is in deterministic form, i.e., it is known without error by the investigator. With discrete choice, an individual's utility maximization problem can be described by the following:

$$
\begin{aligned}
&\textit{Maximize} \quad U = U(x_1, \ldots, x_T, q_1, \ldots, q_T, z) \\
&ST: \\
&\qquad p_j x_j + z = y \\
&\qquad x_i x_j = 0 \quad all\ i \neq j \\
&\qquad x_j = (1, 0) \quad j = 1, \ldots, T
\end{aligned}
\tag{2.10}
$$

This individual consumer has a twice-differentiable, quasi-concave, increasing utility function U defined over the commodities $x_1, x_2, \ldots, x_T$, and z, where z is taken as the numeraire. In addition, the consumer's utility depends on some other variables, $q_1, q_2, \ldots, q_T$, which she/he takes as exogenous; these are, for example, quality attributes of the non-numeraire goods. The consumer chooses (x, z) so as to maximize U.

Once one of the x_j $(j=1, \ldots, T)$ has been selected the quantity of z is fixed because of the budget constraint. Constraint 1 is a budget constraint; constraint 2 describes the discreteness of the consumer choices, which means the x_j $(j=1, \ldots, T)$ are mutually exclusive in consumption; constraint 3 means x_j $(j=1, \ldots, T)$ can be purchased in one unit or not be consumed at all. The above model represents a purely qualitative utility maximizing choice.

It is convenient, but not essential, to make the additional assumption about the utility function in (2.10) that

$$x_j = 0 \rightarrow \partial u / \partial q_j = 0 \qquad j = 1, \ldots, T \tag{2.11}$$

i.e., the attributes of goods do not matter to consumers unless that good is actually consumed.

Suppose that the consumer has chosen good j. Her/his utility conditional on this decision is denoted by U_j, given by

$$U_j = U(0,\ldots,0,\ 1\ ,0,\ldots,0,\ q,\ y-p_j) = V_j(q_j\ ,\ y-p_j) \tag{2.12}$$

where p_j is the price of good j and y is the consumer's income. The consumer's demand function for good j can be written as

$$\begin{aligned} X_j(p,\ q,\ y) \qquad &= 1 \ \text{ if } \ V_j(q_j\ ,\ y-p_j) \geq V_i(q_i\ ,\ y-p_i) \text{ for all } i \neq j \\ &= 0 \qquad\qquad \text{otherwise} \end{aligned} \tag{2.13}$$

Substitution of these demand functions into the utility function yields the unconditional indirect utility function

$$V(p, q, y) = \max[V_1(q_1, y-p_1), \ldots, V_T(q_T, y-p_T)] \tag{2.14}$$

This function measures the utility achieved by the maximizing consumer when confronted with given price, attributes and income. Accordingly, it can be employed to construct monetary measures of the welfare effects of a change in these variables(Hanemann, 1982). By analogy to standard welfare theory, the compensating variation for a change from (p, q^0) to (p, q^1) is defined by

$$V(p, q^1, y+cv) = V(p, q^0, y) \tag{2.15}$$

where *CV* represents compensating variation of quality changes from q^0 to q^1. This equation provides a base for welfare calculation under deterministic discrete choice.

As has been shown in the discussion of the discrete choice model, the random utility model contains some components that are unobservable to the econometric investigator and are treated as random variables. The utility function in this case can be written as

$$U(x, q, z, \epsilon) = u(x, q, z) + \xi_j \epsilon_j \tag{2.16}$$

where ξ_j is 1 if x_j>0, and 0 otherwise. Substituting (2.16) into (2.10) and then solving yields a set of ordinary demand functions and an indirect utility function; these parallel those developed in the deterministic case, except that they involve a random component from the point of view of the econometric investigator.

Suppose that the consumer has selected good j. Conditional on this decision her/his utility is

$$U_j = V_j(q_j, y-p_j) + \epsilon_j \quad j=1, \ldots, T \tag{2.17}$$

The ordinary demand function is

$$\begin{aligned} X_j(p, q, y, \epsilon) &= 1 \quad \textit{if} \quad V_j(q_j , y-p_j) + \epsilon_j \geq V_i(q_i , y-p_i) + \epsilon_i \quad \textit{for all } i \neq j \\ &= 0 \qquad \textit{otherwise} \end{aligned} \tag{2.18}$$

Substituting the ordinary demand functions into the utility function (2.16) yields the unconditional indirect utility function

$$V(p, q, y, \epsilon) = \max[V_1(q_1, y-p_1) + \epsilon_1, \ldots, V_T(q_T, y-p_T) + \epsilon_T] \tag{2.19}$$

Again, $V(p, q, y, \epsilon)$ is the utility attained by the maximizing consumer when confronted with (q, p, y). This is a known number for the consumer, but for the econometric investigator it is a random number. It is natural for the investigator to focus on the mean of the distribution of this random variable. Consequently, the unconditional indirect utility function is evaluated as

$$\begin{aligned} V(p, q, y) &= E[V(p, q, y, \epsilon)] \\ &= E(\max[V_1(q_1, y-p_1) + \epsilon_1, \ldots, V_T(q_T, y-p_T) + \epsilon_T]) \end{aligned} \tag{2.20}$$

Equation (2.20) can be used to calculate welfare effects of changes in the choice set as shown in (2.15). As will be shown, under some assumptions on the distribution of the errors and on the functional form of the indirect utility, the compensating variation (CV) has a closed form solution. However, for some models such as the MNP, the CV calculation could be very complicated and some simulation and numerical techniques are needed to solve for CV.

Equation (2.20) provides the theoretical structure of benefit calculation within RUMs. However, the calculation could be much more complicated in an empirical study, especially in the random utility models with heterogeneity incorporated. Chapter 4 and 5 will develop the empirical benefit estimates in different random utility models.

Chapter 3. The Stated Preference Approach and Survey Data

3.1 The Stated Preference Approach

This study employs two stated preference recreational hunting data sets: one from a study site and the other from a policy site. These two data sets form the base for our experiment and allow us to realize the objectives.

The stated preference approach is a kind of experimental analysis of choice in which an individual is asked to indicate her/his preference or to choose one alternative among a set of hypothetical combinations of attributes that define services and products. Compared with commonly used revealed preference (RP) methods, SP approaches have certain advantages. First, the SP approach uses generated alternatives that can elicit preferences for new (non-existing) alternatives, while RP analysis uses actual alternatives in which responses to non-existing alternatives are not observable. Second, in SP analysis, there is no measurement error of the attributes; multicollinearity can be avoided by experimental design; and the range of attribute levels can be extended. Third, the SP approach prespecifies the choice set, while the choice set in RP analysis is ambiguous in many cases. Fourth, in terms of the number of responses in SP analysis, repetitive questioning is easily implemented, while, in RP approaches, it is difficult to obtain multiple responses from an individual. Fifth, various response formats (e.g., choose one, ranking, rating, matching) can be used in SP analysis, while in the RP approach, the preference information available is "choice". However, SP approach has its disadvantages. A commonly voiced criticism of SP data is that because it is not based on real market behaviour, it may not reflect the current distribution of choices. Models based on SP data may not be appropriate for predicting behaviour if the amount of residual unexplained variation in the SP choice data differs from the amount of residual variation in actual (RP) choices. Recognizing the shortcoming of both SP and RP data, a rescaling approach have been applied to combine SP and RP choice data in order to improve the estimates (see, for example, Adamowicz, et al., 1997 and 1994).

For benefit transfer, the potential advantages of SP methods over RP methods are that SP analysis can avoid the problems associated with the selection of choice sets, and the problem of the differential technical attributes between the transfer site and the policy site. If an activity can be broken down into its attribute components, and if models can be appropriately "segmented" to account for different types of users, the stated preference approach may provide a broad enough response surface to allow for accurate benefit transfer calculations (Adamowicz, Boxall, Louviere and Williams. 1994:19).

Stated preference (SP) analysis has a long history in the marketing and transport literature. It is generally well accepted as a method for eliciting consumer responses to multi-attribute stimuli. While the use of SP techniques in environmental valuation is relatively recent, there have been a few noteworthy examples. Lareau and Rae (1988) studied the value of odour reductions using a type of SP model. They asked individuals to rank alternative combinations of odour contact numbers and increased household costs. Rae (1983) employed SP type techniques in the analysis of benefits from air quality improvements. Mackenzie

(1993) has employed SP type techniques to examine tradeoffs between attributes of recreational hunting experiences. Mackenzie compares a variety of SP methods and illustrates how many of these techniques can be designed to correspond with the Random Utility Model. Opaluch et al. (1993) employed pairwise choices in an SP framework to analyse hazardous waste siting decisions. Viscusi et al. (1991) employed SP type techniques in analysing health risk trade offs. Goodman (1989) examines housing attributes in a SP framework. Adamowicz et al., (1994) employed a choice experiment design to value the impact of a water resource development. This model was constructed to examine recreational site choice.

While the SP approach has several advantages in discrete choice modelling and has been widely used in practice, the process is fairly complicated and certain procedures must be followed. There is a substantial literature on designing SP experiments (Louviere, 1988; Hensher, 1994). Following Hensher (1994), the steps in an experiment can be summarized as follows.

Step 1, Identification of the set of attributes that need to be considered as sources of influence on a consumer's choice.

This step requires a decision on which attributes need to be included in the experimental design and which are to be excluded. This step is quite crucial for SP analysis. If too many attributes are included in the design, the experiment will be too large to handle, and, more importantly, the number of questions needed to be answered by each respondent will be quite large. This may result in a low return rate and accuracy. On the other hand, if some important attributes are missed, there will be no useful information provided by the study. Two things are recommended in this step: First, identifying and including the most important attributes of the commodities and the attributes closely related to policy; second, partitioning the other attributes into generic groups. In practice, the set of attributes is usually identified by focus group discussions combined with the researcher's experience.

Step 2, Selection of the measurement unit for each attribute.

In most cases the measurement unit for an attribute is unambiguous. However for some qualitative or generic attributes such as forest activity on site, the measurement units are not well-specified. The common practice is to define an ordinal scale, e.g., high, medium and low. In this case, the analyst should describe precisely what each level represents.

Step 3, Specification of the number and magnitudes of attribute levels.

For existing attributes, attribute levels should be chosen within the range of current experience. When new alternatives are being evaluated, making the attribute levels believable (and deliverable) becomes a primary consideration. The number of levels for each attribute is decided by the overall complexity of the design. This involves consideration of the combinations of attribute levels generated, the manner in which they are exposed to a respondent, the need to investigate non-linearity, and the extent to which interaction effects between pairs of attributes may be important.

Step 4, Statistical design.

Statistical design involves the compilation of the attribute levels into an experiment. This requires the design of both the "product"descriptions and the choice sets into which these descriptions are placed to satisfy the statistical assumption and properties of various probabilistic discrete choice models. A statistical design could be a full factorial design or a fractional factorial design. A full factorial design contains descriptions of all possible alternatives, enabling one to independently estimate the statistical effects of each attribute on the choice response. In practice the full number of combinations is impractical to evaluate, and so a fractional factorial design is constructed. In designing a fractorial experiment, the analyst has to assume that certain interaction effects among the attributes are not statistically significant. This is a very reasonable maintained hypothesis for a large number of possible interactions. The most common fractional factorial design is a main effects plan. Main effects plans assume that individuals process information in a strictly additive way, such that there are no significant interactions between attributes. A main effects plan does enable the analyst to define linear and high-order dimensions for each attribute. A large number of construction techniques, such as varying choice set double conditional designs, a fixed choice set double conditional design and a fixed choice set design, are now available in practice. They are known to be able to produce designs that satisfy the properties of the models and permit the identification of a wide range of model forms utility specifications (Batsell and Louviere 1991; Louviere 1994; Bunch, Louviere and Anderson 1994).

Step 5, Translation of the experimental design into a set of questions and showcards for execution in the data collection phase.

The experimental design must be translated from a set of orthogonal or near-orthogonal design attribute levels into real information for respondents to comprehend and to which to respond. Where there are a large number of replications, it is popular to block or randomise the experiment in such away that subsets of the respondents are asked to respond to either a fixed subset or random subset in a way that ensures that all replications have equal representation.

Step 6, Model estimation.

As a last step of a SP analysis, RUMs are estimated to represent individual's stated preference. In this step, an appropriate estimation procedure is selected, depending on the metric of the response variable and the level of aggregation of the data for modelling.

3.2 Survey Data

Following the above procedure, two SP recreational moose hunting data sets, one for Alberta and the other for Saskatchewan, were collected by member of the Department of Rural Economy, University of Alberta, and the Canadian Forest Service, respectively. The data collection processes for the two data sets were similar. We use the Alberta SP data to briefly explain this process. The details about the Alberta survey can be found in the project report by Mcleod, Boxall, Adamowicz, Williams and Louviere (1993), and the details about the Saskatchewan survey can be found in the report by MacNab and Mcfarlane (1997).

First, a set of attributes was identified from focus group discussion with hunters and the researchers' knowledge of hunting. In early October 1992, a meeting was held with a group of moose hunters. Most of these individuals were resource management specialists or biologists with high levels of knowledge about moose and forestry. They were also highly experienced moose hunters and all had hunted moose in the study area. In this meeting the researchers, in conjunction with focus group participants, developed a list of hunting attributes for this SP study. This list included the following attributes:

- size and condition of moose populations;
- access within the hunting area both in terms of availability and quality of roads;

- congestion;

- direct presence of forest industry operations.

The distance to the site was chosen to represent travel cost.

Secondly, the levels of each attribute were designed to represent the variations in the real situation of the Wildlife Management Units (WMU) in the moose hunting regions of Alberta.

Combining the information provided by focus group discussion and the real situation in Alberta, attribute levels are constructed as follows.

1. **Evidence of the Size of Moose Populations** - seeing or hearing moose or seeing fresh sign such as tracks, browse or droppings.

A. Less than one moose per day

B. 1 to 2 moose per day

C. 3 moose per day

D. 4 moose per day

2. Access within hunting Area - trail, cutlines or seismic lines.

A. Foot access only

B. ATV or 2-wheel drive vehicles required

C. 4-wheel drive vehicles required

3. Levels of Congestion - encountering (seeing and/or hearing) other hunters during the course of a hunting day.

A. No hunters

B. Other hunters on foot

C. Other hunters on ATVs or other vehicles

4. Quality of Roads

A. Paved surfaces

B. Gravel or dirt, essentially non-paved surfaces

5. Presence of Forest Industry Operations

A. Evidence of logging

B. No evidence of logging

6. Distance from Home to the Hunting Site

A. 50 Km

B. 150 Km

C. 250 Km

D. 350 Km

The attributes and the levels are summarised in Table 1.

Thirdly, an experimental design was constructed based on the attributes and levels described in Table 1. The hunters decision problem was conceptualized as one in which they were offered a choice between pairs of competing Wildlife Management Unit (WMU) descriptions, and given the option of choosing to hunt in one of the described WMUs or to choose not to go moose hunting at all. The design problem involves selecting a sample of WMU profile pairs from the universe of pairs. The Alberta design, for example, was given by a $(2^2\times4^4)\times(2^2\times4^4)\times$(2 versions) factorial, i.e., treating left- and right-hand pairs as a composite set of attributes and levels. As discussed by Louviere and Woodworth (1983), the necessary and sufficient conditions to estimate the parameters of McFadden's (1975) Mother

Table 1. Attributes and Levels Used in the Stated Preference Experiments

Attributes	Levels	Rating
Moose	Evidence of <1 moose per day	1
Population	Evidence of 1-2 moose per day	2
	Evidence of 3-4 moose per day	3
	Evidence of more than 4 moose per day	4
Hunter	Encounters with no other hunters	Cong 1
Congestion	Encounters with others on foot	Cong 2
	Encounters with others on ATV or other vehicles	Cong 3
Hunter	No trails, cutlines or seismic lines	Acc 3
Access	Old trails, passable with ATV or 2 WD vehicle	Acc 2
	Newer trails, passable with 4 WD vehicle	Acc 1
Forestry	Evidence of recent forestry activity	1
Activity	No evidence of recent forestry activity	-1
Road Quality	Mostly paved, some gravel or dirt	1
	Mostly gravel or dirt, some paved	-1
Distance	50 km	
to sites	150 km	
	250 km	
	350 km	

Logit model can be satisfied by selecting the smallest, orthogonal main effects design from this larger factorial to create the WMU profiles and pairs simultaneously. The smallest orthogonal main effects design consists of 32 pairs, which were blocked into two sets of 16 pairs each using a two-level blocking factor added to the design for that purpose.

Fourthly, a questionnaire was constructed. The questionnaire consisted of five parts: i) a trip log outlining all moose hunting trips taken during the 1992 season; ii) a section gathering opinions on hunters' perception of various WMU characteristics such as distance, road quality, access, presence of other hunters, forestry activity, and moose populations; ii) a contingent behaviour question where individuals were asked whether or not they would be willing to travel extra distances to get to a specific WMU if the moose populations in this area were increased; iv) a site choice section where hunters were asked to trade off combinations of attributes within 16 sets of two hypothetical sites; and v) a section collecting information on hunting equipment, preferences and demographic information such as age, income and hunting experience. Sections ii to v of the survey were randomized to allow testing of section ordering bias. Further details of the sampling process and descriptive statistics can be found in McLeod et al. (1993). Since we focus on the random utility models, the most important component of the questionnaire for this thesis is number iv, the choice of hunting site. An example of this part of the survey is displayed in Figure 2.

Fifth, survey interviews were conducted. Samples of Alberta hunters were selected from Alberta Fish and Wildlife records. Some socioeconomic characteristics of the sample can be summarized as follows: the sample mean age was 39.4 years; the oldest individual in the sample was 73 years of age; the hunters had an average of about 20 years of hunting experience and about 16 years of moose hunting experience; about half of those surveyed had completed high school, with 34% reporting some post secondary training; most of the sample reported incomes in the ranges of $20,000-$60,000. The socioeconomic characteristics of the Saskatchewan samples are similar to that of the Alberta sample: over 80% of the Saskatchewan respondents were in the age group of 30 -59; about 70% of the respondents reported incomes in the ranges $20,000-$60,000; about 45% respondents had completed high school, with 36% reporting some post secondary training. The similarity in socioeconomic characteristics may provide a good base for testing benefit transfer.

The Alberta hunters were telephoned and asked to attend a meeting in their town. Of the 422 hunters who were telephoned, 312 confirmed that they would attend the sessions. Of the 312, 271 (87%) actually attended the sessions and 266 completed the survey. There were 8 sessions with group sizes ranging from 20 to 55. The returning rate for the survey was 64.3% (271/422).

A similar survey was conducted in Saskatchewan. Instead of organizing interview sessions for hunters, the Saskatchewan survey was conducted by mail, with 375 individuals in Saskatchewan completing the survey.

Chapter 4. Benefit Transfer Using Heterogenous Multinomial Logit Models

4.1 Introduction

One major difficulty in benefit estimation and transfer is to account for heterogeneity in the choice model. Failure to control for heterogeneity will yield biased and inconsistent parameter estimates, and, more importantly, biased and inconsistent benefit estimates and transfers.

The use of logit models to represent individual choice among alternatives in marketing, transportation and non-market valuation literature has been growing since the late 1970s (e.g., in marketing literature, Guadagni and Little 1983; Louviere and Woodworth 1983; Kamakura and Russel 1989; Gensch 1985; in transportation literature, Ben-Akiva and Lerman 1985, Hensher 1986, and McFadden 1974; in non-market valuation, Admowicz, 1994; Adamowicz, Louviere and Williams. 1994; Adamowicz, Jennings and Coyne, 1990; Bockstael, Hanemann and Strand, 1984; Coyne and Admowicz, 1992). While the application of the multinomial logit model has been widespread, research on proper control for heterogeneity has been limited.

Heterogeneity has a long tradition of importance in consumer choice models, starting with the controversy between Kuehn (1962) and Frank (1962). However, the long tradition of modelling heterogeneity begun by Morrison (1966) has yet to be incorporated into logit models. It has been found that not accounting for heterogeneity when estimating logit models on panel data will lead to biased parameter estimates. Chamberlin (1978, 1980) has shown that severe estimation bias exists in commonly used multinomial models when grouped data are used. Horowitz (1981) has shown in a simulation study that while the estimation bias Chamberlain reports is not too severe in practice for the model parameters, there is severe estimation bias in the estimation of the probabilities of choice.

The possibility of the existence of heterogeneity is quite high in non-market environmental valuation, especially when stated preference data are used. Since only a limited number of attributes (variables) can be included in stated preference experimental design, some alternative-specific and household (individual)-specific variables could be omitted. These unmeasured factors may create some variations in individuals' choice behaviour.

While Adamowicz (1994) developed a rational dynamic model to partially account for heterogeneity in multinomial logit models, only recently, marketing researchers have focused on possible ways of specifying and estimating variations across individuals in multinomial logit models.

In this chapter of the thesis, both a heterogeneous multinomial logit (HMNL) model and a general-purpose estimation procedure are developed. This HMNL model is capable of dealing with multiple sources of heterogeneity in choice data. The calculation of the benefit in HMNL model is also derived.

The remainder of this chapter is organized as follows. In part 2, the heterogeneity in the multinomial logit model is discussed and a HMNL is developed. The benefit calculation formula for the HMNL is derived in part 3, followed by model estimation. Model and benefit transferability are evaluated in Part 5. The final part contains the conclusions.

4.2 Heterogeneity in the Multinomial Logit Model

The original formulation of the multinomial logit model is attributed to Luce (1959). Suppose that individual n faces m choices. Define the utility of alternative i to be $U_{in} = V_{in} + e_{in}$, where V_{in} is the systematic (or explainable) portion of the utility function and e_{in} is an error term. e_{in} is assumed to be independently and identically Gumbel-distributed with a location parameter η, and a scale parameter $\mu>0$.

The probability of individual n choosing alternative 1 is

$$P_n(1) = Pr[V_{1n} + e_{1n} \geq \max_{j=2,\dots,m}(V_{jn} + e_{jn})] \tag{4.1}$$

Define

$$U_n^* = \max_{j=2,\dots,m}(V_{jn} + e_{jn}) \tag{4.2}$$

From the property of Gumbel distribution U_n^* is distributed as

$$U_n^* \sim G(\frac{1}{\mu}\ln\sum_{j=2}^{m} e^{\mu V_{jn}},\ \mu\ ) \tag{4.3}$$

where $\mu = \pi/\sigma_e\sqrt{6}$ is a scale factor. U_n^* can also be written as $U_n = V_n + e_n$, where

$$V_n^* = \frac{1}{\mu}\ln\sum_{j=2}^{m} e^{\mu V_{jn}} \tag{4.4}$$

and $e_n \sim G(0,\ \mu)$.

Since

$$P_n(1) = Pr[V_{1n} + e_{1n} \geq V_n^* + e_n^*)] \\ = Pr[(V_n^* + e_n^*) - (V_{1n} + e_{1n}) \leq 0] \tag{4.5}$$

by the property of Gumbel distribution,

$$P_n(1) = \frac{1}{1+e^{\mu(V_n^*-V_{1n})}} = \frac{e^{\mu V_{1n}}}{e^{\mu V_n^*} + e^{\mu V_{1n}}} \\ = \frac{e^{\mu V_{1n}}}{e^{\mu V_{1n}} + \exp(\ln\sum_{j=2}^{m} e^{\mu V_{jn}})} = \frac{e^{\mu V_{1n}}}{\sum_{j=1}^{m} e^{\mu V_{jn}}} \tag{4.6}$$

If we assume the indirect utility function to be $V_{in} = \beta X_{in}$, Equation (4.6) becomes

$$P_n(1) = \frac{e^{\mu\beta X_{1n}}}{\sum_{j=1}^{m} e^{\mu\beta X_{jn}}} \tag{4.7}$$

Expression (4.7) differs from most published versions of the multinomial logit model by the addition of an imbedded scaler constant μ. The scale factor μ is known to be inversely related to the variance, $\mu = \pi/\sigma_e\sqrt{6}$, but cannot be identified in any particular model because of the confounding with the vector of indirect utility parameters. Ben-Akiva and Lerman (1985) show that as the variance approaches infinity, the scale factor approaches zero, causing the multinomial logit model to predict equal probabilities for all choices. Conversely, as variance goes to zero, μ approaches infinity, causing the multinomial logit model to predict deterministically to the choice with the highest explainable utility.

Since $\sigma_e = \sqrt{6}/\mu\pi = 0.248/\mu$, heterogeneity could be represented by different scale factors. For example, if the source of heterogeneity within a sample is some individual-specific demographic or social-economic variable, s, then a scale function $\mu=\mu(s)$ could be specified to account for the heterogeneity. Substituting the scale function into (4.7) yields

$$P_n(1) = \frac{e^{\mu(s)\beta X_{1n}}}{\sum_{j=1}^{m} e^{\mu(s)\beta X_{jn}}} \tag{4.8}$$

Parameter estimates of both indirect utility function and scale function are obtained by maximizing the following log likelihood function

$$L = \sum_{n} \sum_{\alpha \in c_n} f_{an} \left(\frac{e^{\mu(s)\beta X}}{\sum_{j \in C_n} e^{\mu(s)\beta X}} \right) \tag{4.9}$$

where f_{an} is the observed choice frequency for alternative a and individual n. In this heterogenous multinomial logit model, heterogeneity is incorporated into the likelihood function as a scale function. Under this specification the parameters of the indirect utility function, β, are assumed to be the same within the sample, while the variance is allowed to be different across different groups of the sample.

Recent studies suggest that accounting for the differences in the scale factors is crucial. For example, Adamowicz, Louviere and Williams (1994) found that accounting for scale differences between revealed and stated preference data led to superior model fits, once separate data sources were pooled and rescaled. Swait and Ben-Akiva (1986) and Salomon and Ben-Akiva (1983) suggest improvements in the RUM choice models from appropriate a priori segmentation that accounts for variance heterogeneity and, therefore, scale difference.

For a homogeneous data set, the scale factor μ remains constant and confounded with β, the parameter vector of the indirect utility function. In this case, the scale factor μ cannot be identified from the product $\mu\beta$.

For a heterogeneous data set, however, a set of relative scale factors can be identified. Only relative scale factors are estimable in model (4.9) since $\mu(s)$ cannot be separated from β. The following example explains how a scale factor function is constructed. Assume heterogeneity is associated with an individual's education, and education is categorized into 4 levels. Three dummies, S_1, S_2, and S_3, could be used to represent the last three levels. The scale factor function then can be constructed as $\mu(s) = \exp(\alpha_1 s_1 + \alpha_2 s_2 + \alpha_3 s_3)$,

with the scale factor for the individuals at the first education level being normalized as 1 ($e^0 = 1$). The three relative scale factors for the individuals at the last three education levels can be estimated by substituting $\mu(s)$ into model (4.9), as long as the number of observations in each group is sufficient.

Heterogeneity is also testable in model (4.9). If, for example, α_2 is not significantly different from 0, the individuals with the first and third education level are homogeneous.

One must note that the HMNL model is different from the commonly used MNL model constructed by adding the vector of variables, s, into indirect utility functions. Under HMNL model specification, the consumer's preferences are assumed to be same, but heterogeneity exists within the choice data.

Incorporating heterogeneity into the commonly used multinomial logit model not only is very important for choice model estimation, but also has important implications for benefit estimation in environmental valuation. In the next section, we show that the relative scale factor becomes a part of the benefit calculation formula.

4.3 Benefit Calculation in the Heterogenous MNL model

The formula for benefit calculation in the heterogeneous MNL model can be derived based on Hanemann (1982). Assume the consumer's indirect utility function for visiting site j is $V_j = r(y-p_j) + \beta X_j + \epsilon_j$, where y is the individual's income, p is travel cost and X is a vector of environmental attributes. Using (2.20), the individual's indirect utility function can be written as

$$V(X, y-p) = E(\max[\beta X_1 - rp_1 + \epsilon_1 , \dots , \beta X_M - rp_M + \epsilon_M]) + ry \tag{4.10}$$

If heterogeneity is assumed to exist and expressed as a scale factor function $\mu(s)$, then $\epsilon_i \sim G[\eta_i , \mu(s)]$. By the property of Gumbel distribution, the following distribution can be derived.

$$\begin{aligned} &\max[\beta X_1 - rp_1 + \epsilon_1 , \dots , \beta X_M - rp_M + \epsilon_M] \\ &\sim G[\frac{1}{\mu(s)} \ln \sum_j e^{\mu(s)[\beta X_j - rp_j)]} , \mu(s)] \end{aligned} \tag{4.11}$$

and thus

$$\begin{aligned} &E(\max[\beta X_1 - rp_1 + \epsilon_1 , \dots , \beta X_M - rp_M + \epsilon_M]) \\ &= \frac{1}{\mu(s)} \ln \sum_j e^{\mu(s)(\beta X_j - rp_j)} + \frac{\gamma}{\mu(s)} \end{aligned} \tag{4.12}$$

where $\gamma = 0.5772$ is a constant.

Substituting (4.12) into (4.10) and using the identity

$$V(X^0, y-p\ ) = V(X^1, y-p+CV\ ) \tag{4.13}$$

gives the individual's benefit of the environmental changes from X^0 to X^1 as

$$CV = \frac{1}{\mu(s)\ r}\ [\ \ln\sum_j e^{\mu(s)\ (\beta X_j^0 - rp_j)} - \ln\sum_j e^{\mu(s)\ (\beta X_j^1 - rp_j)}\ ] \tag{4.14}$$

This formula is different from the one commonly used in the literature; as the benefit measure now becomes a function of some prespecified social and/or demographic variables. One must note that since, as shown in (4.10), an expected value is used to calculate an individual's indirect utility function, scale factors (variances) can be expected to enter into the formula of benefit calculation.

An implication of (4.14) for benefit transferability evaluation is that not only the transferability of the consumer's preference but also the variations of the preferences are tested. It is quite possible that the underlying choice processes are the same between the policy site and study site, but one is noisier than the other. It is very important for researchers to identify the sources of heterogeneity within choice data and use heterogeneous models to account for it.

Now we apply the HMNL and equation (4.14) to the stated preference data, and evaluate benefit transferability between Alberta and Saskatchewan.

4.4 Model Estimation

Based on the ρ^2 [1] and the parameter significance of the two provincial models, hunters' indirect utility functions for both provinces are specified as

$$V_j = \alpha_0 + \beta X_j \qquad j = 0,\ 1,\ 2 \tag{4.15}$$

where j = 0 is the nonhunting alternative and j = 1, 2 are the two hunting alternatives (see Figure2), the vector $X_j = (\, P_j\,,\ LnM_j\,,\ ACC_{j1}\,,\ ACC_{j2}\,,\ CON_{j1}\,,\ CON_{j2}\,,\ FOREST_j\,)$ [2] is a vector of explanatory variables. A detailed definition of each variable follows: P_j = travel cost of alternative j, calculated by travel distance (km) multiplied by travel cost per km; LnM_j = log of moose population (evidence of the number of moose per day); ACC_j = hunter access level of alternative j, with ACC_{j1} = newer trails, passable with 4 WD vehicle and ACC_{j2} = newer trails, passable with ATV and 2 WD vehicle; CON_j = hunter congestion of alternative j, with CON_{j1} = no encounters with other hunters and CON_{j2} = encounters with other hunters on foot; and $FOREST_j$ = forestry activity at site j. Another attribute, road quality, is excluded from the model because it is not significant in either provincial models. Attributes

1. ρ^2 is one of the goodness-of-fit measure for discrete choice models. It is defined as $1 - \mathcal{L}(\hat{\beta})/\mathcal{L}(0)$. See Ben-Akiva and Lerman (1987) for details.

2. The variable y-P_j becomes P_j in the model estimation because y is given for a specific individual and thus will not affect the coefficient estimation.

ACC, *CON* and *FOREST* are effects coded[3] rather than dummy coded because dummy coding incorporates the base category into the intercept while effects coding avoids this by making the parameter value for the base equal to the negative sum of the parameter values for the categories. Admowicz, Louviere and Williams (1994) discussed the rationale for using effects codes rather than dummy codes in discrete choice models. A nonhunting alternative is included as a choice and modelled as an alternative specific constant α_0 (plus zero attribute levels for the other variables).

In each provincial sample, several demographic, social and economic variables such as education, residence and income were collected. After estimating several HMNL models, heterogeneity is found to be associated with residence. A scale factor function is then specified as $\mu = e^{\alpha D}$, where D is a dummy variable with D = 0 for rural hunters and D=1 for urban hunters. In this specification, the scale factor for rural hunters is normalized as $\mu_r = e^{\alpha D} = 1$, and the relative scale factor for urban hunters is $\mu_u = e^{\alpha}$. The log likelihood function for each province is then specified as

$$L_i = \sum_{s_i} \sum_{a \in c_n} f_{an} \ln\left(\frac{e^{e^{\alpha D}(\alpha_0 + \beta X)}}{\sum_{j \in C_n} e^{e^{\alpha D}(\alpha_0 + \beta X)}}\right) \tag{4.16}$$

where *i=1* for the Alberta sample and *i=2* for the Saskatchewan sample.

---

3. Effects codes are an alternative to dummy variables for qualitative attributes. If an attribute has 4 levels, the first three levels are coded as dummy variables (3 columns in the design matrix) and 4th is coded -1 for each column. The result is that the coefficient on the 4th is the negative sum of the coefficients on the 3 other levels. The coefficients can be interpreted directly as the impact of that level of the attribute on utility.

If homogeneity is assumed between the data sets of rural and urban hunters, the dummy variable D is excluded from (4.16), and the model becomes the traditional multinomial logit (TMNL) model. To compare with the HMNL models, two provincial TMNL models without heterogeneity specification are also estimated.

GAUSS Maximum Likelihood 4.0 is employed for model estimation. The secant algorithm is set to be of the BFGS (Broyden, Fletcher, Goldfarb, and Shanno) method and the convergence tolerance for the gradient of estimated coefficients is set to be 0.00001. The HMNL model estimation results are presented in Table 2 and the TMNL model estimation results are in Table 3.

The results suggest that the HMNL specification improves the model's goodness of fit. In the Alberta model the HMNL specification increases the $\bar{\rho}^2$ from 0.258 to 0.265, and in Saskatchewan model it increases the $\bar{\rho}^2$ from 0.217 to 0.244. Moreover, both scale factors are statistically significant at a 95% level, which supports the existence of heterogeneity between urban and rural hunters within the two provinces.

As shown in Table 2, almost all parameters in the two HMNL models are statistically significant at a 95% level. As expected, the coefficient on travel cost is negative and significant, and the coefficient on moose population is positive and significant. The congestion level is an important factor in a hunter's decision, since the congestion level with no other hunters is positive and significant. Hunters also prefer easy access to sites since both ACC_1 and ACC_2 are positive and significant. A somewhat surprising result is that forestry activity is a positive factor in hunter's site choice. This may be explained by the fact that forestry activities often improve habitat for moose.

Table 2. Maximum Likelihood Estimates of the HMNL[2] Models

Variables[1]	Coefficient: Saskatchewan Model	Coefficient: Alberta Model	t-value for $\beta_a - \beta_s$
Cost	-0.0199 (12.88)[3]	-0.0253 (22.13)	2.81
Cong 1	0.4283 (15.36)	0.5922 (14.49)	3.31
Cong 2	-0.0217 (1.39)	-0.1421 (3.76)	2.94
Acc 1	0.1862 (5.90)	0.0802 (2.14)	2.16
Acc 2	0.0740 (2.26)	0.1967 (5.24)	2.46
Forestry	0.1983 (8.91)	0.0695 (2.48)	3.59
Moose	1.5313 (16.95)	1.5028 (25.41)	0.26
Scale ($\mu=e^{\alpha}$)	0.946[4] (2.18)[5]	0.844 (1.95)	0.16
Log likelihood	-5471.4	-3461.9	
$\bar{\rho}^2$	0.244	0.265	
Choice Occasions	6000	4256	

1. Attributes *ACC*, *CON* and *Forestry* are effects coded. *ACC 1* = newer trails, passable with 4 WD vehicle and *ACC 2* = newer trails, passable with ATV and 2 WD vehicle; *CON 1* = encounters with no other hunter and *CON2* = encounters with other hunters on foot; *Forestry* = evidence of forestry activity within 5 to 10 years.
2. HMNL model refers to heterogeneous multinomial logit model.
3. Numbers in parentheses are asymptotic *t*-statistics.
4. Scale factor μ calculated from e^{α}.
5. Asymptotic *t*-statistics for the coefficient on the dummies.

Table 3. Maximum Likelihood Estimates of the TMNL Models

Variables	Coefficient: Saskatchewan Model	Coefficient: Alberta Model	t-value for $\beta_a - \beta_s$
Cost	-0.0163 (10.91)	-0.0245 (22.23)	4.41
Cong 1	0.4266 (14.36)	0.5922 (14.46)	3.06
Cong 2	0.0063 (1.12)	-0.1379 (3.76)	3.88
Acc 1	0.1588 (5.91)	0.0800 (2.59)	1.92
Acc 2	0.0725 (1.97)	0.1947 (4.865)	2.25
Forestry	0.2022 (8.13)	0.0678 (2.08)	3.27
Moose	1.4238 (16.78)	1.4692 (26.34)	0.44
Log likelihood	-5512.2	-3471.5	
$\bar{\rho}^2$	0.217	0.258	
Choice Occasions	6000	4256	

1. Attributes *ACC*, *CON* and *Forestry* are effects coded. *ACC 1* = newer trails, passable with 4 WD vehicle and *ACC 2* = newer trails, passable with ATV and 2 WD vehicle; *CON 1* = encounters with no other hunter and *CON2* = encounters with other hunters on foot; *Forestry* = evidence of forestry activity within 5 to 10 years.
2. TMNL model refers to traditional multinomial logit model.
3. Numbers in parentheses are asymptotic *t* -statistics.

4.5 Model and Benefit Transferability Evaluations

Before benefit transferability is evaluated, model transferability between the two provinces is tested. The first test of model transferability is a test of the null hypothesis that the individual coefficients of the two models are the same. For two normal distributions, the null hypothesis that the difference in distribution means equals zero is tested using the classic difference formula

$$z = \frac{\beta_a - \beta_s}{[\sigma^2_{\beta_a} + \sigma^2_{\beta_s}]^{1/2}} \qquad z \sim N(0,\ 1) \tag{4.17}$$

where Z is the test statistic, and β_a and β_s are the estimated individual coefficients of the Alberta and Saskatchewan model, respectively. The test results for the two HMNL models are presented in the last column of Table 2. The results show that while the coefficients on *Ln(M)* and the scale factors are not statistically different between the two HMNL models, all other coefficients are statistically different at a 5% level.

The second test is of the null hypothesis that the set of coefficients for the Saskatchewan model is the same as the set of coefficients for the Alberta model. The likelihood ratio test is constructed as

$$LR = -2\ [\ L_a(\beta_s) - L_a(\beta_a)] \tag{4.18}$$

where $L_a(\beta_s)$ is the log likelihood of Alberta model evaluated at β_s, the coefficients of Saskatchewan model; $L_a(\beta_a)$ is the log likelihood of Alberta model. *LR* is χ^2 distributed with degrees of freedom equal to the number of model parameters. For the two HMNL models the calculated LR = 478.9 {-2 (-3701.60 + 3701.06)}. The probability of exceeding this ratio is less than 1%, so the null hypothesis that the sets of coefficients are the same is strongly rejected.

The same tests are also conducted for the two TMNL models. The results are similar to those of HMNL models. In individual coefficient tests, all coefficients of the transferred model except that on *Ln(M)* are statistically different from those of the "true model". The calculated $LR = 351.9$ {-2(-3647.4+3471.5)}, which also suggests that the null hypothesis is strongly rejected.

Both statistical tests reject the transferability between the "true" and the transferred models, based on a 95% statistical significance level. However, it is important to carefully interpret these tests of statistical significance. As McCloskey and Ziliak (1996) point out, statistical significance and economic significance must be distinguished from each other. A difference can be significant for science or policy and yet be insignificant statistically, and similarly, a statistically significant difference may be insignificant for science and policy. In benefit transfer, economic significance could be interpreted as the percentage difference between the transferred benefit and "true" benefit. Thus, an individual coefficient of a transferred model may be statistically different from that of the "true" model, but not economically different. There are two factors that must be considered when economic and statistical significance are discussed. First, it is quite possible that the magnitudes of two statistically different coefficients are very close. For example, a coefficient of 0.65 with a standard error of 0.0001 is statistically different from a coefficient of 0.64 with standard error of 0.0002, but they may not be significantly different economically when benefits are calculated. Second, as shown in Equation (4.14), different coefficients have been given different weights in benefit calculation.

Assume that Alberta is the policy site and Saskatchewan is the study site. We now evaluate benefit transferability across the two provinces, that is, evaluate the null hypothesis H_0: $B(\beta_s, X_a) = B(\beta_a, X_a)$, where $B(\beta_s, X_a)$ is the transferred benefit calculated from the Saskatchewan model and Alberta policy changes, and $B(\beta_a, X_a)$ is the true benefit calculated from the Alberta (the "true") model and Alberta policy changes. In the region of Alberta examined in this study, there are 14 relevant Wildlife Management Units for recreational moose hunting (see Mcleod et al. for details). Suppose that the proposed policy change will eliminate WMU 346. The benefit of this environmental change is calculated. A representative consumer (hunter) is selected as shown in Table 4. Using the estimated models, the perceived attribute levels for the 14 WMUs (see Table 4), and equation (4.14), the benefits of hunting per trip are calculated in Table 5. Confidence intervals for the mean per trip benefits are also included in Table 5. The standard errors are computed using the Krinsky and Robb (1986) draw procedure. For each model we randomly drew 100 parameter vectors from a multivarate normal distribution with means and a variance-covariance matrix estimated in the logit model. Those draws were used to compute 100 per trip values. The standard error from that distribution was used to calculate the confidence intervals. The minimum and maximum values are also reported. The results show that the transferred benefits of both rural and urban hunters are larger than the true benefits. While the transferred benefit for rural hunters is significantly higher than the true benefit (26% higher),

Table 4. Objective Attribute Levels of Alberta WMUs

WMU	Moose Population	Access	Congestion	Forestry Activity	Distance (Km)
337	2	acc1	cong3	1	87.3
338	3	acc1	cong3	-1	72.6
340	2	acc2	cong3	1	96.5
342	1	acc2	cong3	1	101.9
344	1	acc2	cong3	1	161.2
346	3	acc2	cong3	1	93.9
348	4	acc1	cong3	-1	33.8
350	2	acc2	cong3	1	32.2
352	1	acc2	cong3	1	90.7
354	2	acc2	cong3	1	108.8
356	2	acc1	cong3	1	235.1
437	2	acc2	cong2	1	128.6
438	2	acc1	cong3	1	141.8
439	1	acc2	cong3	-1	48.9

Table 5. Benefit Estimates from True and Transferred Models (Dollar/Trip)

Models	Mean	Min.	Max.	95% CI	Deviation from True Benefit* (%)
True Alberta HMNL					
Rural	-5.44	-6.72	-3.67	[-6.59, -3.84]	
Urban	-6.07	-7.44	-4.17	[-7.32, -4.38]	
True Alberta TMNL	-5.57	-6.74	-3.91	[-6.46, -4.53]	
Transferred HMNL					
Rural	-6.84	-8.12	-3.98	[-8.04, -4.13]	25.7%
Urban	-6.93	-8.26	-5.75	[-8.17, -5.96]	14.1%
Transferred TMNL	-8.24	-10.33	-6.92	[-9.89, -6.83]	47.8%

1. The confidence interval is calculated using the Krinsky-Robb procedure.
2. The deviation from true model is calculated as $[cv^c/cv^t - 1]$, where c and t represent, respectively, the transferred and the true benefit.

transferred benefit for urban hunters is closer to the true value (14% higher).

Benefit transferability using the traditional MNL models is also evaluated. The results reveal that using the traditional MNL model specification yields poorer benefit transfer. The transferred benefit is 48% larger than the "true" benefit. The HMNL model improves benefit transferability in our experiment (Table 5).

4.6 Conclusions

Two goals have been accomplished to this point. First, a heterogeneous multinominal logit model and its benefit calculation have been developed. Second, this new model has been applied to evaluate benefit transferability across regions. It has been shown that, since the variance of choice is a part of the benefit formula, properly controlling for heterogeneity in choice data is particularly important when a choice model is used in environmental valuation. It is also suggested, by comparing the HMNL and TMNL models, that the heterogeneous specification improves the model's goodness of fit. The developed HMNL model is capable of accounting for heterogeneity in choice data.

Using the HMNL model, a SP moose hunting experiment is designed to evaluate benefit transferability between Alberta and Saskatchewan. The results show that, even with an assumption of homogeneous preference, hunter's benefits could differ due to the heterogeneity in choice data. Between the two provinces, the benefit difference for rural hunters is quite high (26%), while that for urban hunters is much closer (14% difference). Using the HMNL model to account for heterogeneity has improved benefit transfer in our experiment.

The developed HMNL model is capable of dealing with multiple sources of heterogeneity, once these sources are identified and the data are available. However, the sources of heterogeneity in choice data may not be pre-identified and the data may not be available for the researchers. In this case, more advanced techniques are needed. The next chapter develops such a model.

Chapter 5. Benefit Transfer Using Random Coefficient Multinomial Probit Models

5.1 Introduction

In the previous chapter, a heterogeneous multinomial logit (HMNL) model was developed and applied to benefit transferability evaluation. This model is an advancement of the basic structure of the commonly used multiple choice logit model. It is capable of partially accounting for heterogeneity within choice data. However, this heterogeneous multinomial logit model still has some undesirable properties. These undesirable properties may result in bias in benefit estimation and transferability evaluation.

First, the well known Independence of Irrelevant Alternatives (IIA) still maintained by the heterogeneous multinomial model restricts the pattern of substitution across alternatives and thus makes the model less likely to reflect reality. The odds ratio between any two alternatives, say $P(i|X,Z)/P(j|X,Z)$, takes the form of $\exp[V_i(X_i, Z) - V_j(X_j, Z)]$, which is independent of the attributes or even the existence of any alternative other than i and j. This implies that, if a new alternative were introduced, all the choice probabilities would be reduced in the same proportion. This is surely unreasonable as a general property of a choice model: one would expect $P(j \mid X, Z)$ to be affected by an amount dependent on the degree to which the individual regards the new alternative as a substitute for alternative j. Moreover, the importance of the IIA property is not confined to applications involving the forecasting of demand for new alternatives: for example, it implies a very restrictive pattern of elasticities of $P(j \mid X, Z)$ with respect to the attributes X_i. Differentiation of $\log[P(j \mid X, Z)]$ with respect to $logX_i$ for any $i \neq j$ reveals

$$\frac{\partial \; lnP(j|X \, , \, Z)}{\partial \; lnX_i} = -P(i|X \, , \, Z) \; \frac{\partial \; V_i(X_i \, , \, Z)}{\partial \; lnX_i} \tag{5.1}$$

This elasticity is dependent only on choice i, and not on j. All choice probabilities share a uniform set of cross-elasticities. This imposes a severe restriction on the type of substitution responses that can be modelled successfully.

Second, the heterogeneous multinomial logit model can account for heterogeneity within choice data only if the sources of heterogeneity are observable, that is, in order to apply the HMNL model, the sources of heterogeneity need to be identified first. However, in many cases the sources of heterogeneity in choice data could be unobservable. Many unmeasured, individual-specific factors may influence an individual's choice behaviour. Even with the specification of demographic and social variables, individuals may differ in their responses to travel cost and environmental attributes. If there are multiple unobservable sources of heterogeneity, the previously developed HMNL model will yield biased and inconsistent model parameters and benefit estimates.

A random coefficient multinomial probit (RCMP) model is developed and applied in this chapter, in response to the above two problems. The RCMP model is an appealing alternative to account for the heterogeneities and release the restrictive IIA assumption in the MNL model. It can account for heterogeneity by specifying a random component for each coefficient of the indirect utility function and release the IIA restriction by allowing a general covariance matrix for the errors in the indirect utility function.

This chapter is organized as follows. In part 2 , the RCMP model is developed and the way it accounts for heterogeneity is discussed. Part 3 provides the techniques for RCMP model estimation and identification. The procedure of benefit simulation is presented in Part 4, followed by empirical model estimation and testing. Part 6 includes benefit simulation and transferability evaluation. The final part contains concluding remarks.

5.2 The RCMP model and Heterogeneity in Choice Data

The random utility model of recreational choice is often specified as

$$U_j^n = \gamma\ (y^n - p_j^n) + X_j\beta + \epsilon_j^n \tag{5.2}$$

where p_j^n is the travel cost to site j for individual n ; y^n is the income of individual n; X_j is a vector of environmental quality attributes of alternative j; γ and β are the constant preference parameters; and ϵ_j^n is an error term associated with joint random variation across both individuals and alternatives. The commonly used MNL model assumes that ϵ_j^n follows the type I extreme value distribution, and that γ and β are constant over individuals.

The RCMP model incorporates heterogeneous preferences into (5.2) in two ways. First, it specifies a perception error for each individual. In this specification, consumers are allowed to perceive the attributes of alternatives differently. For example, in a recreational site choice model, the environmental site quality variables X_j in the model specification are often measured as technical numbers or indices by the environmental management agency or researcher. These numbers or indices may vary across sites, but not individuals. However, when individual choice is modelled, perception errors have to be considered. By taking a trip to site j, the perceived site quality Z_j (X_j) can be different from the number (indices) X_j for different individuals. The RCMP model reflects this perception variation by specifying

$$Z_j^n = X_j (1+\eta^n) , \tag{5.3}$$

where η^n allows perception error for individual n. The perceived attributes Z_j^n are the "real" attributes in the choice decision of individual n, but they are not observable by researchers. Substituting the "real" attributes Z_j^n into (5.2) results in

$$\begin{aligned} U_j^n &= \gamma(y^n - p_j^n) + X_j (1 + \eta^n)\beta + \epsilon_j^n \\ &= \gamma(y^n - p_j^n) + X_j \beta + X_j (\eta^n \beta) + \epsilon_j^n \end{aligned} \tag{5.4}$$

The second way to reflect heterogeneous preferences is to specify a varying term for the respondents' tastes for site attributes. To reflect the varying tastes for the site attributes across individuals, one can specify a random parameter model $\beta^n = \beta + \delta^n$, where β is the "average" taste, and δ^n is the individual specific taste variation. The heterogeneous taste or heterogeneous marginal utility of the site attribute can then be accommodated into the indirect utility function (5.2) as

$$\begin{aligned} U_j^n &= \gamma(y^n - p_j^n) + X_j \beta^n + \epsilon_j^n \\ &= \gamma(y^n - p_j^n) + X_j \beta + X_j \delta^n + \epsilon_j^n \end{aligned} \tag{5.5}$$

With $\beta\eta^n = \delta^n$, model (5.4) and (5.5) are identical. The two interpretations, varying perception and varying taste, are econometrically indistinguishable in the RCMP model. Both of them could be seen as a form of heterogeneity within choice data (Chen and Cosslett, 1996).

Since the deviations of each individual's utility function $\beta\eta^n$ or δ^n are unobservable to the econometric investigator, Equation (5.4) and (5.5) are rewritten[1] as:

$$U_j = \gamma(y - p_j) + X_j\beta + e_j \qquad where \quad e_j = X_j\delta + \epsilon_j \tag{5.6}$$

where there are J alternatives and k attributes, X_j is J by k. If the preferences and the errors are distributed multivariate normal, then the following are distributed

[1] Hereafter, the superscript n is omitted to simplify the notation in this part.

multivariate normal, with the indicated mean and covariance matrix.

$$
\begin{aligned}
&\tilde{\beta} \sim MVN(\beta\ ,\ \Sigma_\delta)\ ,\quad \Sigma_\delta \in \Re^{k\times k} \quad where \quad \tilde{\beta} = \beta + \delta \\
&\epsilon \sim MVN(0\ ,\ \Sigma_\epsilon)\ ,\quad \Sigma_\epsilon \in \Re^{J\times J} \\
&e_j \sim MVN(0\ ,\ \Omega)\ ,\quad \Omega = X_j\Sigma_\delta X_j' + \Sigma_\epsilon, \quad \Omega \in \Re^{J\times J} \\
&U_j \sim MVN(X_j\beta\ ,\ \Omega)\ ,\quad \Omega = X_j\Sigma_\delta X_j' + \Sigma_\epsilon, \quad \Omega \in \Re^{J\times J}
\end{aligned}
\tag{5.7}
$$

Equation (5.6) and (5.7) constitute the RCMP model. With a random component for each coefficient of the individual's indirect utility function, every individual could have a different utility function. Thus, the RCMP model is capable of handling multiple sources of heterogeneity. Moreover, since the random components are assigned to the coefficients of the attributes, there is no need for investigators to obtain additional data and identify the sources of heterogeneity in choice data. So, the RCMP model is capable of accounting for unobservable multiple sources of heterogeneity in choice data.

5.3 RCMP Model Estimation and Identification

Although the advantages of the RCMP model have been known for some time, it has rarely been used in practice. This is mainly because the estimation of the RCMP model involves numerical integration of a multiple dimensional multivariate normal density function. This, combined with the necessity of using an iterative technique to maximize the likelihood function, has made the application of the RCMP essentially impractible. However, with the development of computational capacity and various approximation and simulation techniques, estimation of the RCMP has become feasible. It can be expected that, instead of continuing to apply the MNL model, more and more applications of the RCMP model will emerge in the marketing, transportation and environmental literature.

To estimate the parameters in (5.6), the common practice is to assume that δ and ϵ are independently distributed. Each of them is assumed to follow a multivariate normal distribution with $\epsilon \sim N(0, \Sigma_\epsilon)$ and $\delta \sim N(0, \Sigma_\delta)$. When $\Sigma_\delta = \mathrm{diag}(\sigma^2_{\delta 1}, \ldots, \sigma^2_{\delta k})$, the covariance matrix for e is given by $\Sigma_e = \Sigma_\epsilon + \Sigma_{x\delta}$, where

$$\Sigma_\epsilon = \begin{bmatrix} \sigma^2_{\epsilon_{11}} & \cdots & \sigma^2_{\epsilon_{1J}} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \sigma^2_{\epsilon_{J1}} & \cdots & \sigma^2_{\epsilon_{JJ}} \end{bmatrix} \tag{5.8}$$

and

$$\Sigma_{x\delta} = \begin{bmatrix} \Sigma_k \sigma^2_{\delta k} x^2_{1k} & \dots & \Sigma_k \sigma^2_{\delta k} x_{1k} x_{Jk} \\ . & . & . \\ . & . & . \\ . & . & . \\ \Sigma_k \sigma^2_{\delta k} x_{1k} x_{Jk} & \dots & \Sigma_k \sigma^2_{\delta k} x^2_{Jk} \end{bmatrix} \tag{5.9}$$

The correlation across alternatives can be introduced in two ways in this specification. One is due to the varying taste or perception errors for site quality with $\Sigma_k \sigma^2_{\delta k} x_{kj} x_{j'k} \neq 0$ for $j \neq j'$. The other is due to the correlations contained in ϵ with $\sigma^2_{\epsilon j,j'} \neq 0$ for $j \neq j'$.

Now, the probability that individual n selects alternative j is given by the MNP model:

$$P(j|V_U(\beta, \gamma, y-p, X), \Sigma_e(\Sigma_\delta, \Sigma_\epsilon, X)) = Prob([U_{nj} > U_{nm} \;\; for\ all \;\; j \neq m]$$

$$= \int_{u_j=-\infty}^{\infty} \int_{u_1=-\infty}^{u_j} \dots \int_{u_{j-1}=-\infty}^{u_j} \int_{u_{j+1}=-\infty}^{u_j} \dots \int_{u_J=-\infty}^{u_j} \phi(u|V_U, \Sigma_e) du_1 \dots du_J \tag{5.10}$$

where

$$V_U(\beta, \gamma, y-p, X) = X\beta + \gamma(y-p)$$

$$\Sigma_e(\Sigma_\delta, \Sigma_\epsilon, X) = \Sigma_{x\delta} + \Sigma_\epsilon$$

and $\phi(u|V_U, \Sigma_e)$ is the multivariate normal density function with mean V_U and covariance Σ_e . Since Σ_δ and Σ_ϵ are assumed to be positive definite, it is straightforward to show that $\Sigma_e(\Sigma_\delta, \Sigma_\epsilon, X)$ is also positive definite. This is desirable for establishing regularity conditions for the above integration (Daganzo, 1979).

One of the difficulties associated with (5.10) is that it requires the evaluation of a multivariate integral, which does not have a closed form solution. The usual first step is to reduce the dimension of the integral from J to J-1 using the transformation discussed by Bunch (1991). For a multinomial choice problem, it is equivalent to write the choice probabilities in utility levels or in utility differences, i.e.,

$$
\begin{aligned}
&Prob\ (\ U_{nj} > U_{nm}\ ,\ \ for\ all\ j \neq m) \\
&= Prob\ (\ U_{nj} - U_{nm} > 0,\ \ for\ all\ j \neq m)
\end{aligned}
\tag{5.11}
$$

Thus it is possible to reduce the dimension of the integral in Equation (5.10) from J to J-1 by using some prespecified differencing matrix.

To simplify the problem, assume that the model has only three choices for each individual. To compute $P(j|\ V_U,\ \Sigma_e)$, the differencing matrices $\Delta_i = \Re^{2\times3}$ can then be defined as[2]

$$
\Delta_1 = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad \Delta_2 = \begin{bmatrix} 0 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad \Delta_3 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \tag{5.12}
$$

[2]See Daganzo (1979) and Bunch (1991) for more detailed discussions.

Using these differencing matrices, the choice probabilities can be transformed into utility differences. For example, the probability of choosing alternative 1 becomes

$$\begin{aligned} Prob\ (1) &= Prob\ (\ Z = \Delta_1 U < 0) \\ &= Prob\ (\ U_2 - U_1 < 0\ ,\ U_3 - U_1 < 0) \end{aligned} \tag{5.13}$$

The transformation $Z = \Delta_i U$ applied to (5.10) gives

$$P(i|V_U(\beta,\ \gamma,\ y-p,\ X),\Sigma_e(\Sigma_\delta,\ \Sigma_\epsilon,\ X))=Prob([U_j - U_i < 0,\ \ j \neq i] = \Phi(0|V_z,\ \Sigma_Z)$$

$$\begin{aligned} & where \\ V_z &= \Delta_i V_U = \Delta_i[X\beta + \gamma(y-p)] = \Delta_i X\beta + \Delta_i \gamma(y-p) \\ \Sigma_Z &= \Delta_i \Sigma_e \Delta_i^T = \Delta_i (X\Sigma_\delta X^T)\Delta_i^T + \Delta_i \Sigma_\epsilon \Delta_i^T \end{aligned} \tag{5.14}$$

where Z, $V_z \in \Re^{J-1}$, $\Sigma_z \in \Re^{(J-1)\times(J-1)}$, and $\Phi(.)$ is the cumulative distribution function for the (J-1) dimensional multivariate normal distribution. For a choice model with 3 alternatives, $\Phi(.)$ becomes the cumulative distribution of the bivariate normal density. For a choice model with 4 alternatives, $\Phi(.)$ becomes the cumulative distribution of the trivariate normal density.

Using maximum likelihood procedures, for a data set with the number of choices $J<5$, one should be able to actually calculate the multivariate integral and estimate the indirect utility function. However, for a data set with $J>5$, direct computation of the multivariate integral is infeasible. Simulation or approximation methods have to be used.

Another problem in RCMP model estimation is the identification of the covariance matrix Σ_e . Σ_e is composed of two separate parts; the covariance matrix of preferences, Σ_δ , and the covariance matrix of alternative specific errors, Σ_ϵ . Theoretically Σ_δ is fully identifiable with sufficient data observations. However, Σ_ϵ can not be fully identified. There are J(J+1)/2 parameters in Σ_ϵ, but only J(J-1)/2 of them are identifiable (Daganzo, 1979). Bunch (1991) gives an in depth discussion of the identification of Σ_ϵ. He also recommends some matrix specifications for it. In most applications, researchers have specified either Σ_δ or Σ_ϵ, but not both of them. Whether the researcher specifies Σ_δ or Σ_ϵ, the number of parameters in the covariance matrix proliferate extremely quickly with increasing numbers of alternatives or explanatory variables. Thus it is important to consider parameterizations that limit the number of covariance parameters. One such specification was suggested by Hausman and Wise (1978). They assume that the random coefficients are uncorrelated, and that the alternative specific errors are independently distributed, that is, the off-diagonal elements in both Σ_δ and Σ_ϵ are assumed to be 0. This is a reasonable assumption, especially when the number of parameters is large.

5.4 Benefit Estimation in the RCMP Models

Benefit estimation in the RCMP model does not have a closed form. Simulation techniques are required. Applying the indirect utility function (5.6) to the commonly used expected maximum utility formula (2.20) yields

$$V(p,\ X^0,\ y) = E(V(p,\ X^0,\ y,\ e))$$
$$= \gamma y + E(\max[-\gamma p_1 + X_1^0\beta + e_1,\ \ldots,\ -\gamma p_J + X_J^0\beta + e_J])$$
$$V(p,\ X^1,\ y+CV) = E(V(p,\ X^1,\ y+CV,\ e))$$
$$= \gamma(y+CV) + E(\max[-\gamma p_1 + X_1^1\beta + e_1,\ \ldots,\ -\gamma p_J + X_J^1\beta + e_J]) \tag{5.15}$$

where X^0 and X^1 are, respectively, the site attributes before and after environmental changes, and CV is compensating variation. The benefit of environmental changes can then be estimated from $V(p, X^0, y) = V(p, X^1, y + CV)$ as

$$CV = -\frac{1}{\gamma}\,([E(\max[-\gamma p_1 + X_1^1\beta + e_1,\ \ldots,\ -\gamma p_J + X_J^1\beta + e_J])$$
$$- E(\max[-\gamma p_1 + X_1^0 + e_1,\ \ldots,\ -\gamma p_J + X_J^0\beta + e_J])) \tag{5.16}$$

Since the errors e are of a very complex form, a closed form for CV does not exist. A simulation method can be used to solve for CV. As the first step, the expected maximum utility needs to be simulated. Chen and Cosslett (1996) constructed a frequency simulator to estimate the expected maximum utility

$$\bar{U}_m = \frac{1}{R}\sum_{r=1}^{R}\sum_{j=1}^{J}(\gamma p_j + \beta X_j + e_j^r)\ I[\gamma p_j + \beta X_j + e_j^r > \gamma p_l + \beta X_l + e_l^r\ ,\ \forall\ l] \tag{5.17}$$

where R is the number of replications, and I[A] is an indicator, with I[A] = 1, if A is true, and

I[A] = 0 otherwise. Their practical procedure is described as the following 5 steps:

(1) Consider the model in (5.6). Let the covariance matrix of a representative observation in the sample be Σ_e. Compute the Cholesky decomposition $\Sigma_e = LL'$, where L is the lower triangular matrix. Then one can have $e = L\eta$, where η has an independent standard normal distribution $\eta \sim N(0, I)$.

(2) For replication r, draw a vector $\bar{\eta}^r$ from the normal random variables η, and thus $\bar{e}^r = L\bar{\eta}^r$.

(3) Compute the utility $\bar{U}^r = \gamma p + \beta X + \bar{e}_i^r$ for all alternatives. Let $\bar{U}_m^r$ be the maximum utility of all the alternatives for replication r.

(4) Repeat steps 2 and 3 for replications r = 1 , ..., R, and take the average $\bar{U}_m = (1/R)\Sigma \bar{U}_m^r$. Chen and Cosslett showed that this simulated expected maximum utility $\bar{U}_m$ is unbiased.

(5) Substituting $\bar{U}_m$ into (5.16) gives the simulated benefit function:

$$CV(X^1|X^0) = \frac{\bar{U}_m(X^1) - \bar{U}_m(X^0)}{-\gamma} \tag{5.18}$$

5.5 RCMP Model Estimation Results

The RCMP model estimation technique is used to estimate the parameters of an individual's utility function while allowing for unobserved heterogeneity. The individual's utility function is specified as follows:

$$
\begin{aligned}
U_j^n &= \alpha_0^n + Z_j\beta^n + \epsilon_j^n \\
&= \alpha_0 + Z_j\beta + u_j^n \\
&\textit{where} \\
u_j^n &= Z_j\delta^n + \epsilon_j^n \qquad j = 0,\ 1,\ 2 \\
\beta^n &= \beta + \delta^n
\end{aligned}
\tag{5.19}
$$

where j = 0 is the nonhunting alternative; and $Z_j = \{P_j, LnM_j, ACC_{j1}, ACC_{j2}, CON_{j1}, CON_{j2}, FOREST_j\}$ is a vector of explanatory variables. A detailed definition of each variable follows: P_j = travel cost of alternative j; LnM_j = log of moose population (evidence of the number of moose per day); ACC_j = hunter access level of alternative j; CON_j = hunter congestion of alternative j; and $FOREST_j$ = forest activity at site j. Attributes *ACC*, *CON* and *FOREST* are effects coded rather than dummy coded because dummy coding incorporates the base category into the intercept, while effects coding avoids this by making the parameter value for the base equal to the negative sum of the parameter values for the categories.

Using the two data sets discussed in Chapter 3, four models are estimated. For each data set, an independent probit model and a RCMP model are estimated. Since the independent probit model assumes that there is no heterogeneity in preferences, and that the alternative specific errors are *iid* N(0,1), it should behave very closely to the MNL model (which has the property of IIA). The important difference between the RCMP and the MNL models is not the choice between normal versus type I extreme value errors, but independent versus correlated errors.

All four models are estimated by GAUSS Maximum Likelihood 4.0. The secant algorithm is set to be of the BFGS (Broyden, Fletcher, Goldfarb, and Shanno) method and the convergence tolerance for the gradient of estimated coefficients is set to be 0.00001. Since there are only 3 choices in the SP data sets, the probability functions are constructed as a cumulative distribution of the bivariate normal density, using the transformation from (5.12) to (5.14). Then a GAUSS maximum likelihood program is developed to estimate the parameters. Unlike other approximation or simulation techniques, this procedure should produce unbiased and asymptotically efficient and consistent estimators.

The covariance matrix of the errors term for the three alternative independent probit models (with no preference heterogeneity) is

$$\Omega = Z_i \Sigma_\beta Z_i' + \Sigma_\epsilon = Z_i [0] Z_i + \sigma^2 I^3 \tag{5.20}$$

where I^3 is a 3 by 3 identity matrix. After scaling σ^2 to unity, this independent probit model should be very similar to the MNL model with the exception that the MNL model parameter estimates are scaled by the standard deviation of the Gumbel distribution, $\pi/6^{1/2}$.

Tables 6 and 7 present the estimation results of the independent multinomial probit (IMNP) models of Alberta and Saskatchewan samples. The estimates of MNL models for the two samples are also presented in Table 6 and Table 7, in order to compare the IMNP and MNL models. After re-scaling the parameter estimates of the IMNP models by $\pi/6^{1/2}$, it can be seen that the estimated MNL model and the IMNP model are quite similar. Most of the parameter estimates are fairly close to each other.

Table 6. **Maximum Likelihood Estimates of the MNL and IMNP Models, Alberta**

Variables	Coefficient IMNP	Coefficient MNL	Re-scaled IMNP
Constant	-0.7005 (8.13)	-0.8979 (11.51)	-0.8966
Cost	-0.0211 (15.78)	-0.0245 (22.23)	-0.0270
Cong 1	0.4597 (9.66)	0.5922 (14.46)	0.5885
Cong 2	-0.1111 (2.67)	-0.1379 (3.76)	-0.1422
Acc 1	0.0642 (1.72)	0.0800 (2.59)	0.0821
Acc 2	0.1473 (3.19)	0.1947 (4.86)	0.1886
Forestry	0.0522 (1.90)	0.0678 (2.08)	0.0668
Moose	1.1398 (18.39)	1.4692 (26.34)	1.4590
Log likelihood	-1515.9	-3471.5	
Choice Occasions	4256	4256	

1. MNL model refers to traditional multinomial logit model and IMNP refers independent mutinomial probit model.
2. Numbers in parentheses are t -statistics.
3. Parameters of the re-scaled IMNP model are rescaled by $\pi/6^{1/2}$ for comparison with the MNL model.

Table 7. Maximum Likelihood Estimates of the MNL and IMNP Models, Saskatchewan

Variables	Coefficient IMNP	MNL	Re-scaled IMNP
Constant	0.1059 (1.87)	0.1228 (1.89)	0.1355
Cost	-0.0137 (8.03)	-0.0163 (10.91)	-0.0175
Cong 1	0.3724 (12.38)	0.4266 (14.36)	0.4766
Cong 2	0.0051 (2.52)	0.0063 (1.12)	0.0065
Acc 1	0.1143 (6.42)	0.1588 (5.91)	0.1463
Acc 2	0.0571 ((3.52)	0.0725 (1.97)	0.0731
Forestry	0.1490 (2.87)	0.2022 (8.13)	0.1910
Moose	1.0922 (17.30)	1.4238 (16.78)	1.3981
Log likelihood	-2581.3	-5512.2	
Choice Occasions	6000	6000	

1.MNL model refers to traditional multinomial logit model and IMNP refers independent mutinomial probit model.

2. Numbers in parentheses are *t* -statistics.

3.Parameters of the re-scaled IMNP model are rescaled by $\pi/6^{1/2}$ for comparison with the MNL model.

The log likelihood values of IMNP models increase dramatically due to the change of distribution assumption for ϵ. For the Alberta model, the log likelihood value increases from -3471.5 (MNL model) to -1515.9. For the Saskatchewan model, the log likelihood value increases from -5512.2 to -2581.3. These indicate that the normal distribution fits the data sets better than the extreme value distribution. This finding is different from those of previous studies. For example, using Michigan recreational fishing data, Chen and Cosslett (1996) found that the log likelihood value of an IMNP model was smaller than that of a MNL model. Layton (1996) used a stated preference data of public preferences for Superfund hazardous waste site cleanup and found that the two models (MNL and IMNP) produce almost the same log likelihood values.

The multinomial probit model presented in (5.6) is extremely flexible. The most general model that can be estimated would consist of two separate covariance matrices, one for the random coefficients Σ_β, and one for the alternative specific random errors Σ_ϵ (subject to the identification restrictions discussed early). Here we focus on the random coefficient model, both because preferences are likely to be heterogeneous, and because the random coefficients model offers a clear approach to parsimoniously modelling shared unobservables in the choice data. Following Hauseman and Wise (1978), we consider a model that allows each of the k preference parameters to be independently normally distributed, each with their own variance. This specification with the normalization $\sigma_\epsilon = 1$ imposed[3] is

[3] As we discussed under model identification, a more general model with $\sigma_\epsilon \neq 0$ *for* $j \neq j$ can be estimated if we have enough data observations to recover the (J(J-1))/2 number of parameters in the covariance matrix Σ_ϵ.

$$\Omega = Z_i \Sigma_\beta Z_i' + \Sigma_\epsilon$$

where

$$Z = \begin{bmatrix} 0 & p_1 & C11 & C12 & A11 & A12 & F_1 & M_1 \\ 0 & p_2 & C21 & C22 & A21 & A22 & F_2 & M_2 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Sigma_\beta = \begin{bmatrix} \sigma_\alpha^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{c1}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{c2}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{A1}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{A2}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_F^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_M^2 \end{bmatrix} \qquad (5.22)$$

$$\Sigma_\epsilon = I^3$$

where α, p, $c1$, $c2$, $A1$, $A2$, F, M stand for non-hunting alternative constant, travel cost, congestion 1, congestion 2, access 1, access 2, forestry activity and log moose population, respectively. The coefficient of travel cost is not specified as a random variable $\sigma_p = 0$ for two reasons: (1) since the travel cost is already site specific and individual specific, specifying the coefficient as a random number is not likely to be as important as that for site attributes; and (2) specifying the coefficient of the travel cost as a random variable makes the benefit calculation untractable. This specification avoids the proliferation of covariance terms that are associated with the full specification. One should note that although

Table 8. **Maximum Likelihood Estimates of the RCMP Models, Alberta and Saskatchewan**

	Alberta Model		Saskatchewan Model	
Variables	**Coefficient**	**t - Statistics**	**Coefficient**	**t - Statistics**
Constant	-2.5440	2.93	-2.7416	2.07
Cost	-0.0568	5.79	-0.0523	3.29
Cong 1	1.4941	4.02	1.2524	2.54
Cong 2	-0.3326	2.38	-0.4141	3.20
Acc 1	0.2741	1.96	0.2324	1.82
Acc 2	0.4867	3.04	0.4905	2.11
Forestry	0.0567	0.48	0.0660	1.32
Moose	3.9066	4.43	3.6711	3.89
σ_{α}	3.1615	1.61	5.1769	2.50
σ_{c1}	1.4679	0.99	1.9405	1.98
σ_{c2}	1.2376	0.83	1.8856	1.06
σ_{A1}	1.5206	2.32	2.0483	2.36
σ_{A2}	1.2457	0.83	1.4176	1.88
σ_{F}	0.8966	2.44	1.5109	1.83
σ_{M}	3.3330	2.08	4.3211	4.29
Log likelihood	-1497.6		-2379.6	
Choice Occasions	4256		6000	

the β coefficients are independently distributed, the errors are not, and therefore this model is not subject to the IIA assumption.

Table 8 reports the parameter estimates for the two RCMP models: one for Alberta and one for Saskatchewan. Almost all coefficients in the two models are significant at a 0.05 level. Travel cost and moose population are the two most significant factors in hunter's site choice. Hunters in both provinces prefer less congestion and better access. Forestry activity is a positive but insignificant factor in moose hunting. It is noticed that since the IMNP model and RCMP model have different covariance structures, comparison of the parameter estimates between the two models is not as straightforward as the comparison between the IMNP and the MNL models. However, compared to the IMNP models, one can find that both RCMP models have higher log likelihood function values. For the Alberta sample, the log likelihood value increases from -1515.9 (IMNP model) to -1497.6 (RCMP model); and for the Saskatchewan sample, the log likelihood function value increases from -2581.3 (IMNP model) to -2379.6 (RCMP model). The higher log likelihood values of the two RCMP models suggest that the RCMP specification fits the two data sets better than the IMNP specification. These also could be seen as an indication of the existence of heterogeneous preferences within the two choice data sets.

The results in Table 8 also reveal significant variation in hunters' tastes (preferences) over hunting demand. In the Alberta model, σ_{A1}, σ_F , σ_M are significant at a 0.05 level. This indicates that Alberta hunters have significant heterogeneous preferences for access, forestry activity and moose populations. In the Saskatchewan model, all varying components but σ_{c2} are significant at a 0.05 level. This suggest that Saskatchewan hunters have much more variations in their preferences.

The most important and interesting finding comes from the comparison between the two RCMP models. The constant or average parts ($\bar{\beta}X$) of the indirect utility functions are very close in the two provinces, while varying or heterogeneous parts ($\tilde{\beta}X$) are significantly different from each other. Comparing the varying parameters σs, one can find that all the σs in the Saskatchewan model are much higher than those in Alberta model. This finding may have important implications in the evaluation of benefit transferability. It is quite possible that, since the so-called "intrinsic" indirect utility functions are similar in the two provinces, benefits may be more transferable after removing the heterogeneous preference component. In the next part, benefit transferability is evaluated.

5.6 Benefit Transferability Evaluation

Following the procedure presented in Part 5.4, benefits are simulated for different policy scenarios. Again, assume Alberta is the policy site and Saskatchewan is the study site. We now evaluate benefit transferability across the two provinces, that is, evaluate the null hypothesis H_0: $B(\beta_s, \Omega_s(X_a, \beta_s), X_a) = B(\beta_a, \Omega_a(X_a, \beta_a), X_a)$, where $B(\beta_s, \Omega_s(X_a, \beta_s), X_a)$ is the transferred benefit calculated from the Saskatchewan model and Alberta policy changes, and $B(\beta_a, \Omega_a(X_a, \beta_a), X_a)$ is the true benefit calculated from the Alberta (the "true") model and Alberta policy changes.

In the region of Alberta examined in this study, there are 14 relevant Wildlife Management Units for recreational moose hunting (see Mcleod et al. for details). They are WMU337, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 437, 438, and 439. The objective attribute levels of these WMUs, along with the distances to these WMUs for a representative hunter, are listed in Table 4. Using the data in Table 4[4], the "true" and transferred benefits of 8 policy scenarios are simulated. These scenarios are:

Scenario 1: closing WMU337.

Scenario 2: closing WMU338.

Scenario 3: closing WMU340.

Scenario 4: closing WMU344.

Scenario 5: closing WMU348.

Scenario 6: increasing moose population in WMU344 from 1 to 3.

Scenario 7: reducing moose population in WMU348 from 4 to 1.

Scenario 8: reducing the congestion level of WMU348 from *cong3* to *cong1*.

Benefit simulation for these 8 policy scenarios could be grouped into two categories: closing a site or changing a site attribute. For scenario 1 to 5, benefit is expressed as

$$CV_j = \frac{\max(U_1\ ,\ U_2\ ,\ \dots\ ,\ U_{14}\ )-\max(U_1\ ,\ \dots,\ U_{j-1}\ ,\ U_{j+1}\ ,\ \dots\ ,\ U_{14}\ )}{-\gamma} \qquad (5.23)$$

and for scenario 6 to 8, the benefits are calculated as

---

4. Two items should be noted when the data in Table 4 are used to simulate benefits. (1) Instead of using objective levels of attributes, perceived levels of the attributes in the WMUs could be used since the information on the perceptions of attributes of the WMUs was also collected. However, the objective levels are used here for two reasons: first, preliminary analysis (McLeod, et al, 1993) suggests a reasonable degree of correlation between objective and perceived measures; second, the RCMP model has considered perception errors among hunters. (2)Although the selection of hunter's distances to sites will have no impact on Ωs since there is no varying component assigned to travel cost, it does affect the "intrinsic" utility. As a result, the selection of the distances will affect benefit transferability. A randomly selected representative hunter in Alberta is used in this study.

$$CV_j = \frac{\max(U_1 ,\dots , U_j(X^0) , \dots , U_{14})-\max(U_1 , \dots, U_j(X^1) , \dots , U_{14})}{-\gamma} \quad (5.24)$$

For the first category (Scenario 1 to 5), benefits are simulated using the following procedure: (1) calculate the "intrinsic" utility ($\bar{\beta}X$) for all the 14 WMUs; (2) simulate the varying utility $\tilde{\beta}X$ by employing the step 1 and step 2 in Part 5.4; (3) calculate the individual's utility $U_j = \bar{\beta}X_j + \tilde{\beta}X_j$; (4)calculate $U_m^0 = \max(U_j)$ *for* $j=1, \dots, 14$ and $U_m^1 = \max(U_j)$ *for* $j=1,\dots, j-1 , j+1 ,\dots, 14$; and (5) calculate benefits using (5.23).

For the second category , $U_m^0 = \max(U_j (X^0))$ *for* $j=1, \dots, 14$ and $U_m^1 = \max(U_j (X^1))$ *for* $j=1, \dots, 14$.

The simulated benefits for the 8 policy scenarios are reported in Table 9. The results in Table 9 suggest that the benefits of most of the policy scenarios are reasonablely close between the "true" model and the transferred model. For example, for policy scenarios 2, 5, 7 and 8, the deviations of the transferred benefits from the "true" benefits are less than 10%. This number is much smaller than that in the previous chapter where the heterogeneous MNL and the traditional MNL were used. It is suggested that the varying parameter specification improves benefit transferability.

It is found that the simulated benefit distributions have very large variances. The standard errors of the "true" (transferred) benefits for policy scenarios 1 to 8 are, respectively, 2.27 (3.18), 3.14 (3.72), 1.58 (1.86), 0.82 (1.41), 21.35 (24.65), 18.55 (22.50), 18.45 (19.62) and 35.37 (38.34). For all policy scenarios, transferred benefits are more variable than the

Table 9. Benefit Estimates from the True and Transferred RCMP Models

Policy Scenarios	Mean	Min.	Max.	Deviation*
Scenario 1				
True Model	-0.280	-39.76	0	
Transferred Model	-0.527	-45.210	0	85.2%
Scenario 2				
True Model	-0.573	-45.54	0	
Transferred Model	-0.577	-55.01	0	0.7%
Scenario 3				
True Model	-0.137	-36.67	0	
Transferred Model	-0.160	-35.35	0	16.7%
Scenario 4				
True Model	-0.071	-18.25	0	
Transferred Model	-0.141	-22.94	0	98.5%
Scenario 5				
True Model	-13.844	-120.17	0	
Transferred Model	-14.357	-136.13	0	3.7%
Scenario 6				
True Model	3.267	-79.76	99.00	
Transferred Model	4.009	-105.96	118.50	22.7%
Scenario 7				
True Model	-13.126	-80.94	18.30	
Transferred Model	-12.787	-90.39	22.48	-2.5%
Scenario 8				
True Model	33.78	-31.99	145.73	
Transferred Model	29.82	-47.01	167.45	-3.9%

* The deviation is calculated as $[cv^c/cv^t - 1]$, where c and t represent, respectively, the transfer model and the true model.

Table 10. **Simulated Benefits of Eliminating a Site and Probability of Visiting the Site**

Policy Scenarios	Mean Benefit	Probability (%)
Site WMU 337		
True Model	-0.280	2.7
Transferred Model	-0.527	4.4
Site WMU 338		
True Model	-0.573	5.0
Transferred Model	-0.577	4.7
Site WMU 340		
True Model	-0.137	1.6
Transferred Model	-0.160	1.3
Site WMU 344		
True Model	-0.071	1.0
Transferred Model	-0.141	1.6
Site WMU 348		
True Model	-13.844	46.6
Transferred Model	-14.357	38.5

"true" benefits. This reflects the fact that the transferred model has more significant variation than the "true" model.

The simulated impacts of scenario 1 to 5 are actually truncated distributions. The maximum welfare is always zero. This is due to the way CV is simulated. If, in a replication, the site that is going to be eliminated is not chosen, CV will be zero. Thus, the maximum impacts of eliminating a site are zero. It is interesting to explore the relationship between the probability of visiting a site and the cost of removing access to the site. Theoretically, it is expected that the higher the probability, the larger the cost. The results presented in Table 10 reflect this theory. The probability of hunting in WMU 348 is the highest, at 46.6%. The loss of eliminating this site is the largest, $13.84 per trip. It is also found that the simulated probabilities between the "true" and transferred models are very close.

5.7 Conclusion

This chapter has employed an econometric model to account for individual heterogenous preferences of recreational demand. An estimation procedure has been developed for the stated preference data. The empirical results of the random coefficient probit models have revealed significant variation among individual preferences. The methods applied in this chapter are important because they are currently tractable and allow the modelling of flexible correlation structures. It is expected that random coefficient probit model will become important in analysing individual choice models.

By comparing the independent model and the random coefficient model, one can see that the varying parameter specification improves the model's goodness of fit. The specification can be important because in many cases the explanatory variables such as the site attributes, are measured by a set of numbers that do not vary across individuals. The concern is whether the constant parameters of the indirect utility function can adequately explain each individual's taste or perception of site quality. From the estimation results, the random coefficient specification provides a significant improvement over the constant parameter specification.

A frequency simulator is also employed in this chapter to calculate the benefits of different environmental changes. The results show that benefits to the representative consumer have much larger variation after the heterogenous preferences are accounted for. The simulated "true" and transferred benefits are very close in the random coefficient probit models, compared to those in the previous chapter. It is also found that the calculated benefits are consistent with the simulated probabilities, that is, the higher the probability of visiting the site, the larger the loss of eliminating the site.

Chapter 6 Statistical Tests for Benefit Transferability Using Nonparametric Procedures

6.1 Introduction

In Chapter 5, the "true" and transferred benefits of eight policy scenarios are simulated. It is shown that, when the random parameter multinomial probit model is employed to account for heterogeneous preferences, benefits of environmental change could be approximately simulated as a random series. The means, maximums and minimums of these empirical distributions are given in Chapter 5. Benefit transferability can be evaluated by comparing the means of the "true" and transferred benefit distributions. These evaluations are quite preliminary, and more rigorous statistics are required to test for benefit transferability.

To test benefit transferability is to test the similarity of the underlying distributions. Unfortunately, there is little theoretical foundation for defining the underlying distributions of the simulated random series. These random series can only be seen as the samples of some unknown distributions. For this purpose, comparison of the means of two unknown distributions is definitely not a rigorous test, because it is quite possible that two significantly different distributions can have identical means. Nonparametric or distribution-free procedures are required.

This chapter employs nonparametric procedures to test benefit transferability across regions, using the simulated samples of the empirical benefits from Chapter 5. Two nonparametric procedures, the Mann-Whitney test and the method of convolutions approach, are used to evaluate the significance of the difference of the empirical distributions. The question asked here is: Is the difference of the two distributions of point estimates $B(\beta_s, \Omega_s(X_a, \beta_s), X_a]$ - $B(\beta_a, \Omega_a(X_a, \beta_a), X_a]$ significantly different from zero? Here, $B(\beta_s, \Omega_s(X_a, \beta_s), X_a]$ is the transferred benefit distribution simulated from the Saskatchewan model and the Alberta policy changes, and $B(\beta_a, \Omega_a(X_a, \beta_a), X_a]$ is the true benefit distribution simulated from the Alberta ("true") model and Alberta policy changes.

The Mann-Whitney test is a frequently used nonparametric test that is equivalent to another well-known test, the Wilcoxon sum-of-ranks test. The Mann-Whitney test is simple to use for any size samples, and tables of the exact null distribution are widely available. The large-sample approximation is quite adequate for most practical purposes (Breiman, 1973).

The convolutions approach is based on the method of convolutions. This technique is used in mathematics and statistics to calculate the distribution of a sum of random variables and series (Feller, 1957; Mood, Graybill, and Boes, 1974). Based on this technique, Poe, Severance-Lossin and Welsh (1994) developed a statistical test for the difference of simulated distributions. They demonstrated that the classic parametric method that invokes a normality assumption is generally not appropriate for evaluating differences in simulated distributions. They argued that, because the convolution formula provides an exact statistical test of the difference of empirical distributions, it is preferable to the parametric normality based classic tests.

Since each of these two tests has advantages over the other, both of them will be used. Given the widespread use of resampling or simulation techniques in welfare measures, elasticities and flexibilities, economies of size, and non-market values, it can be expected that nonparametric methods will be used more frequently in applied economics.

This chapter is organized into 4 sections. The second section briefly reviews and evaluates some commonly used parametric and nonparametric tests for underlying distributions. Section 3 presents the theoretical foundation of the empirical convolutions approach and uses a simple distribution to demonstrate the test procedure. The final section presents test results.

6.2 Methods of Testing for Underlying Distributions

The problem of deciding whether a number of samples come from the same underlying distribution is a fundamental and frequent problem in statistics. The problem could be stated as follows: Given a number of samples from different populations, decide whether the populations have the same distribution. Precisely, suppose $x_1,..., x_n$ is a sample drawn from one underlying distribution and $y_1,..., y_m$ a sample drawn, independently of the first, from some other underlying distribution, how can these data be used to decide whether the two underlying distributions are the same? In statistics, various methods have been developed to solve this problem. These methods can be divided into two categories: parametric tests and nonparametric tests.

1. Parametric Test

The entire body of parametric techniques is based on fairly specific assumptions regarding the nature of the underlying population distribution; usually its form and some parameter values must be stated. Given a set of assumptions, certain test statistics can be developed. As long as the assumptions themselves can be substantiated, the conclusions reached using these techniques are valid.

The most commonly used parametric test for underlying distributions is the student t test. Assume there is a sample $x_1,\ldots, x_n$ from one distribution P_1 and an independent sample $y_1,\ldots, y_m$ from another distribution P_2. Suppose further that $P_1 \sim N(\mu_1, \sigma^2)$ and $P_2 \sim N(\mu_2, \sigma^2)$. Let $\hat{\mu}_1$, $\hat{\mu}_2$ be the sample means, and examine the difference $\hat{\mu}_1 - \hat{\mu}_2$. Under H_0: $\mu_1 = \mu_2$, $\hat{\mu}_1 - \hat{\mu}_2$ has mean zero. The student t test is constructed as follows:

$$t = \frac{\hat{\mu}_1 - \hat{\mu}_2}{Sd} \tag{6.1}$$

where Sd is the standard deviation of the difference $\hat{\mu}_1 - \hat{\mu}_2$. The degrees of freedom *(df) equals* $n+m-2$ in this case. Comparison of the calculated $|t|$ and the critical $|t^*|$ provides the test result. If $|t| < |t^*|$, H_0 can not be rejected at a given significance level, and if $|t| > |t^*|$, H_0 is rejected. Sd is calculated as

$$Sd = [\, S_p^2\,(1/n + 1/m)]^{1/2}$$

where

$$S_p^2 = \frac{(n-1)\hat{\sigma}_1^2 + (m-1)\hat{\sigma}_2^2}{n+m-2} \tag{6.2}$$

In order to apply the classic t test, one must assume $\sigma_1 = \sigma_2$. If these variances are not equal, the t test can exhibit serious problems (Feller, 1957). Since equality of variances is probably not a safe assumption in simulation or resampling, an old but reliable approximate solution, the Welch confidence interval, is recommended by Law and Kelton (1982), instead of using the classic t test.

Assume two normal distributions with unequal and unknown variances. The Welch confidence interval is constructed as follows:

$$(\hat{\mu}_1 - \hat{\mu}_2) \pm t_{\hat{df},\, 1-\alpha/2} \, (\hat{\sigma}_1^2/n + \hat{\sigma}_2^2/m)^{1/2}$$

$$\text{where} \qquad \hat{df} = \frac{(\hat{\sigma}_1^2/n + \hat{\sigma}_2^2/m)^2}{(\hat{\sigma}_1^2/n)^2/(n-1) + (\hat{\sigma}_2^2/m)^2/(m-1)} \tag{6.3}$$

Since the degrees of freedom will not, in general, be an integer, interpolation in the t tables will probably be necessary. If the confidence interval contains zero, the null hypothesis H_o: $\mu_1 = \mu_2$ can not be rejected.

The above parametric test is surprisingly robust. It maintains its stated level accurately even if the underlying distributions are not exactly normal (Breiman, 1973). The larger the sample size, the more robust it is, and the two-sided test is more robust than the one-sided.

However, as Poe et al. demonstrated, the normal distribution assumption may not be appropriate for simulated distributions in most cases. In this study, objections to the normality assumption occur at both a theoretical and empirical level. First, due to the structure of the multiple choice model, the distributions of consumer's benefits from some environmental changes are truncated. For example, the welfare distributions of eliminating a hunting site in policy scenarios 1 to 5 (Chapter 5) are theoretically truncated at 0. Consumers (hunters) will never get positive welfare from the elimination of a hunting site.

Second, the simulated empirical distributions do not support the normality assumption. Figures 3 to 10 present the probability density functions (pdf) of the simulated true and transferred benefit distributions. From the shapes of these *pdfs*, one should not conclude that the simulated points are samples drawn from normal distributions. To actually test the normality assumption, a Wald statistic is constructed for each empirical distribution. Under the hypothesis of normality, the test statistic would be

$$W = n\left[\frac{b_1}{6} + \frac{(b_2 - 3)^2}{24}\right] \sim \chi^2(2)$$

where

$$b_1 = \textit{skewness coefficient} = \frac{E[(x - \bar{x})^3]}{Var[x])^{3/2}} \qquad (6.4)$$

$$b_2 = \textit{kurtosis coefficient} = \frac{E[(x - \bar{x})^4]}{Var[x])^2}$$

The normality test results are reported in Table 11. All calculated Wald statistics are much larger than the 5 percent critical value 5.99. Thus, the null hypothesis is rejected and none of the empirical distributions could be treated as a sample for a normal distribution.

Table 11. **Wald Statistic[a] of the Normality Test for the Simulated Empirical True and Transferred Benefit Distributions**

Scenario	True Models	Transferred Models
Scenario 1	17388.2	15428.8
Scenario 2	13267.7	17405.6
Scenario 3	5189.9	4319.9
Scenario 4	12504.6	14792.3
Scenario 5	73.8	110.3
Scenario 6	17.6	17.7
Scenario 7	23.5	35.5
Scenario 8	16.7	14.9

[a] The 5 percent critical value from the Chi-squared table for two degrees of freedom is 5.99, so all empirical distributions do depart significantly from normality.

In applied economics, many researchers employ simulation techniques because they capture the inherent nonlinearities of the functions of the parameters used to calculate the desired point estimate. It seems rather counter-intuitive to impose arbitrary assumptions such as normality when one of the primary advantages of the simulation method is to avoid unnecessary parametric restrictions. In practice, it remains an empirical question if the "approximately normal" condition holds for simulated economic variables. For example, while Dorfman, Kling and Sexton (1990), and Krinsky and Robb (1991) found that Taylor's approximations and normality assumptions closely approximate the simulated elasticity confidence intervals, Green, Hahn and Rocke (1987), and Anderson and Thursby (1986) found such equivalence to hold only under restrictive conditions. Non parametric techniques provide an alternative set of tests when the underlying distributions are unknown.

2. Nonparametric Tests

Nonparametric techniques have (certain) desirable properties that hold under relatively mild assumptions regarding the underlying population from which the data are obtained. In particular, nonparametric procedures forego the traditional assumption that the underlying populations are normal. Although at first glance most nonparametric procedures seem to sacrifice too much of the basic information in the samples, theoretical investigations have shown that this is not the case. More often than not, the nonparametric procedures are only slightly less efficient than their normal theory competitors when the underlying populations are normal, and they can be mildly and wildly more efficient than these competitors when the underlying populations are not normal (Hollander and Wolfe, 1973).

There are several nonparametric tests. The often used median test and Mann-Whitney test are discussed below.

The median test generally is not very efficient, but can be computed very easily. Assume $n+m$ is even, and let v be the median of the combined sample $x_1,\ldots, x_n, y_1,\ldots, y_m$. Let N be the number of $x_1,\ldots, x_n$ that are less than v. If the two samples come from the same distribution, then N should be around $n/2$. For a two-sided test, the hypothesis of homogeneity is accepted if N is neither too large nor too small. If one is testing against the one-sided alternative that the distribution underlying the x values has generally larger outcomes, then the hypothesis is accepted if N is not too small -- and analogously for testing against the alternatives that the y values are generally larger. For sufficiently large sample sizes, N has an approximately normal distribution with

$$E(N) = \frac{n}{2}, \qquad V(N) = \frac{1}{4}\left(\frac{nm}{n+m-1}\right) \tag{6.5}$$

This test could be extended to test k samples. The advantage of this median test is that it is very easy to calculate. The drawback of this test is that its efficiency is not up to par, when used in the normal case and compared to the student t test (Breiman, 1973).

The Mann-Whitney (Mann and Whitney, 1947) test is the most commonly used nonparametric test for underlying distributions. It is used to test H_0: two different samples are drawn from the same distributions. The Mann-Whitney test is based on the idea that the particular pattern exhibited when mX random variables and nY random variables are arranged together in increasing order of magnitude provides information about the relationship between their populations. The Mann-Whitney criterion is based on the magnitudes of the Ys in relation to the Xs. A sample pattern of arrangement where most of the Ys are greater than most of the Xs, or vice versa, would be evidence against a random mixing and thus tends to discredit the null hypothesis of identical distributions. It does not require the samples to be drawn from normal distributions and works well when the sample size is small.

The Mann-Whitney test is performed in 5 steps:

(1) Combine two samples into a single group and keep track of which sample each point comes from;

(2) Sort the combined data list from smallest to largest;

(3) Assign each point in the list a number corresponding to its position in the data list;

(4) Add the rank from each sample;

(5) Calculate the test statistics as follows:

$$
\begin{aligned}
&Z = \frac{R_1 - U_{RI}}{\sigma_R} \sim N(0,\ 1) \\
&\textit{where} \\
&R_1 = \textit{total rank of sample } 1 \\
&U_{RI} = \frac{n(n+m+1)}{2} \\
&\sigma_R = [\frac{nm(n+m+1)}{12}]^{1/2}
\end{aligned}
\qquad (6.6)
$$

The calculated Z is compared to the critical value from the standard normal distribution to give the test result.

The above expression of the Mann-Whitney statistic does not allow for the possibility of ties across the sample. If, however, there exist ties among samples, the modified definition of the Mann-Whitney statistic is

$$Z = \frac{U_T - E(U_T|H_0)}{[Var(U_T|H_0)]^{1/2}} \sim N(0,\ 1)$$

where

$$U_T = \sum_{i=1}^{n} \sum_{j=1}^{m} D_{ij}$$

$$D_{ij} = \begin{cases} 1 & \text{if } x_i > y_j \\ 0 & \text{if } x_i = y_j \\ -1 & \text{if } x_i < y_j \end{cases} \tag{6.7}$$

$$E(U_T|H_0) = 0$$

$$Var(U_T|H_0) = \frac{mn(N+1)}{12}[1 - \frac{\sum_{i=1}^{e}(t_i^3 - t_i)}{N(N^2 - 1)}]$$

Here t denotes the multiplicity of a tie and the sum is extended over all sets of t ties. Lehmenn (1975) shows that the normal approximation is supported by a limit theory, which states that the null distribution of Z tends to the standard normal distribution provided both m and n tend to infinity and $max(d_1/N,\ d_2/N,\ ...,\ d_e/N)$ is bounded away from 1 as $N \to \infty$.

The Mann-Whitney test is a frequently used nonparametric test that is equivalent to another well-known test, the Wilcoxon sum-of-ranks test. The test is simple to use for any size samples, and tables of the exact null distribution are widely available. The large-sample approximation is quite adequate for most practical purposes, and corrections for ties can be incorporated in the test statistic. The test has been found to perform particularly well as a test for equal means (or medians), since it is especially sensitive to differences in location. For our purpose, this test is quite appropriate for the tests of the benefit distributions of scenarios 6 to 8. However, due to the precondition $\lim_{N \to \infty} Max[d_1/N, d_2/N, ..., d_j/N] \gg 1$, this test may not be appropriate for the tests of benefit distributions of scenarios 1 to 5. Since the benefit distributions of scenarios 1 to 5 are truncated at zero, it is possible for them to have a fairly large proportion of zeros. In this case, the calculated Z will not be standard-normally distributed, and the Mann-Whitney test cannot be used. A more general distribution-free nonparametric test based on the convolutions approach can be used in this case.

6.3 Convolutions Approach

Another test is based on the method of convolutions and presented by Poe, Severance-Lossin and Welsh (1994). The convolutions approach is used in mathematics and statistics to calculate the distribution of a sum of random variables and series. The following discussion is heavily drawn from Poe, Severance-Lossin and Welsh (1994).

Assume that two independent random variables, X and Y, have respective probability density functions $f_x(X)$ and $f_y(Y)$. Define V=X-Y to be another random variable. The probability of the event V=v is defined as the union of all the possible combinations of x and y that result in X-Y=v. For continuous functions $f_x(X)$ and $f_y(Y)$, the probability density function of V is explicitly given as

$$f_v(v) = \int_{-\infty}^{\infty} f_Y(x-v)\, f_X(x)dx$$
$$= \int_{-\infty}^{\infty} f_X(v+y)\, f_Y(y)dy \tag{6.8}$$

Using only the far right-hand side of (6.8), the corresponding cumulative distribution $F_V(v=x-y=v^0)$ is

$$F_v(v^0) = \int_{-\infty}^{v^0} f_V(v)dv$$
$$= \int_{-\infty}^{v^0}\int_{-\infty}^{\infty} f_X(v+y)\, f_Y(y)dydv \tag{6.9}$$

In empirical applications with discrete observations, both $f_X(.)$ and $f_Y(.)$ have no explicit solutions but can be numerically approximated in discrete manner. Using max(.) and min(.) to replace the infinites, the dimensions of (6.9) can be reduced substantially. Imposing finite width windows (Δ) upon the continuum of the values associated with X and Y, the approximate cumulative empirical distribution of $F_v(v^0)$ is given by

$$\hat{F}_v(v^0) = \sum_{\min(\hat{x}-\hat{y})}^{v^0} \sum_{\min(\hat{y})}^{\max(\hat{x})} \hat{f}_X(\hat{v}+\hat{y})\, \hat{f}_Y(\hat{y})\Delta y \Delta v \tag{6.10}$$

where min(.) and max(.) denote minimum and maximum convoluted values, and "^" indicates that the distributions or values are a discrete approximation of a true underlying distribution or value. Since the distribution of $V=X-Y$ is generally unknown, an empirical approach to estimating confidence intervals is needed. Adopting a percentile approach (Efron, 1982), equation (6.10) can be directly applied to test the null hypothesis $H_0 : V=X-Y=0$, by calculating the lower bound, $\hat{L}_{1-\alpha}(\hat{V})$, and upper bound, $\hat{U}_{1-\alpha}(\hat{V})$, of the 1-$\alpha$ confidence intervals of V:

$$\begin{aligned} \hat{L}_{1-\alpha}(\hat{v}) &= \hat{F}^{-1}_{\hat{v}}(\alpha/2) \\ \hat{U}_{1-\alpha}(\hat{v}) &= \hat{F}^{-1}_{\hat{v}}(1-\alpha/2) \end{aligned} \tag{6.11}$$

The null hypothesis H_0 : V=X-Y=0 cannot be rejected at the size α if the approximate 1-α confidence interval of the convolution includes zero, and is rejected otherwise. Using the convolutions approach, the approximate two-sided significance level of the difference of the distributions is determined by $2\hat{F}_{\hat{v}}(0)$ *if* $\hat{F}_{\hat{v}}(0) \leq 0.5$, and $2[1 - \hat{F}_{\hat{v}}(0)]$ otherwise.

This approach can also be applied as a one tailed test. For example, assume a policy maker wants to test whether the difference between the "true" and transferred benefits is less than a given number v^0. A null hypothesis H_0: X-Y $\leq v^0$ can be tested. The significance level of the difference is given by $\hat{F}_{\hat{v}}(0)$.

Poe, Severance-Lossin and Welsh (1994, p 914) presented a simple example to demonstrate the application of the empirical convolution formula (6.10) and the suggested statistical test for estimating the significance of the difference of two approximate empirical distributions. Suppose that the following two hypothetical distributions are approximated from two random series.

Values	-2	-1	0	1	2	3	4	5
$f_x(.)$	0.00	0.00	0.00	0.00	0.10	0.40	0.40	0.10
$f_y(.)$	0.00	0.00	0.05	0.30	0.60	0.05	0.00	0.00

where $f_x(.)$ and $f_y(.)$ are the pdf of the random variable x and y respectively. We are interested in evaluating the difference ($v=X-Y$) between the two simple distributions. The first step is the calculation of the probability density function at a given point v^0. It is given by $f_v(v^0) = \Sigma\,[f_x(x)\ f_y(y)]$ for all $x\text{-}y = v^0$. For example, $f_v(2) = f_x(2)f_y(0)+f_x(3)f_y(1)+f_x(4)f_y(2)+f_x(5)f_y(3)$ $= (0.1)(0.05) + (0.4)(0.3) + (0.4)(0.6) + (0.1)(0.05) = 0.370$. The calculated probability density and cumulative density function are presented as follows:

Values	-2	-1	0	1	2	3	4	5
$f_v(.)$	0.000	0.005	0.080	0.290	0.370	0.200	0.050	0.005
$F_v(.)$	0.000	0.005	0.085	0.375	0.745	0.945	0.995	1.000

Evaluating $F_v(0)$ indicates that the distributions are different at a 17% ($=2\times0.085$)% level.

For large samples, the convolutions approach could be computationally costly. Poe, Severance-Lossin and Welsh kindly provided us their program. Their GAUSS program for performing a convolution of two vectors centres on the convolution (CONV) routine in GAUSS. It is in 6 steps[1] : (1) Input matrices: two GAUSS (*.fmt) matrices that contain the samples obtained from the simulations are read. (2) Determination of Interval (finite Windows) size: The interval is the Δ in Equation (6.10). Selecting or determining the size of the window is based on experience and "feel", much like determining the range and intervals to be considered in graphing a bar chart. The smaller the size, the more precise the approximation, but the longer the computational time. (3) Specification of the Size for Confidence Intervals: This section allows the user to choose the two-sided central confidence intervals ($1-\alpha$) by specifying a size α . (4) Organize Bounds for Convolution: This section reorders the matrices to conform to a large - small ordering, creating a pseudo data set by rounding to the upper bound of the corresponding interval, and prints out bounds and precision of the convolution. (5) Set up matrices for convolution: this section calculates probability density functions and eliminates unneeded series of zeros below the lowest point in which a non-zero probability is observed. (6) Conduct and Report the Convolution: this section uses GAUSS CONV. It is the heart of the program. This section places upper and lower limits on the convolution, and identifies and reports two-sided values of convolution at zero and the upper and lower confidence bounds.

Both the Mann-Whitney and the Convolutions Approach are capable of testing two

---

[1]See Poe, G.L., M.P.Welsh and E.K.Lossin (1994) for details.

random series. However, they are different in the statement of the null hypothesis. Assume there are two samples, $x_1, \ldots, x_n$ generated from the "true" benefit $B[\beta_e, \Omega_e(X_e, \beta_e), X_e]$, and $y_1, \ldots, y_m$ generated from the transferred benefit $B[\beta_s, \Omega_s(X_e, \beta_s), X_e]$. The Mann-Whitney test examines the null hypothesis, H_0 : $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ are drawn from identical populations, that is $B(\beta_s, \Omega_s(X_e, \beta_s), X_e] = B(\beta_e, \Omega_e(X_e, \beta_e), X_e]$. The convolutions approach tests the null hypothesis $H_0 : X - Y = 0$, that is $B(\beta_e, \Omega_e(X_e, \beta_e), X_e]$ and $B(\beta_s, \Omega_s(X_e, \beta_s), X_e]$ generate identical empirical distributions. Strictly speaking, the convolutions approach is not a rigorous statistical test for distributions. But since our purpose is to determine how different are the benefits generated from the "true" and transferred models, this approach is appropriate.

6.4 Test Results and Discussion

The test results for the 8 policy scenarios are presented in Table 12. The tests based on the convolutions approach are performed for all policy scenarios, while the Mann-Whitney test is only performed for scenarios 6 to 8 due to its limitation in the situation of ties. When the Mann-Whitney test is performed, Equation (6.7) is used to adjust the statistic for ties.

The following conclusions can be drawn from the test results in Table 12:

(1) The results of the test based on the convolutions approach suggests that, for policy scenarios 2, 5, 6, 7 and 8, the null hypothesis: benefits are transferable across the two regions, cannot be rejected. For example, in the test for scenario 8, α_{CONV} =85.001, which means that $\hat{F}_V(0) = 100 - \alpha_{CONV}/2 = 57.4995$, that is 57.5% of the X-Y is less than or equal to zero.

Table 12. Statistical Test Results for Benefit Transferability

Policy Scenarios	MW[1]	α^2_{CONV}	CI of Size α^3
Scenario 1		4.266	[-8.530, -0.008]
Scenario 2		9.768	[-55.010, 45.54]
Scenario 3		3.173	[-0.190, 0.000]
Scenario 4		1.981	[-0.020, -0.000]
Scenario 5		72.382	[-77.260, 66.760]
Scenario 6	1.921	99.189	[-59.760, 64.391]
Scenario 7	1.441	90.771	[-58.420, 55.250]
Scenario 8	7.847	85.001	[-104.32, 102.23]

1. Mann-Whitney Statistical Test
2. Significance level of the difference X-Y in the convolutions approach.
3. $\alpha = 0.05$.

(2) The confidence interval of size α is consistent with the significance level α_{CONV}. For all scenarios with a confidence interval containing zero, α_{CONV} are greater than 5.

(3) Although the Mann-Whitney test results are consistent with those of the convolution approach for scenarios 6 and 7, the Mann-Whitney test seems to be more conservative than the convolutions approach. The Mann Whitney test rejects the null hypothesis at a 15% level when the convolutions approach cannot.

(4) If the confidence interval of size $\alpha = 0.05$ is used to judge benefit transferability, we may conclude that the most of benefits generated from the two models are transferable.

We have discussed and applied two nonparametric tests for the purpose of testing benefit transferability. The Mann-Whitney test has been used in statistics for quite a long time and has been demonstrated to be a very efficient nonparametric test. Unfortunately, the null hypothesis is not exactly the same as the one we want to test. Since the two benefit samples are not directly generated from consumers but from two different models, the null hypothesis of identical distributions (populations) is much more restrictive than the null hypothesis of X-Y=0. Moreover, due to its limitation in dealing with a large number of ties, the Mann-Whitney test is not capable of testing some empirical distributions in applied economics.

The test based on the convolutions approach provides a general-purpose tool of testing two simulated empirical distributions. Although it may involve intensive computation for widely dispersed distributions, it is easy to perform in some computing packages that offer convolutions routines. However, the power and efficiency of this test are unknown. No statistical theory provides a base to perform this test. Serious problems could result from several aspects. First, efficiency could be lost in approximating $f_x(x)$, $f_y(y)$ and $\hat{F}_{\hat{v}}(0)$. Second, since the shape of $\hat{F}_{\hat{v}}$ is unknown, the consistency of the approximated confidence interval and significance level is in question.

Nevertheless, it should be reasonable to accept the approximations for our purpose. We have applied different test procedures and find similar results. Our results suggest that most of the benefits are transferrable from the study site to the policy site in our study.

Chapter 7. Test for the Economic Significance of Benefit Transfers

7.1 Introduction

Chapter 6 provided the nonparametric test results of benefit transfers. The tests examine the statistical significance of the differences between the "true" and the transferred benefits. However, it is extremely important for economists to carefully interpret these tests of "statistical significance". A statistical test can only deal with the question of whether a difference appears to be a chance variation or not. It is not designed to see whether the difference is important. Most importantly, economists would like to know whether the difference is important in economics or policy. A test for whether the difference is important in economic policy may be called the test of economic significance.

Statistical significance and economic significance are two different concepts. As McCloskey and Ziliak (1996) point out, statistical significance and economic significance must be distinguished from each other. A difference can be significant for science or policy and yet be insignificant statistically, and similarly, a statistically significant difference may be insignificant for science and policy. For example, a calculated "true" benefit of $5.65 with a standard error of 0.0021 is statistically different from a transferred benefit of $5.55 with a standard error of 0.001, but they may not be significantly different economically. On the other hand, a "true" benefit of $2.65 with a standard error of 1.35 is not statistically significantly different from a transferred benefit of $3.15 with a standard error of 1.55, but they may be significantly different economically. In economics this is a question of "How large is large".

While the term "statistical significance" has been widely used in economics, economic significance gets little attention. In many cases, statistical significance is used to decide nearly everything. McCloskey and Ziiliak (1996: p102) write: "Using ambiguously the very word "significance" implies there is no difference between economic significance and statistical significance, that nothing or little else matters. Of the 96 (*American Economic Review*) papers that use only the test of statistical significance as a criterion of importance at its first use, 90 percent imply (or state) that it is decisive in an empirical argument, and 70 percent use the "significance" ambiguously. .... Only seven of 96 distinguish statistical significance from economic or policy or scientific significance in the conclusions and implications sections."

In the previous chapter, the statistical significance of benefit transfer was tested, asking the question: Is the difference $B[\beta_s, \Omega_s(X_a, \beta_s), X_a] - B[\beta_a, \Omega_a(X_a, \beta_a), X_a]$ [1] of the two distributions of point estimates significantly different from zero? In this chapter, the economic significance of benefit transfer is tested by using policy simulations. The question asked here is: How important is the difference $B[\beta_s, \Omega_s(X_a, \beta_s), X_a] - B[\beta_a, \Omega_a(X_a, \beta_a), X_a]$ in a policy decision?

Benefit transfer is a method used by policy makers to estimate the benefit of a policy or project. Its economic significance depends on how the transferred benefit affects the policy maker's decision, compared to the "true" benefit. Two issues are essential for policy makers

---

1. $B[\beta_s, \Omega_s(X_a, \beta_s), X_a]$ is the transferred benefit distribution simulated from the Saskatchewan model and Alberta policy changes, and $B[\beta_a, \Omega_a(X_a, \beta_a), X_a]$ is the true benefit distribution simulated from the Alberta (the "true") model and Alberta policy changes.

in benefit transfer. First, if the transferred benefits are used, what is the probability of making an incorrect decision? Second, if the transferred benefits are used, instead of undertaking a new study to investigate the true benefit, what benefit can be expected? This chapter investigates these two issues to answer the question: How important is the difference $B(\beta_s, \Omega_s(X_e, \beta_s), X_e]$ - $B(\beta_e, \Omega_e(X_e, \beta_e), X_e]$ in a policy decision? If, for example, the probability of making an incorrect decision is very high, or the expected benefit is significantly negative, the difference $B(\beta_s, \Omega_s(X_e, \beta_s), X_e]$ - $B(\beta_e, \Omega_e(X_e, \beta_e), X_e]$ is very important in policy decisions. The conditions for the difference $B(\beta_s, \Omega_s(X_e, \beta_s), X_e]$ - $B(\beta_e, \Omega_e(X_e, \beta_e), X_e]$ to be economically significant are also investigated.

This chapter is organized as follows. The second part presents the policy maker's problem with respect to benefit transfer, and then from this problem, derives the calculations of the probability of making an incorrect decision and the expected benefit of benefit transfer. The third part actually calculates the probability and the expected benefit in a simulated policy setting, using the calculated true and transferred benefits of chapter 5. The final part contains the summary and conclusions.

7.2 The Policymaker's Problem in Benefit Transfer

Consider the policymaker's problem with respect to benefit transfer. When an environmental policy is evaluated, the policymaker needs to know the benefits and costs of the policy. The benefits consist of two components: market benefits and non-market benefits. By assuming both the cost and the market benefit are known, the policymaker focuses on the calculation of the non-market benefit. She has two options: conduct new research to investigate the "true" benefit, or use a transferred benefit. If the first option is taken, the "true" benefit is obtained. Then the policymaker makes the decision with full information and there is no risk of making the incorrect decision. However, benefit assessments are costly and time consuming. There are two types of cost associated with option 1: the money cost of the benefit assessments, C_m, and the value of the time needed for benefit assessments, C_v. If the benefit transfer option is taken, there will be no cost[2], that is, $C_m = C_v = 0$. However, there exists a risk that the policymaker will make an incorrect decision, because the transferred benefit and the true benefit may not be close enough.

The economic significance of benefit transfer can be expressed as the probability of making an incorrect decision. To simplify the calculation of this probability, several assumptions are made:

Assumption 1: All consumers are identical to the representative consumer, and the per trip benefit is an exogenous variable. If the benefit per trip is B, the aggregated benefit is nB, where n is the total number of trips.

Assumption 2: The calculated non-market benefit is the only non-market benefit associated with the proposed policy (project). The cost and the market revenue of the policy (project) are known and fixed as C and R, respectively.

Assumption 3: The policymaker's decision is made solely based on the total benefits and costs. For example, if the total cost is larger than the total

2. This is an assumption made to simplify the analysis. However, benefit transfer can be costless in practice.

benefit, the policy (project) will not be implemented, and vice-versa.

The incorrect decision is defined as: the policy or project is implemented when the total cost is larger than the total benefit, or that the policy or project is not implemented when the total benefit is larger than the total cost.

Based on the above assumptions, there are four possible outcomes when the transferred benefit is used. Assume B_a is the true benefit and B_t is the transferred benefit. The four outcomes can be shown as follows:

	$nB_t - C + R > 0$	$nB_t - C + R < 0$
$nB_a - C + R > 0$	Correct Decision	Incorrect Decision
$nB_a - C + R < 0$	Incorrect Decision	Correct Decision

There are two situations in which the policymaker will make the wrong decision: $nB_a - C + R < 0$ and $nB_t - C + R > 0$, or $nB_a - C + R > 0$ and $nB_t - C + R < 0$. In the other two situations where $nB_a - C + R < 0$ and $nB_t - C + R < 0$, or $nB_a - C + R > 0$ and $nB_t - C + R > 0$, the policymaker's decision will be the same, no matter whether the true benefit or the transferred benefit is used.

When the transferred benefit is used for the policy decision, the probability of making an incorrect decision can be expressed as:

$$P_w = P_1 + P_2 = P(\ nB_a - C + R > 0\ ,\ nB_s - C + R < 0\ ) \\ + P(\ nB_a - C + R < 0\ ,\ nB_s - C + R > 0\ ) \tag{7.1}$$

and the probability of making the right decision is

$$P_r = P_3 + P_4 = P(\ nB_a - C + R > 0\ ,\ nB_s - C + R > 0\ ) \\ + P(\ nB_a - C + R < 0\ ,\ nB_s - C + R < 0\ ) \tag{7.2}$$

Assume that the representative consumer's "true" and transferred benefits are normally and independently distributed[3] as $B_a \sim N(\mu_a , \sigma_a)$, and $B_s \sim N(\mu_s , \sigma_s)$, respectively. The probability P_1 can be calculated as

3. Considering the normality test results in Chapter 6, this is a restrictive assumption. However, it is necessary for the simplification of the demonstration.

$$P_1 = P(\, nB_a - C + R > 0 \,,\, nB_s - C + R < 0 \,)$$

$$= P(B_a > \frac{C-R}{n}) \, P(\, B_s > \frac{C-R}{n})$$

$$= P(\frac{B_a - \mu_a}{\sigma_a} > \frac{\frac{C-R}{n} - \mu_a}{\sigma_a}) \, P(\frac{B_s - \mu_s}{\sigma_s} > \frac{\frac{C-R}{n} - \mu_s}{\sigma_s})$$

$$= P(\, Z > \frac{\frac{C-R}{n} - \mu_a}{\sigma_a}) \, P(Z < \frac{\frac{C-R}{n} - \mu_s}{\sigma_s}) = (1 - \Phi(z_1)) \, \Phi(z_2) \tag{7.3}$$

where

$$z_1 = \frac{\frac{C-R}{n} - \mu_a}{\sigma_a} \,,\, z_2 = \frac{\frac{C-R}{n} - \mu_s}{\sigma_s} \,,$$

$$\Phi(z) = P(\, Z \le z) \,,\quad Z \sim N(0 \,,\, 1)$$

Similarly, the probabilities of the three other events could be calculated as:

$$\begin{aligned} P_2 &= \Phi(z_1) \, (1 - \Phi(z_2)) \\ P_3 &= (\, 1 - \Phi(z_1)) \, (\, 1 - \Phi(z_2)) \\ P_4 &= \Phi(z_1)\Phi(z_2) \end{aligned} \tag{7.4}$$

For any given set of C, R, and n, P_1, P_2, P_3, and P_4 can be calculated. The probability of making an incorrect decision $P_w = P_1 + P_2$ could be used as an indicator of the economic significance of benefit transfer.

Another indicator for economic significance may be the expected benefit of benefit transfer. The policymaker may want to consider what is the benefit of benefit transfer compared to investigating the true benefit. To calculated the expected benefit, four situations need to be considered:

1. If $nB_a - C + R < 0$ and $nB_s - C + R > 0$, the expected benefit from benefit transfer is $B_1 = C_m + C_v + (n\mu_a - C + R)$, where $C_m + C_v$ is the money saved by using benefit transfer, and $nB_a - C + R$ is the money lost due to making an incorrect decision.

2. If $nB_a - C + R > 0$ and $nB_s - C + R < 0$, the benefit from benefit transfer is $B_2 = C_m + C_v - (n\mu_a - C + R)$, where $C_m + C_v$ is the money saved by using benefit transfer, and $nB_a - C + R$ is the money lost due to making an incorrect decision.

3. If $nB_a - C + R < 0$ and $nB_s - C + R < 0$, the benefit from benefit transfer is $B_3 = C_m + C_v$. There is no loss from making an incorrect decision in this case.

4. If $nB_a - C + R > 0$ and $nB_s - C + R > 0$, the benefit from benefit transfer is $B_4 = C_m + C_v$. There is no loss from making an incorrect decision in this case.

The expected benefit then could be calculated as

$$E(B) = P_1B_1 + P_2B_2 + P_3B_3 + P_4B_4 \tag{7.5}$$

If $E(B) > 0$, policymaker will benefit from benefit transfer, and if $E(B) < 0$, she will lose from benefit transfer.

The two indicators: the probability of making an incorrect decision and the expected benefit of benefit transfer, should provide sufficient information for testing the economic significance of benefit transfer.

In addition to the two indicators, one condition may also be useful in the policymaker's decision. If the policymaker wants to know under what conditions the probability of making an incorrect decision is limited to 5%, the following formula could be used:

$$
\begin{aligned}
0.05 &= P_1 + P_2 = (1 - \Phi(z_1))\Phi(z_2) + (1 - \Phi(z_2))\Phi(z_1) \\
&= \Phi(z_1) + \Phi(z_2) - 2\Phi(z_1)\Phi(z_2) \\
&\quad \textit{where} \qquad\qquad (7.6) \\
\Phi(z) &= \int_{-\infty}^{z} \frac{1}{(2\pi)^{1/2}} e^{-\frac{x^2}{2}} dx
\end{aligned}
$$

Since $z_1 = (C - R)/n - \mu_a/\sigma_a$ and $z_2 = (C - R) - \mu_s/\sigma_s$, a numerical solution for (C-R)/n could be found by using specialized computer software, given μ_a/σ_a and μ_s/σ_s. The variable (C-R)/n may be interpreted as per trip market cost of the policy.

7.3 Test for the Economic Significance of Benefit Transfer

Based on the methodology discussed above, this section actually tests the economic significance of benefit transfer, using the results of Chapter 6. As the first step, two benefit distributions, the benefit distribution of policy scenario 5 and the benefit distribution of policy scenario 8, are simulated[4]. To approximate normal distributions, batch means are used as the

4. Policy scenarios 5 and 8 are defined as the same as they were in Chapter 6. Policy scenario 5 is defined as: eliminating WMU348, and policy scenario 8 is defined as: reducing the congestion level of WMU348 from *cong3* to *cong1*.

point estimates of the benefits. That is, the mean of each 1000 replications is used as one point estimate, and 1000 means are simulated. It is suggested that nonnormality will not be a problem if the number of batches is larger than 40 (Law, 1977). This batch mean technique is consistent with the technique used in Chapter 6. The only difference is that the expected benefit of each 1000 replications in Chapter 6 is treated as one point estimate.

The simulation results of the "true" and transferred benefit distributions of the policy scenarios 5 and 8 are as follows:

	Scenario 8	Scenario 5
True Benefit	$B_a \sim N(30.72, 4.08)$	$B_a \sim N(-14.58, 0.67)$
Transferred Benefit	$B_s \sim N(24.04, 5.53)$	$B_s \sim N(-14.27, 0.76)$

It can be seen that, while the difference between the "true" and transferred benefit of policy scenario 5 is small, the difference between the "true" and transferred benefit of scenario 8 is relatively large. However, the importance of these differences are determined by the probabilities of making an incorrect decision.

Recalling equation (7.1), (7.2) and (7.3), the probability of making an incorrect decision could be written as

$$P_w = P_1 + P_2 = \Phi(z_1) + \Phi(z_2) - 2\Phi(z_1)\Phi(z_2)$$

where

$$z_1 = \frac{\frac{C-R}{n} - \mu_a}{\sigma_a} \qquad (7.7)$$

$$z_2 = \frac{\frac{C-R}{n} - \mu_s}{\sigma_s}$$

Given the distributions of the "true" and transferred benefits, P_w is determined by (C - R)/n. (C-R)/n could be interpreted as the per trip market cost of the policy or project. This number is a key factor in determining the probability of making an incorrect decision when benefit transfer is conducted. If, for example, this number is close to the "true" per trip non-market benefit μ_a, then a small deviation of the transferred benefit μ_s from the "true" benefit will lead to a high probability of making an incorrect decision. On the other hand, if $(C-R)/n \gg \mu_a$ or $(C-R)/n \ll \mu_a$, the difference between μ_a and μ_s may not be large enough to affect the policymaker's decision. So, when the transferred benefit μ_s is used to replace the "true" benefit μ_a, the smaller the value of $|(C-R)/n - \mu_a|$, the higher the possibility of making an incorrect decision; and the larger the value of $|(C-R)/n - \mu_a|$, the lower the possibility of making an incorrect decision.

To calculate the probability of making an incorrect decision, (C-R)/n has to be known. In order to have a better understanding of the relationship between the size of (C-R)/n and the probability of making an incorrect decision, a variable is defined as

$$y = \left(\frac{C - R}{n} - \mu_a\right)/\mu_a \tag{7.8}$$

Here y is the percentage deviation of the (C-R)/n from the "true" mean benefit. Again, the larger the $|y|$, the lower the possibility of making an incorrect decision.

Rewrite (7.8) as

$$\frac{C - R}{n} = (1 + y)\,\mu_a \tag{7.9}$$

Substituting (7.9) into (7.7) gives an equation describing the relationship between y and P_w, the probability of making an incorrect decision.

Using the statistical software Splus, the relationship between y and P_w is calculated. The results for policy scenarios 5 and 8 are displayed in Table 13. For policy scenario 5, benefit transfer is quite promising. As long as the per trip market cost *(C-R)/n* is 10% larger or 10% less than the per trip benefit, the probability of making an incorrect decision is very low, less than 2.5% and 7.9% respectively. The highest possibility of making an incorrect decision is 35%, when the per trip market cost is 5% less than the "true" benefit. When the per trip market cost (C-R)/n is 10% larger or 15% less than the "true" benefit, the probability of making an incorrect decision is 0, that is, benefit transfer will result in the exactly same decision as the "true" benefit.

For policy scenario 8, benefit transfer provides a less promising result. This is because the difference between the "true" and transferred benefit is larger in scenario 8 than in scenario 5. The highest possibility of making wrong decision is 63.3%, when the per trip

Table 13. The Probabilities of Making an Incorrect Decision at Different Per Trip Market Costs

	Scenario 8		Scenario 5	
y	Market Cost/Trip[1]	P_w	Market Cost/Trip	P_w
0.05	32.56	37.3%	-15.31	20.0%
0.10	33.79	24.7%	-16.04	2.5%
0.15	35.33	12.8%	-16.77	0.0%
0.20	36.86	7.5%	-17.50	0.0%
0.30	39.94	1.4%	-18.95	0.0%
0.40	43.01	0.2%	-20.41	0.0%
0.50	46.08	0.0%	-21.87	0.0%
0.60	49.15	0.0%	-23.33	0.0%
0.65	50.69	0.0%	-24.06	0.0%
-0.05	29.18	60.7%	-13.85	35.0%
-0.10	27.65	63.3%	-13.12	7.9%
-0.15	26.11	60.8%	-12.39	0.7%
-0.20	24.58	53.3%	-11.66	0.0%
-0.30	21.50	32.7%	-10.21	0.0%
-0.40	18.43	15.6%	-8.75	0.0%
-0.50	15.36	5.8%	-7.29	0.0%
-0.60	12.29	1.6%	-5.83	0.0%
-0.65	10.75	0.0%	-5.10	0.0%

1. Market cost/trip is (C-R)/n. It is calculated as $(1+y)\mu_s$, where μ_s =30.72 in Scenario 8 and -14.58 in Scenario 5.

market cost is 10% less than the "true" benefit. However, when (C-R)/n is in the intervals [13, -∞] and [43, +∞], the probability of making an incorrect decision is 0.

Using the software Splus, the relationships between y and P_w for the two policy scenarios are also presented in Figure 11 and 12. These figures are drawn by calculating 200 points of y and P_w. The x-axis is y, the deviation of (C-R)/n from the mean benefit, and the y-axis is P_w, the probability of making an incorrect decision. Figure 11 is for policy scenario 5 and Figure 12 for policy scenario 8.

Now, the problem is considered inversely. Assume that the probability of making an incorrect decision is given. One can calculate the per trip market cost (C-R)/n or the deviation y by using (7.6) and (7.8).

The mathematical software program MatLab is used for this task. Assume $P_w = 5\%$. The solutions for y and (C-R)/n are obtained by solving (7.6) numerically. For policy scenario 5, $y_1 = -8.6\%$ ([C-R]/n = -15.81), and $y_2 = 9.1\%$ ([C-R]/n = -12.95). For policy scenario 8, $y_1 = -51.4\%$ ([C-R]/n = 14.94), and $y_2 = 25.9\%$ ([C-R]/n = 38.68). The interpretations of these numbers are as follows: For Scenario 5, the probability of making an incorrect decision is less than 5%, as long as the per trip market cost (C-R)/n is less than -15.81 or larger than 12.59. For Scenario 8, the probability of making an incorrect decision is less than 5%, as long as the per trip cost is less than 14.94 or larger than 38.68.

The second indicator of economic significance - policymaker's benefit from benefit transfer - is calculated using (7.5). The probabilities of the four events P_1, P_2, P_3, P_4 are calculated by Splus. To be consistent with the first indicator, the same combinations of the per trip market cost and benefit are assumed.

Table 14. **Policymaker's Benefits from Benefit Transfer Under Different Market Costs**

	Scenario 8 Expected Benefit	Scenario 5y Expected Benefit
0.05	$C_m + C_v - 0.4377n$[1]	$C_m + C_v - 0.0385n$
0.10	$C_m + C_v - 0.5738n$	$C_m + C_v - 0.0072n$
0.15	$C_m + C_v - 0.5009n$	$C_m + C_v - 0.0011n$
0.20	$C_m + C_v - 0.3431n$	$C_m + C_v$
0.30	$C_m + C_v - 0.0916n$	$C_m + C_v$
0.40	$C_m + C_v - 0.0123n$	$C_m + C_v$
0.50	$C_m + C_v - 0.0015n$	$C_m + C_v$
0.60	$C_m + C_v - 0.0000n$	$C_m + C_v$
-0.05	$C_m + C_v - 0.7229n$	$C_m + C_v - 0.1121n$
-0.10	$C_m + C_v - 1.5848n$	$C_m + C_v - 0.0758n$
-0.15	$C_m + C_v - 2.7312n$	$C_m + C_v - 0.0139n$
-0.20	$C_m + C_v - 2.9834n$	$C_m + C_v - 0.0009n$
-0.30	$C_m + C_v - 2.8652n$	$C_m + C_v$
-0.40	$C_m + C_v - 1.8898n$	$C_m + C_v$
-0.50	$C_m + C_v - 0.8896n$	$C_m + C_v$
-0.60	$C_m + C_v - 0.3077n$	$C_m + C_v$
-0.65	$C_m + C_v - 0.1612n$	$C_m + C_v$

1. C_m and is C_v are the money cost and value of time of benefit investigation; n is the total number of trips.

Multiplying the calculated P_1, P_2, P_3, P_4 by the benefits of different situations, the policymaker's benefit is calculated and presented in Table 14. These results are consistent with the probabilities of making an incorrect decision. Theoretically, the highest expected benefit from benefit transfer is $C_m + C_v$, the sum of the money cost and the value of time of doing benefit investigation. For Policy Scenario 8, when the per trip market cost is 60% greater than the benefit, the expected benefit from benefit transfer is $C_m + C_v$. For Policy Scenario 5, the highest expected benefits occur when the per trip cost is 20% larger or 30% less than the benefit.

In general, the higher the probability of making an incorrect decision, the lower the policymaker's expected benefit from benefit transfer. However, since the expected loss also depends on $n\mu_a - C + R$, it is possible that when the deviation y increases the expected benefit decreases. For example, the expected benefit for Scenario 8 at y =10% is $C_m + C_v - 0.5738n$, which is less than $C_m + C_v - 0.4377n$, the expected benefit at y =5%.

From the formula of expected benefit, it can be seen that given the per trip market cost, the higher the money cost and the value of time of doing benefit investigation, the higher the policymaker's expected benefit from benefit transfer, and the larger the number of trips, the lower the expected benefit.

7.3 Summary and Conclusions

It is important for economists to distinguish economic significance from statistical significance. This chapter has developed and applied a method to test the economic significance of benefit transfer. Two indicators, the probability of making an incorrect decision and the expected benefit of benefit transfer, are used to test the economic significance of benefit transfer.

Applying the "true" and transferred benefits of Policy Scenarios 5 and 8, it is found that the possibility of making an incorrect decision is reasonablely low when benefit transfer is applied. For Policy Scenario 5, the probability of making an incorrect decision is lower than 5% when the per trip market cost is 10% higher or 15% lower than the per trip benefit. For Policy Scenario 8, the probability of making an incorrect decision is less than 5% when the per trip market cost is 20% higher or 50% lower than the per trip benefit. These conditions are not very restrictive, considering the relative importance of nonmarket benefits in decision making. In most practical policy issues, the market value of the policy (project) may be much larger or smaller than the nonmarket benefit. In these cases, policymaking will gain from benefit transfers.

The calculated expected benefits of benefit transfer are consistent with the probability of making an incorrect decision. When the probability of making an incorrect decision is 0, the expected benefit reaches the highest point $C_m + C_t$. Also, the expected benefit is positively related to the money cost and the value of time in benefit investigation.

Using specialised software, an equilibrium condition is also calculated. It is suggested that for Policy Scenario 5, the probability of making an incorrect decision is less than 5%, as long as the per trip market cost is in the interval $[-\infty, -15.81]$ or $[12.59, +\infty]$; and for Scenario 8, the probability of making an incorrect decision is less than 5%, as long as the per trip cost is in the interval $[-\infty, 14.94]$ or $[38.68, +\infty]$.

Since many of the policy parameters are unknown, all the calculations are made in a simulated policy environment. Fortunately, with the aid of a variety of softwares, a range of solutions is obtained.

This chapter has developed a procedure to examine the economic significance of benefit transfers. It provides a tool for policymakers and managers to evaluate the feasibilities of benefit transfers. The test results in this specific experiment show that benefit transfers are quite promising. However, caution is required when these results are used in general applications. More work needs to be done to increase the stock of studies before general conclusions can be made.

Chapter 8. Conclusions and Discussions

Motivated by both intensive resource management and assorted legislative and juridical mandates, public and private agencies are continuing to expend considerable resources to quantify economic consequences of altering service flows and stocks of non-marketed features of the natural environment. The increasing demand for non-market valuation of environmental assets has brought about the idea of using benefit transfer as a time-saving and cost-effective alternative to environmental valuation. This thesis employs two stated preference surveys to investigate the issues in benefit transfers. It focuses on three aspects: benefit estimation with multiple sources of heterogeneity, statistical tests for benefit transferability and economic significance test of benefit transfers.

One major problem in benefit estimation is heterogeneous preferences among individuals. In order to obtain unbiased and consistent model and benefit estimates and thus provide a sound base for benefit transferability tests, two models are developed to account for multiple sources of heterogeneity in choice data. The first one is the heterogenous multinomial logit model. This model accounts for heterogeneity by specifying a relative scale factor for each group of individuals with specific characteristics. It is capable of dealing with observable multiple sources of heterogeneity. The results of the applications of this model in both the Alberta and the Saskatchewan data sets suggest that the heterogeneous model specification improves the model's goodness of fit, compared to the traditional multinomial logit model. The benefit calculation formula in the heterogeneous multinomial logit model is also derived. It has been shown that since the variance of choice is a part of the benefit formula, properly controlling for the heterogeneity in choice data is particularly important when the model is used for environmental valuation. Using the heterogeneous multinomial model to accounting for heterogeneity has improved benefit transfer in our experiment. For the designed policy scenario, the deviations of the transferred benefits from the true benefit decreased from 48% in the traditional multinomial logit model to 14% (urban) and 26% (rural) in the heterogeneous model.

The second model employed in this thesis is the random coefficient multinomial probit model. This model accounts for the heterogeneity in choice data in a more flexible way. It accounts for multiple sources of heterogeneity by specifying a random component for each coefficient of the indirect utility function and releases the IIA restriction by allowing a general covariance matrix for the errors. Using GAUSS Maximum Likelihood 4.0, an estimation procedure was developed for the stated preference data. The results of the random coefficient probit models have revealed significant variation among individual preferences. The most important and interesting finding is that the constant or average parts of the indirect utility functions are very close between the "true" and transferred sites, while the varying or heterogeneous parts are significantly different from each other. This finding implies that, since the so-called "intrinsic" indirect utility functions are very similar, benefits are expected to be more transferable after removing the heterogeneous preferences. It is also found that the random coefficient multinomial probit model provides the best goodness of fit. The result of the comparison between the independent model and the random coefficient model suggests that the random coefficient specification more adequately explains each individual's taste or perception of the site quality.

A consistent frequency simulator is employed to calculate the benefits of different environmental changes. The results from the eight policy scenarios show that benefits of the representative consumer have much larger variations after the heterogenous preferences are accounted for. The simulated benefits are much more transferrable in the random coefficient probit models than in those other models. It is also found that the calculated benefits are consistent with the simulated probabilities, that is, the higher the probability of visiting the site, the larger the loss of eliminating the site.

After obtaining the benefits from the random coefficient multinomial probit model, the statistical and economic significance of benefit transfer are examined in Chapters 6 and 7. Chapter 6 employs two nonparametric procedures, the Mann-Whitney test and the convolutions approach, to test benefit transferability statistically. The tests based on the convolutions approach are performed for all policy scenarios, while the Mann-Whitney test is only performed for scenarios 6 to 8 due to its limitation in the situations of ties. The results of the test based on the convolutions approach suggest that, for policy scenarios 2, 5, 6, 7 and 8, the null hypothesis: benefits are transferable across the two regions, cannot be rejected at a 5% level. Although the Mann-Whitney test results are consistent with those of the convolution approach for scenarios 6 and 7, the Mann-Whitney test seems to be more conservative than the convolutions approach. The Mann Whitney test rejects the null hypothesis at 15% level when the convolutions approach cannot. If the confidence interval of size $\alpha = 0.05$ is used to judge benefit transferability, it may be concluded that most of the benefits generated from the two models are transferable statistically.

It is extremely important for economists to distinguish economic significance from statistical significance. Chapter 7 has developed and applied a method to test the economic significance of benefit transfer. Two indicators, the probability of making an incorrect decision and the expected benefit of benefit transfer, are used for the test of economic significance of benefit transfer.

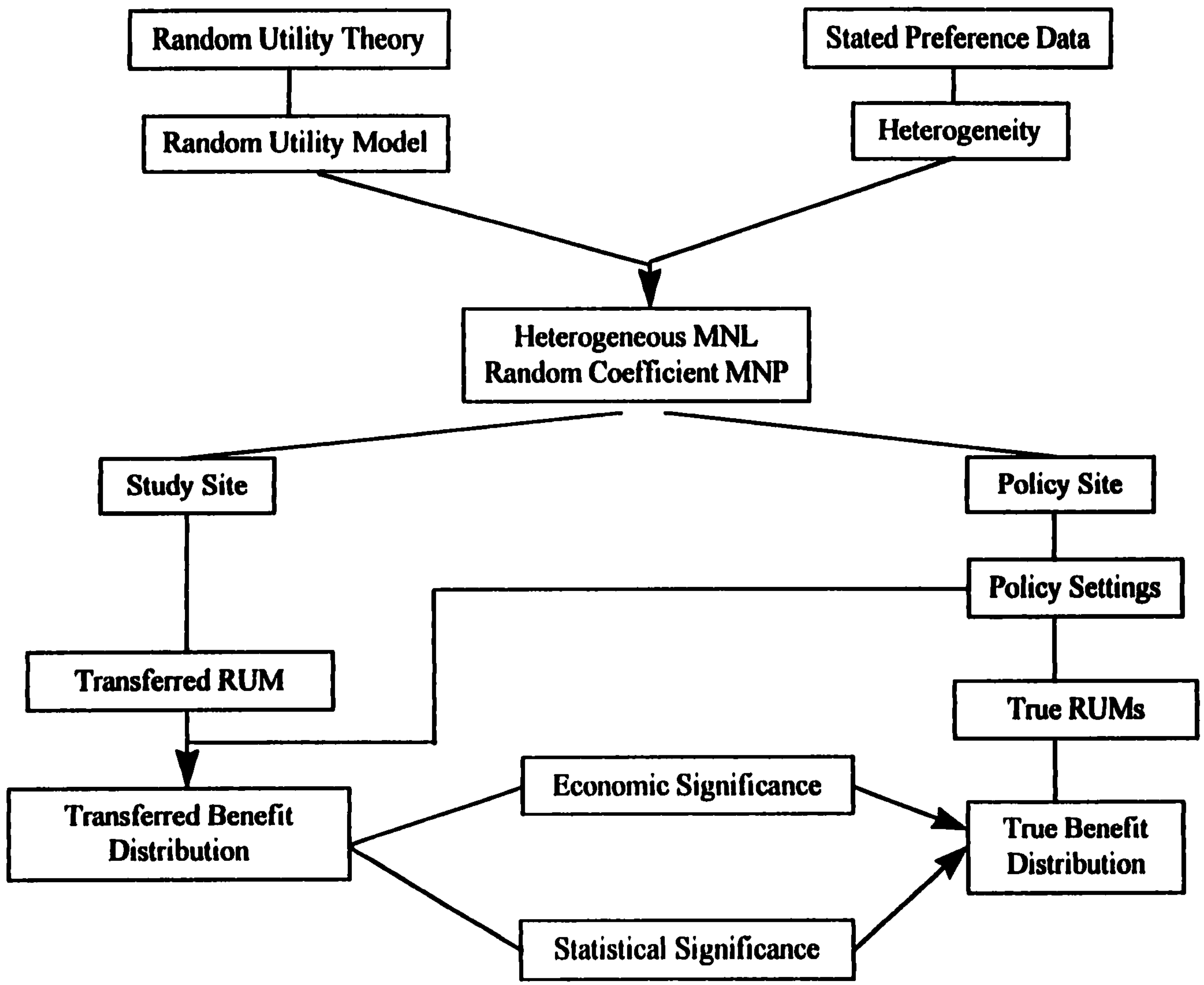
Applying the "true" and transferred benefits of policy scenario 5 and 8 in a case study, it is found that the possibility of making an incorrect decision is reasonablely low when benefit transfer is applied. In this case for policy scenario 5, the probability of making an incorrect decision is lower than 5% when the per trip market cost is 10% higher or 15% lower than the per trip benefit. For policy scenario 8, the probability of making an incorrect decision is less than 5% when the per trip market cost is 20% higher or 50% lower than the per trip benefit.

The calculated expected benefits of benefit transfer are consistent with the probability of making an incorrect decision. When the probability of making an incorrect decision is 0, the expected benefit reaches the highest point. It is also suggested that the expected benefit is positively related to the money cost and the value of time of benefit investigation, but negatively related to the total number of trips.

This thesis has provided a case study of benefit transfer and examined several important issues in the application of benefit transfer. The following important recommendations can be made from this study: (1) Benefit transfer can only be as accurate as benefit estimates. More accurate benefit estimates are more transferable. In order to obtain more accurate benefit estimates, heterogeneous preferences need to be accounted for if the computing capacity is available. If computing capacity is limited, scale factors (variances) could be considered since they account for heterogeneity and may enhance benefit calculations. (2) The more important the non-market value of a project (policy) in decision making, the higher the risk of making an incorrect decision. Thus, more caution should be used when applying benefit transfer to a project (policy) with a large non-market benefit (cost) component. (3) Benefit transfer is not as simple as a transformation of a number or a model. Doing benefit transfer correctly may in fact be quite costly in terms of time and effort.

This thesis has studied several important issues in benefit transfers. Several advanced econometric and simulation techniques have been developed and applied. These techniques are very important tools in the practical use of benefit transfers. However, in order to apply benefit transfers in practices, more work has to be done. First, more nonmarket valuation studies are needed to increase the availability of a stock of studies for consideration in the benefit transfer area. The suggestion of Boyle and Bergstrom (1992) and Atkinson et al. (1992) to build a nonmarket library may be a good solution. Second, the accuracy of the transferred nonmarket studies is critical. Benefit transfers are only as accurate as the initial studies. Further development of the existing knowledge base of both the models and benefit calculations is required to increase accuracy. Third, systematic comparisons of multisite models for different regions are needed to investigate the robustness of the benefit transfer process. Finally, some updating techniques may be used to increase the accuracy of benefit transfer when partial information about the policy site is available. In a previous study (Xu and Adamowicz, 1996), we find that Bayesian updating provides more transferrable benefits. Ben-Akiva, Bolduc and Pene(1995) suggest a combined estimator approach to model transferability and updating. They have shown that this estimator has superior accuracy, in a mean square error sense, to an unbiased direct estimator whenever the transfer bias is relatively small. This method is quite useful in practical benefit transfers.

Figure 1: Study Plan



CHOICE OF MOOSE HUNTING SITE

In this section you will examine 16 different scenarios which offer you the choice of hunting moose at two different sites or not hunting, Please assume that the two sites presented in each scenario are the only sites that you can choose from for your next hunting trip. We want you to indicate for each scenario which site you would choose, if either.

The enclosed information sheet entitled "Glossary of Terms" provides detailed information about the terms used in this section of the survey.

---

1. Assuming that the following areas were the **ONLY** areas available, which one would you choose on your next hunting trip, if either?

Features of Hunting Area	Site A	Site B	
[redacted]	50 kilometres	[redacted]	Neither Site A or Site B I will NOT go moose hunting
	Mostly gravel or dirt, some paved		
	Newer trails, cutlines or seismic lines, passable with a 2WD vehicle		
	No hunters, other than those in my hunting party, are encountered		
	Some evidence of recent logging found in the area		
	Evidence of less than 1 moose per day		

Check ONE and only one box ☐ ☐ ☐

Please complete all 16 of the scenarios that follow. <u>Missing any of these questions will not allow us to properly analyze your choices!</u>

Figure 2 Example of the instrument used to gather stated preference data

Figure 3. PDF of the True and Transferred Benefits for Policy Scenario 1.

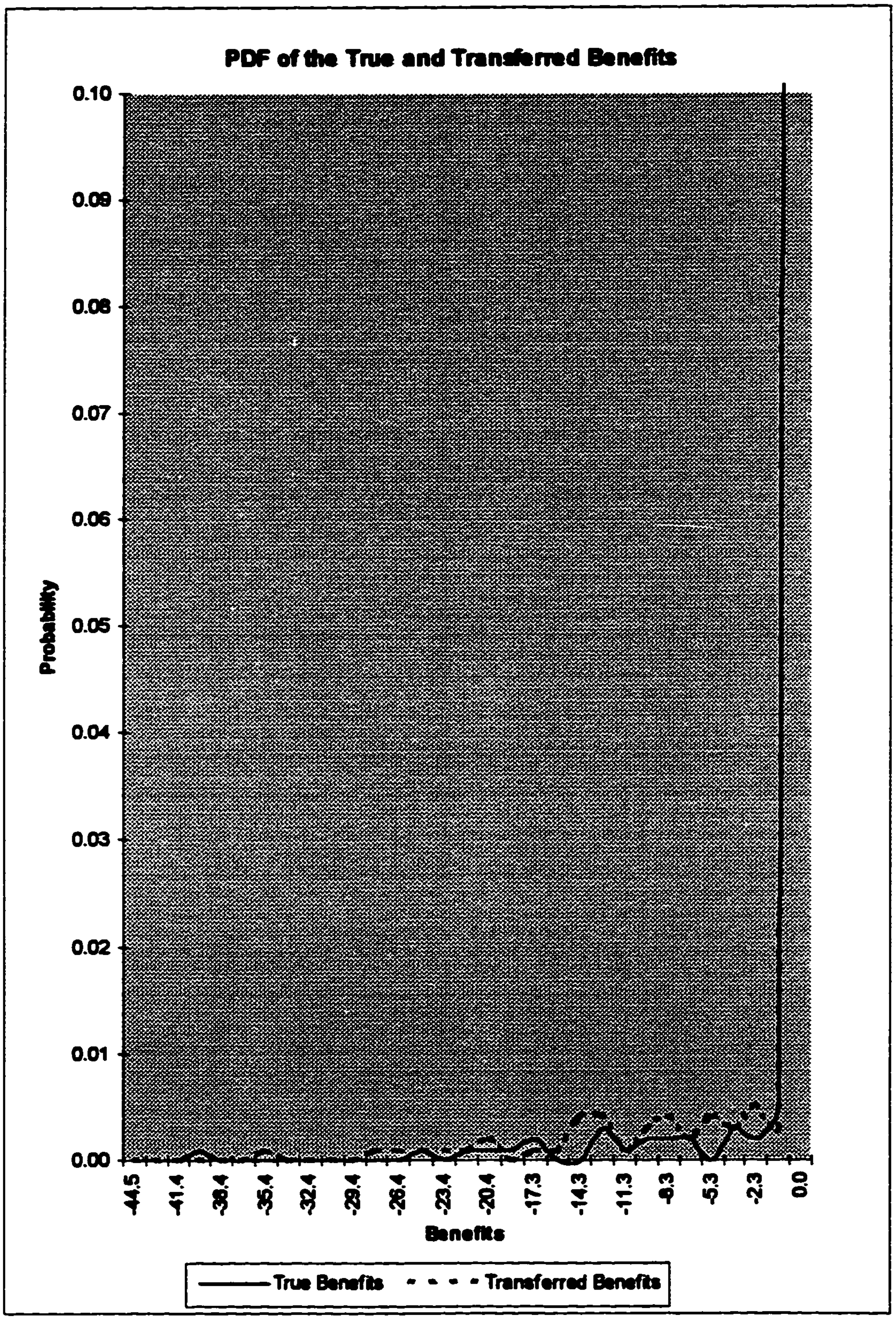


Figure 4. PDF of the True and Transferred Benefits for Policy Scenario 2.

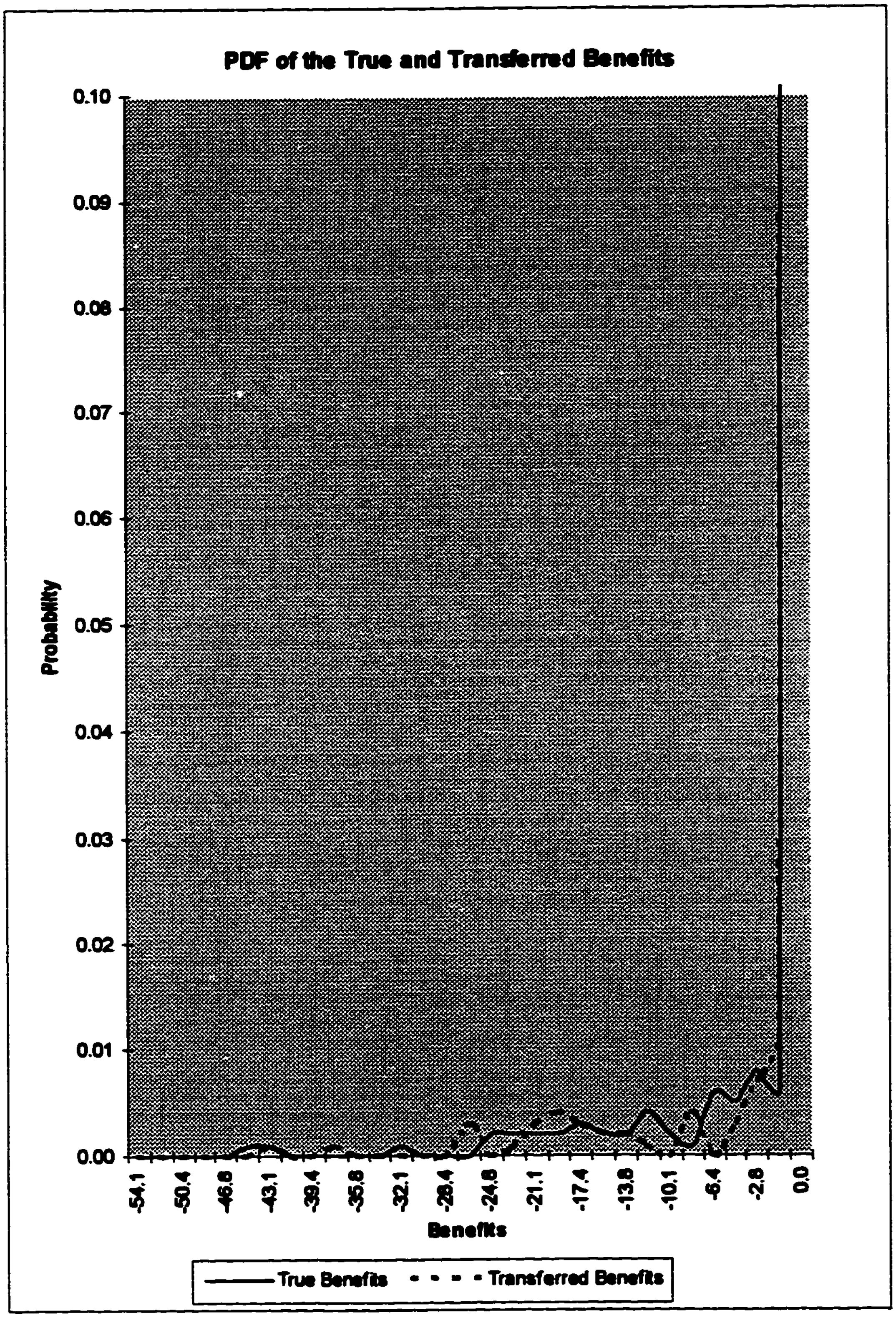


Figure 5. PDF of the True and Transferred Benefits for Policy Scenario 3.

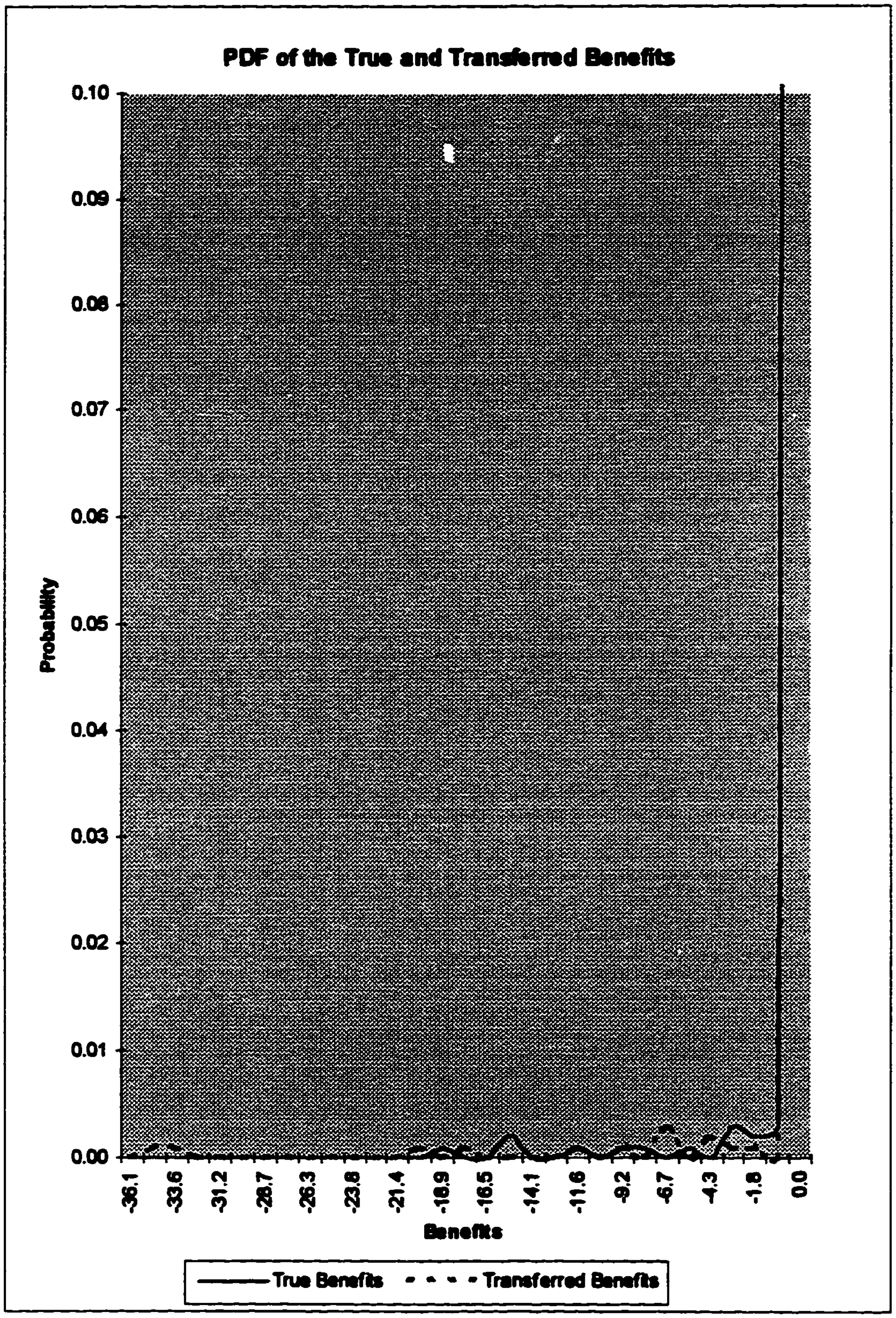


Figure 6. PDF of the True and Transferred Benefits for Policy Scenario 4.

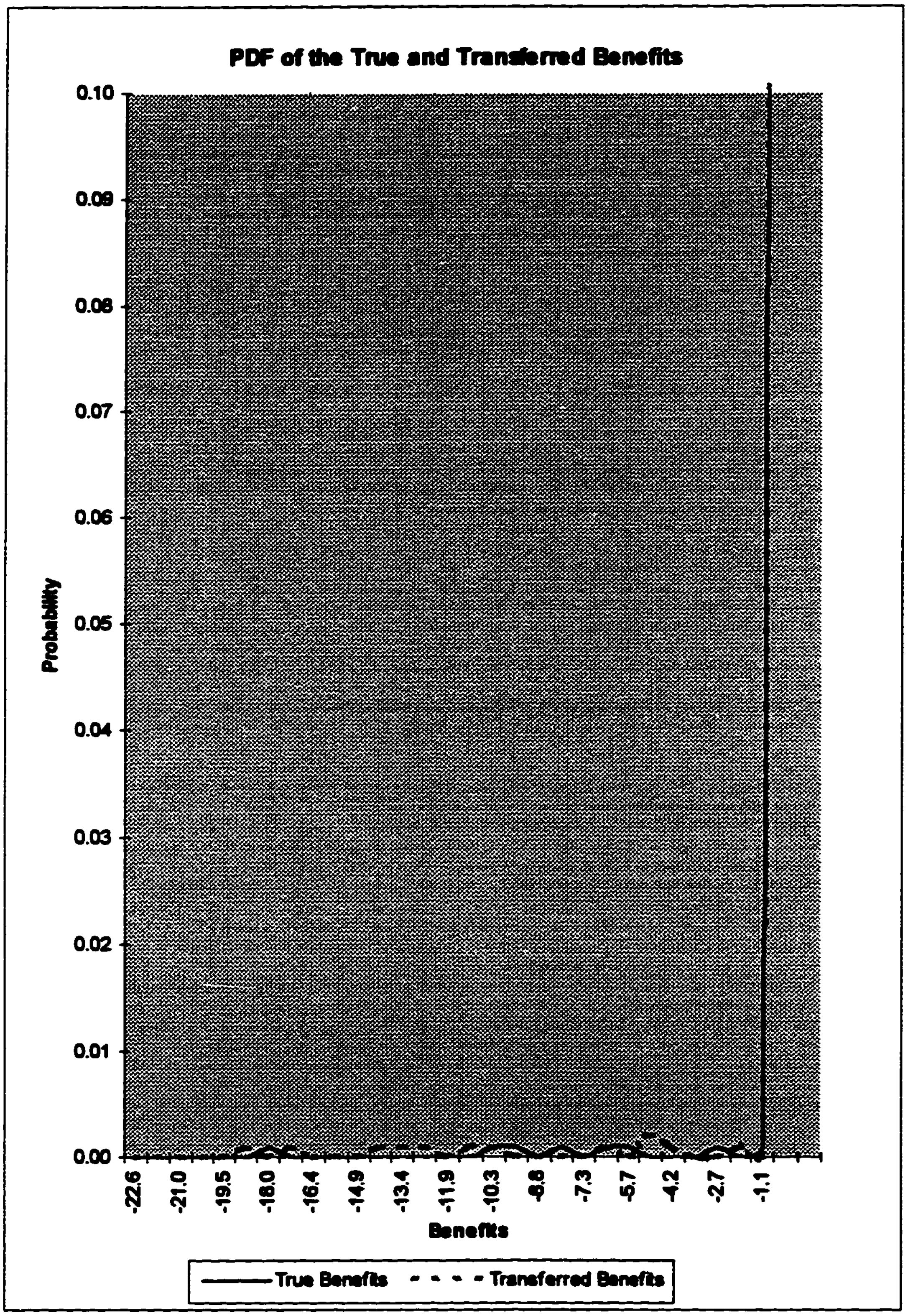


Figure 7. PDF of the True and Transferred Benefits for Policy Scenario 5.

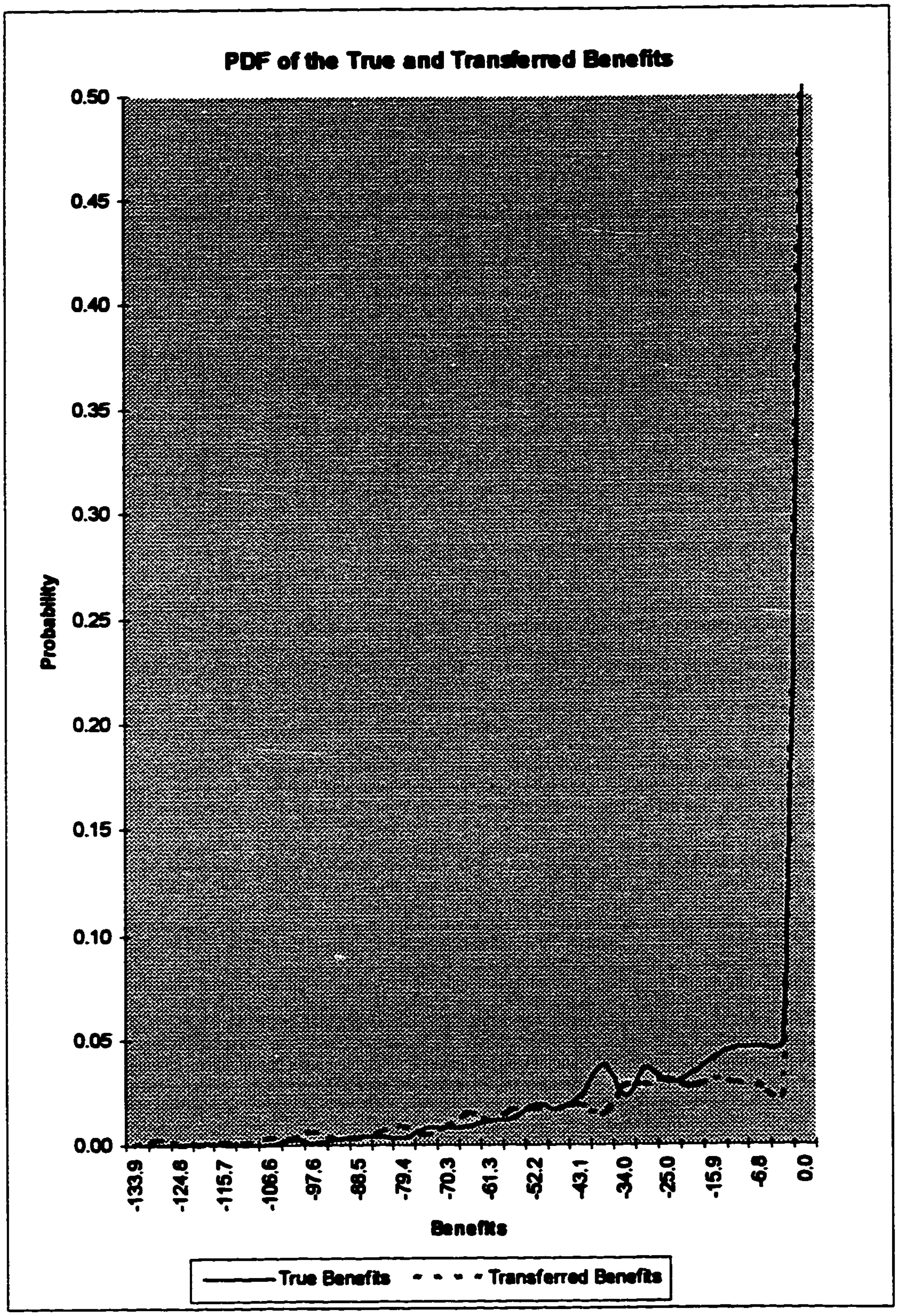


Figure 8. PDF of the True and Transferred Benefits for Policy Scenario 6.

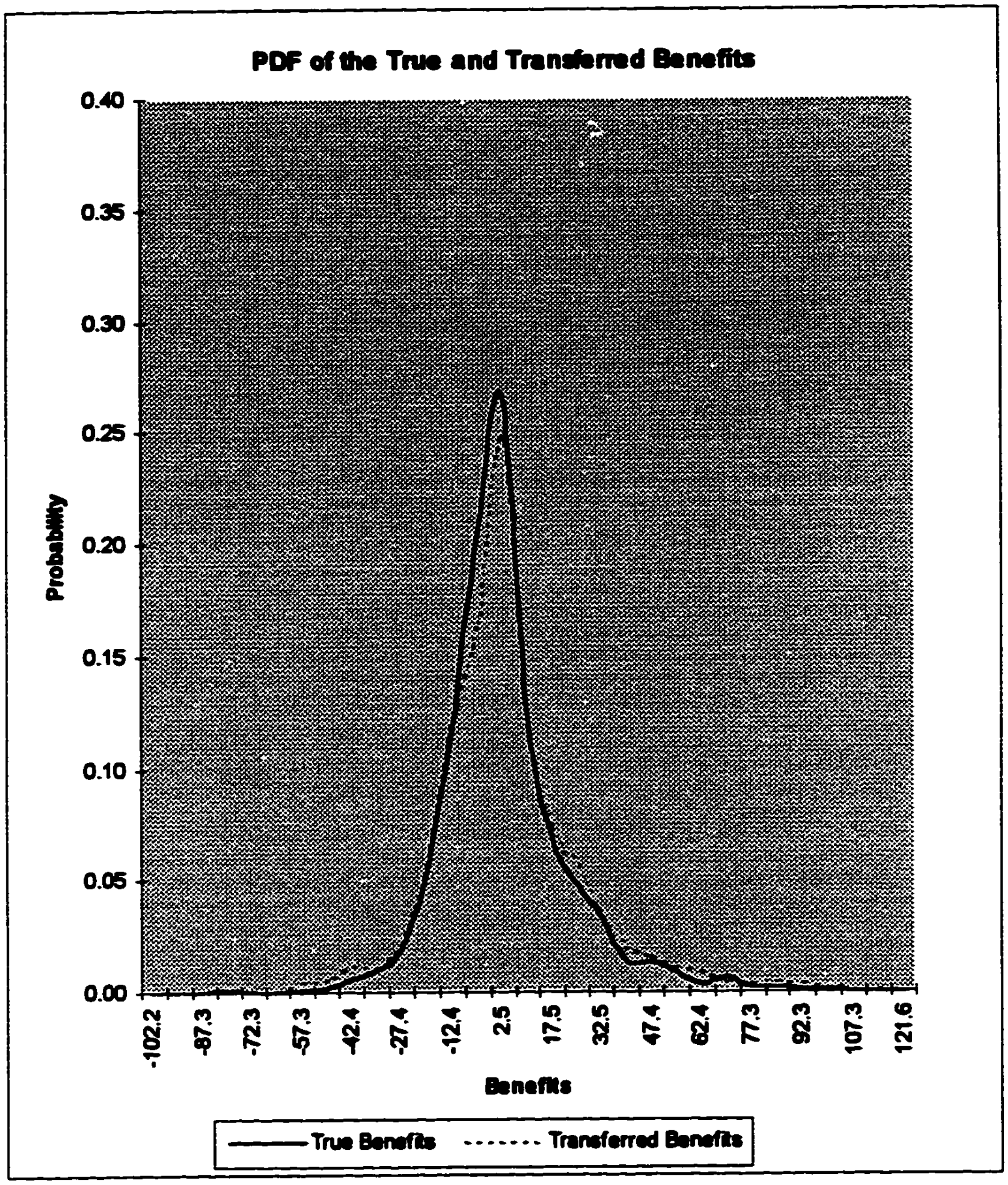


Figure 9. PDF of the True and Transferred Benefits for Policy Scenario 7.

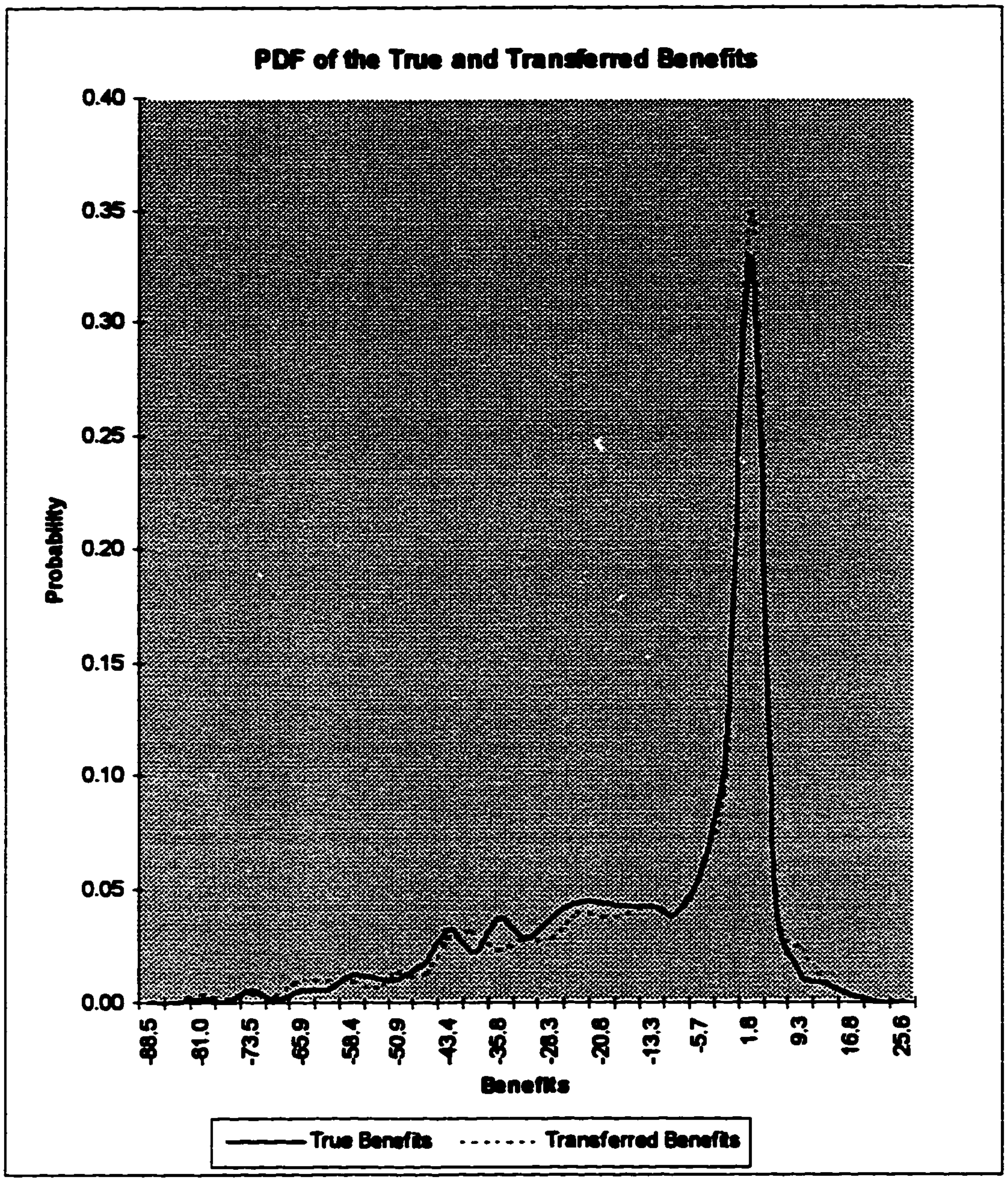


Figure 10. PDF of the True and Transferred Benefits for Policy Scenario 8.

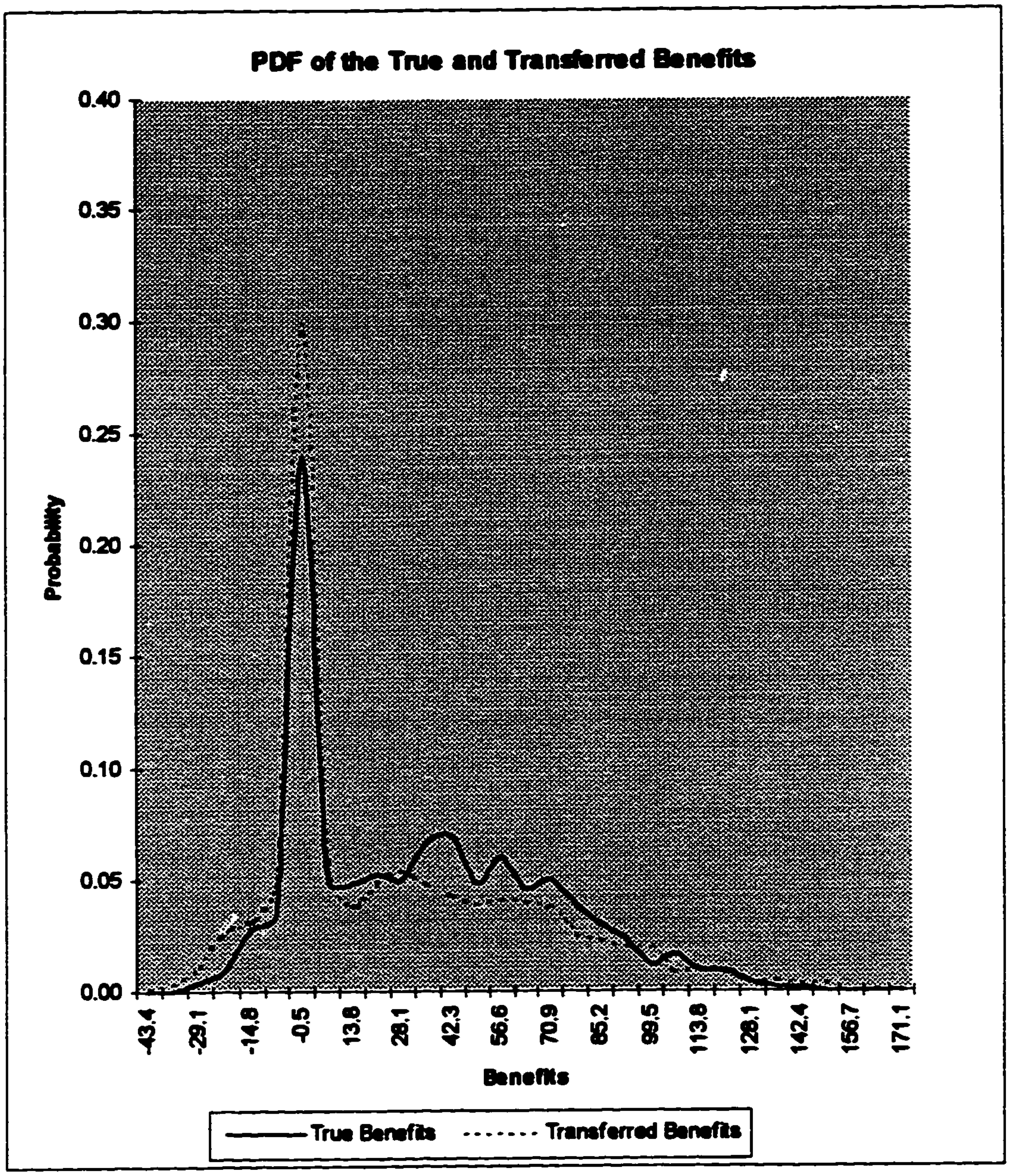


Figure 11. Probability of Making an Incorrect Decision under Different Market Costs (Policy Scenario 8)

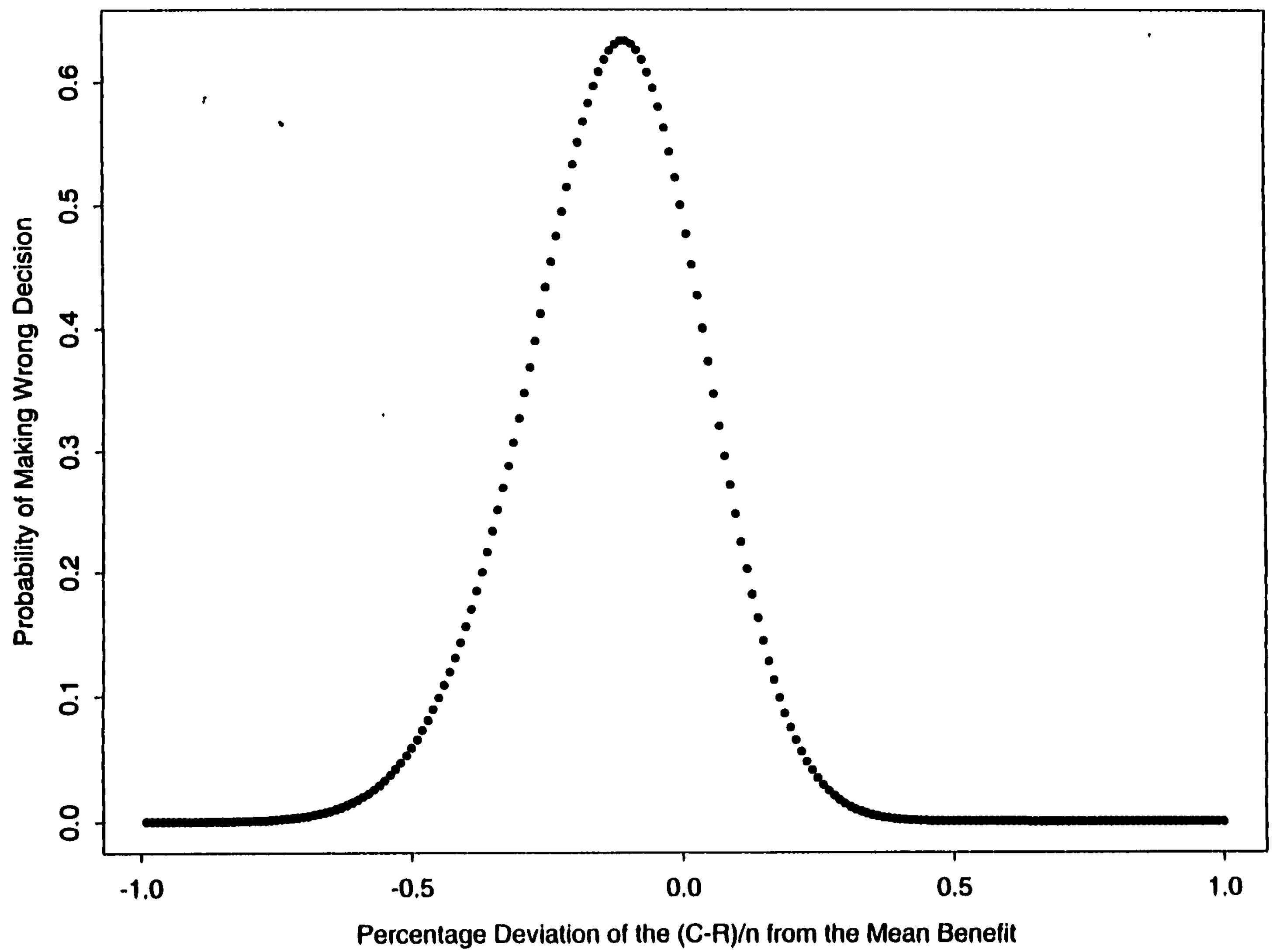
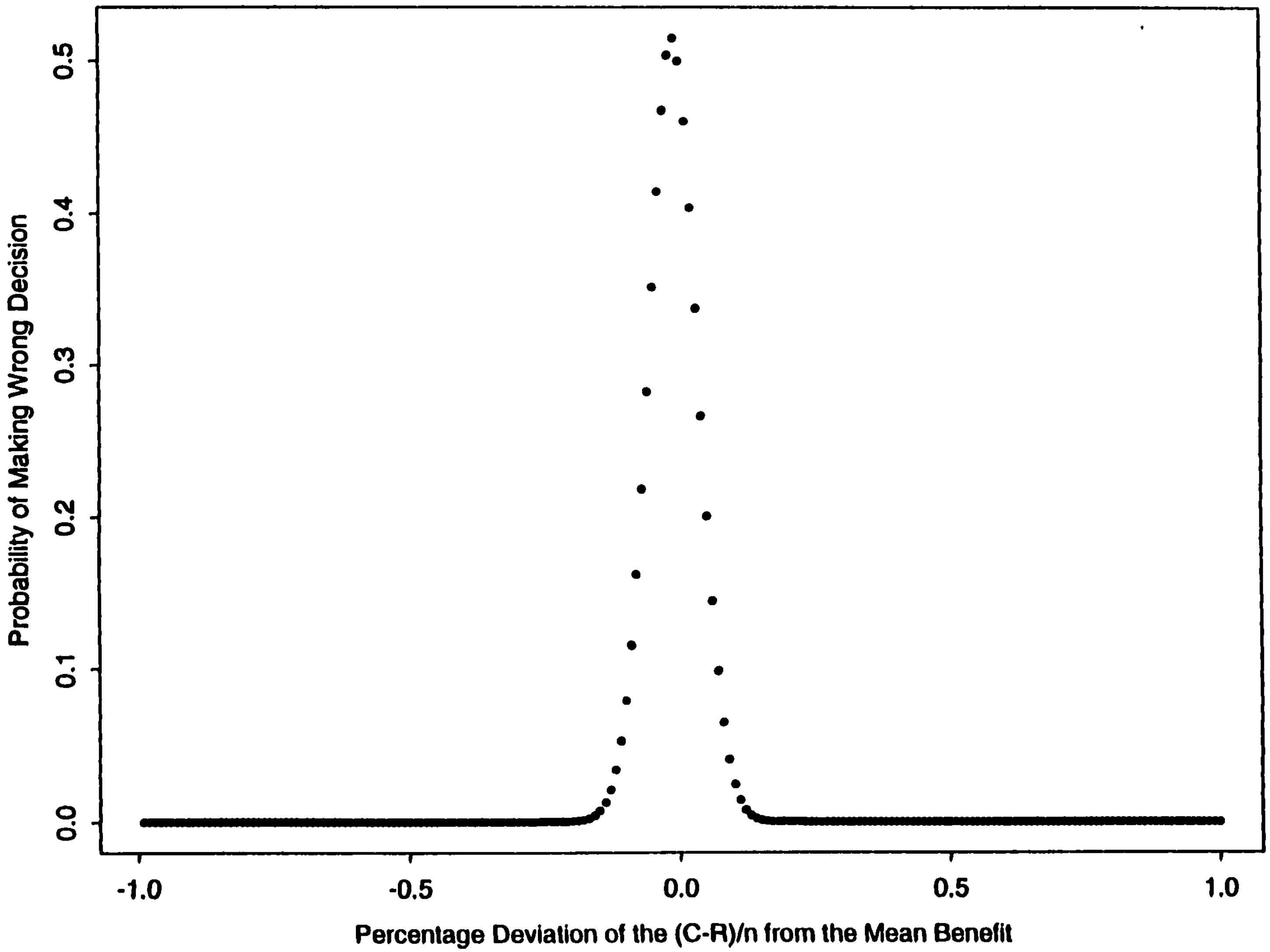


Figure 12. Probability of Making an Incorrect Decision under Different Market Costs (Policy Scenario 5)



References

Abdelwahab, W.M. "Transferability of Intercity Disaggregate Mode Choice Models in Canada." 1991. *Can. J. Civ. Eng.* 18, 20-26.

Adamowicz, W.L., J. Swait, P. Boxall, J. Louviere, and M. Williams. "Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation." *Journal of Environmental Economics and Management*.32. 1997.

Adamowicz, W.L. "Habit Formation and Variety Seeking in a Discrete Choice Model of Recreation Demand." 1994. *Journal of Agric. And Reso. Economics*. 19(1).

Adamowicz, W.L., S.Jennings ans A.G. Coyne. "A Sequential Choice Model of Recreational Behaviour." *Western Journal of Agric. Economics*. 1990. 15:91-99.

Adamowicz, W.L., P.C. Boxall, J.J.Louviere, J.Swait and M.Williams. "Stated Preference Methods for Environmental Valuation."1994. *Staff Paper 94-07*, Dept. of Rural Economy, University of Alberta.

Adamowicz, W.L., J.J.Louviere and M.Williams. "Combining Stated and Revealed Preference Methods for Valuing Environmental Amenities." 1994b. *Journal of Environmental Economics and Management*.(26):271-92.

Anderson, E.B.. "The Numerical Solution of a Set of Conditional Estimation Equations." 1972. *Journal of the Royal Statistical Society, SeriesB*, 34.

Atherton,T. and M.Ben-Akiva. "Transferability and Updating of Disaggregate Travel Demand Models." 1976. *Trans. Res. Rec.*, 610, 12-18.

Atkinson, S.E.. T.D. Crocker and J.F. Shogren. "Bayesian Exchangeability, Benefit Transfer and Research Efficiency." 1992.*Water Resources Research*, Vol. 28, No. 3.

Ben-Akiva,M. and S.Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand.* 1985. MIT Press, Cambridge, Mass.

Ben-Akiva M., D. Bolduc and K. Pene. "The combined estimator approach to model transferability and updating". *Empirical Economics*. (1995) 20:167-75.

Bockstael, N.E., W.M. Hanemann, and I. Strand. "Measuring the Benefits of Water Quality Improvements Using Recreational Demand Models." Draft Report EPA No.CR-811043-01-0. Dept. Of Agricultural and Resource Economics. University of Maryland, College Park, Maryland.

Boyle, K.J. and J.C. Bergstrom. "Benefit Transfer Studies: Myths, Pragmatism, and Idealism." 1992. *Water Resources Research*, Vol. 28, No. 3.

Braden, J.B and C.D. Kolstad. *Measuring the Demand for Environmental Quality*. Elsevier Science Publishers B.V., 1991.

Breiman, L. Statistics: With a View Toward Application. Boston: Houghton Mifflin Company. 1973.

Bromley, D.W. (ed) The Handbook of Environmental Economics. Blackwell publishers, Inc., 1995.

Brookshire, D.B. and H.R. Neill. "Benefit Transfers: Conceptual and Empirical Issues." 1992. *Water Resources Research*, Vol. 28, No. 3.

Bunch, D.S.. "Estimability in the multinomial probit model." *Transpn. Res. B.* Vol.25B, No.1, 1991.

Carpenter, G.S. and D.R. Lehmann. "A model of Marketing Mix, Brand Switching, and Competition." 1985. *Journal of Marketing Research*. 22 (Aug.)

Chamberlain, G. "Analysis of Covariance with Qualitative Data." 1980. *Review of Economic Studies*, 47. 225-238.

Chen, H.Z. and S.R. Cosslett. "Heterogenous preference of environmental quality and benefit estimation in multinomial probit models." presented at the 1996 AERE workshop, June 2-4, Tahoe City, California

Cooper, J.H. "A comparison of Approaches to Calculating Confidence Intervals for Benefit Measures from Dichotomous Choice Contingent Valuation Surveys." 1994. *Land Economics* 70(1):111-22.

Coyne,A. and W. Adamowicz. Modelling Choice of Site for Bighorn Sheep Hunting." 1992. *Wildlife Society Bulletin*. Spring.

Daganzo, C. Multinomial Probit: Theory and Its Application to Demand Forecasting. Academic Press, New York. 1979.

Desvousges, W.H., M.C. Naughton and G.R. Parsons. "Benefit Transfer: Conceptual Problems in Estimating Water Quality Benefit Using Existing Studies." 1992. *Water Resources Research*, Vol. 28, No. 3.

Downing, M. and JR. Ozuna. "Testing the Reliability of the Benefit Function Transfer Approach". *Journal of Environmental Economics and Management*. Vol. 30. No. 3, May 1996.

Feller, W. An Introduction to Probability Theory and Its Application. Value I, 2nd edition.New York:

John Wiley and Sons, 1957.

Flath,D. and E.W. Leonard. "A Comparison of Two Logit Models in the Analysis of Qualitative Marketing Data." 1979. *Journal of Marketing Research*, 22(Nov.)

Gensch,D.H. and W.W.Recker. " The Multinomial, Multiattribute Logit Choice Model." 1979. *Journal of Marketing Research*, 16(Feb.)

Gibbons, J.D. Nonparametric Statistical Inference. 2nd edition. New York and Basel: Marcel Dekker, Inc., 1971.

Gonul, F and K. Srinivasan. "Modelling Multiple Sources of Heterogeneity in Multinomial Logit Models: Methodological and Managerial Issues." 1993. *Marketing Science* Vol. 12. No. 3.

Goodman, A.C., Identifying Willingness-to-pay for Heterogeneous Goods with Factorial Survey Methods. *Journal of Environmental Economics and Management* 16, 58-79. 1989.

Gourieroux, C. and A. Monfort. "Simulation-based inference." *J. of Econometrics* 59 (1993) 5-33.

Guadagni,P.M. and J.D.C. Little. "A logit Model of Brand Choice." 1983. *Marketing Research*. 2(Summer).

Hanemann, W.M. "Applied welfare analysis with qualitative response models". Working paper No.241. 1982. University of California.

____, "Welfare Evaluation in Contingent Valuation Experiments with Discrete Response." *Amer. J. Agr. Econ.* 73 (August 1984): 332-41.

Hellerstien, D. And R. Mendelsohn. "A theoretical Foundation for Count Data Models. 1993.0 *American Journal of Agricultural Economics*. 1993. 75:604-611.

Hensher, D.A. ""Sequential and Full Information Maximum Likelihood Estimation of a Nested Logit Model." *Review of Economics and Statistics* LXVIII(4), 657-667.

Hensher, D.A. and M.Bradley. "Using Stated Response Choice Data to Enrich Revealed Preference Discrete Choice Models." 1993. *Marketing Letters* 4:2.

Hollander, M. And D.A. Wolfe. Nonparametric Statistical Methods. New York: John Wiley & Sons, Inc., 1973.

Horowitz, J. "Sampling, Specification and Data Errors in Probabilistic Discrete-Choice Models." 1981. *Applied Discrete-Choice Modelling*. D.A. Hensher and L.W.Johnson, Appendix, 217-

236, Halsted Press, New York: John Wiley and Sons.

Jones,J.M. and F.S. Zufryden. " An Approach for Assessing Demographic and Price Influences on Brand and Purchase Behaviour." 1982. *Journal of Marketing*. 36.

Jones, J.M. and J.T. Landwehr. "Removing Heterogeneity Bias from Logit Model Estimation." 1988. *Marketing Science*. Vol. 7. No. 1.

Keane M.P.. "A note on identification in the multinomial probit model." *J. of Busi. And Econ. Statis.* 1992 Vol.10 No.2.

Koppelman,F.S. and C.G. Wilmot. "Transferability Analysis of Disaggregate Choice models." *Transportation Research Record* 895.

Krinsky,I. and A.L.Robb. "On Approximating the Statistical Properties of Elasticities." 1986. *Rev. Econ. Stat.* 68, 715-19. (Correction, Rev. Econ. Stat., 72, 189-90, 1990).

Lareau, T.J. and D.A. Rae. Valuing WTP for Diesel Odor Reductions: An Application of Contingent Ranking Technique. *Southern Economic Journal*. 1988. 728-42.

Law, A.M. and W.D. Kelton. Simulation Modeling and Analysis. McGraw-Hill, Inc., 1982.

Layton D.F.. "Rank ordered, random coefficients multinomial probit models for stated preference surveys." presented at the 1996 AERE workshop, June 2-4, Tahoe City, California.

Lehmann, L.E., and H.J.M.D'abrera. Nonparametrics. San Francisco: Holden-Day, Inc., 1975.

Louviere,J.J. "Analysing Decision Making: Metric Conjoint Analysis." 1988.Sage University Papers Series No. 67. Newbury park: Sage Publication.

____. "Conjoint Analysis Modelling of Stated Preferences: A Review of Theory, Methods, Recent Developments and External Validity." 1988. *Journal of Transportation, Economics and Policy*. 10, 93-119.

____. "Experimental Choice Analysis: Introduction and Overview". *Journal of Business Research*, 23, 291-297. 1991.

____. Conjoint Analysis. In R. Bagozzi (ed.) Advances in Marketing Research. Blackwell Publishers. 1994.

Louviere,J.J. and G.G Woodworth. "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregated Data." 1983. *Journal of Marketing Research*, 20, 350-67.

Loomis, J.B.. "The Evaluation of a More Rigorous Approach to Benefit Transfer: Benefit Function Transfer." 1992. *Water Resources Research*, Vol. 28, No. 3.

Luken, R.A., F.R. Johnson and V. Kibler. "Benefit and Costs of Pulp and Paper Effluent Controls under the Clean Water Act."1992. *Water Resources Research*, Vol. 28, No. 3.

Luckert, M and W. Adamowicz. "Empirical Measures of Factors Affecting the Social Rate of Time Preference." 1993. *Env. and Resource Economics*. 3(Feb.):1-22.

Mackenzie, J. A Comparison of Contingent Preference Models, *American Journal of Agricultural Economics*. 75, 593-607. 1993.

MacNab, B and B.L. Mcfarlane. "The Benefits of the Saskatchewan Forest Habitat Study: Moose Hunting and Wildlife Viewing". 1997, Report in Progress, Canadian Forest Service, Northern Forestry Centre, Edmonton, Alberta.

Maddala,G.S. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge, England: Cambridge University Press.

McCloskey D.N. and S.T.Ziiliak. "The standard error of regressions". *J. of Economic Literature*. Vol.XXXIV (March, 1996).

McConnell, K.E. "Model Building and Judgment: Implications for Benefit Transfers with Travel Cost Models." 1992. *Water Resources Research*, Vol. 28, No. 3.

McCoomb,L.A. "Analysis of the Transferability of Disaggregate Demand Models among Ten Canadian Cities." 1983. *Transportation Forum* 3(1): 19-32.

McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behaviour." *Frontiers of Econometrics*, P. Zarembka (Ed.), New York: Academic Press, 105-142.

______. "Econometric Models of Probabilistic Choice" in *Structure Analysis of Discrete data*, eds. C.F.Manski and D.McFadden, 1981, pp. 192-272. Cambridge, MA:MIT Press.

______. "Econometric Analysis of Qualitative Response Models." in *Handbook of Econometrics*, eds. Z.Griliches and M.D. Intriligator, vol.2, chapter 24. Pp. 1395-457. Amsterdam: North-Holland Publishing Company.

Mcleod,K., P.C. Boxall, W.L. Adamowicz, M.Williams and J.J.Louviere. "The Incorporation of Nontimber Goods and Services in Integrated Resource Management. I. An Introduction to the Alberta Moose Hunting Study." 1993. *Project Report 93-12*, Dept. of Rural Economy, University of Alberta.

Meyer, R and E.J. Johnson. "Empirical Generalizations in the Modeling of Consumer Choice." *Marketing Science*. Vol. 14. No. 3. part 2of 2. 1995.

Morey,E.R., R.D. Rowe and M.Watson. "A Repeated Nested-Logit Model of Atlantic Salmon Fishing." 1993. *American Journal of Agricultural Economics*. 75:578-592.

Morton,K., W.L. Adamowicz and P.C. Boxall. "Economic Effects of Environmental Quality Change on Recreational Hunting in Northern Saskatchewan." 1994. *Project Report 94-02*, Dept. of Rural Economy, University of Alberta.

Mood, A.M., F. A. Graybill and D.C. Boes. Introduction to the Theory of Statistics. 3rd edition. San Fransico: McGraw Hill, 1974.

Noether, G.E. Elements of Nonparametric Statistics. New York: John Wiley & Sons, Inc., 1967.

Opaluch, J.J., S. Swallow, T. Weaver, C. Wessels, and D. Wichlens. Evaluating Impacts from Noxious Waste Facilities: Including Public Preferences in Current Mechanisms. *Journal of Environmental Economics and Management*. 24. 41-59. 1993.

Parsons,G.R and M.J.Kealy. "Benefits Transfer in a Random Utility Model of Recreation." 1994. *Water Resource Research*, Vol.30, No. 8, 2477-84.

Poe, G.L., E.K. Severance-Lossin and M.P. Welsh. "Measuring the Difference (X-Y) of Simulated Distributions: A Convolutions Approach." 1994. *Amer. J. Agr. Econ*. 76(Nov.)

Poe, G.L., M.P. Welsh, and E.K. Lossin. "Implementation of the Covolutions Approach to Measuring Differences in Simulated Distributions." Cornell University Agricultural Resource, and Managerial Economics Working Paper 94-05.

Punj, G.N. and R. Staelin. "The Choice process for Graduate Business Schools." 1978. *Journal of Marketing Research*, 15(Nov.)

Rae, D.A. The Value to Visitors of Improving Visibility at Mesa Verde and Great Smokey National Parks, in Rowe, R.D. and L.G. Chestnut (eds), Managing Air Quality and Scenic Resources at National Parks and Wilderness Areas. Boulder, CO, Westview Press. 1983.

Smith, V.K. "On Separating Defensible Benefit Transfers from 'Smoke and Mirrors'." 1992. *Water Resources Research*, Vol. 28, No. 3.

Smith, V.K. and Y. Kaoru. "Singles or Noise? Explaining the Variation in Recreationa Benefit Estimates". *American Journal of Agric. Econ.*, 72(2). 1990.

Stynes,D.J. and G.L. Peterson. "A review of Logit Models with Implications for Modelling Recreation Choices." 1984. *J. Leisure Research*. 16:295-310.

Swait,J. and J.J.Louviere. "The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models." 1993. *Journal of Marketing Research*, Aug. 305-14.

Takayuki Morikawa. " Correcting State Dependence and Serial Correlation in the RP/SP Combined Estimation Method." 1994. *Transportation* 21: 153-65.

Talvitie,A. And D. Kirshner. "Specification, Transferability, and the Effect of Data Outliers in Modelling the Choice of Mode in Urban Travel." 1978. *Transportation* 7:311-331.

Viscusi, W.K., W.A. Megat, and J. Huber. Pricing Environmental Health Risks: Survey Assessments of Risk-Risk and Risk-Dollar Tradeoffs for Chronic Bronchitis. *Journal of Environmental Economics and Management*, 21, 32-51, 1991.

Xu, Xiaosong and W.L. Adamowicz.1996. Assessing the Transferability of Stated Preference Environmental Valuation Models. Selected contributed paper presented at the Agricultural Institute of Canada Annual Conference, Lethbridge, July 7-11.

Yen, S.T. and W.L. Adamowicz. "Participation, Trip Frequency and Site Choice: A Multinomial-Possion Hurdle Model of Recreation Demand." 1994. *Cdn J. of Agricultural Economics*. 42 March.