Testing Criterion Validity of Benefit Transfer Using Simulated Data

Nilam Prasai

Thesis submitted to the faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Agricultural and Applied Economics

Chris Parmeter, Chair
Kevin J. Boyle, Co-chair
Jaren Pope
Nicolai Kuminoff

August 5, 2008

Blacksburg, Virginia

Keywords
Criterion Validity, Benefit Transfer, Discrete Choice Modeling, Scale Parameter, Substitution Possibilities, Characteristics Difference, Preference Difference, Measurement Error, Random Parameter

Testing Criterion Validity of Benefit Transfer Using Simulated Data

Nilam Prasai

Abstract

The purpose of this thesis is to investigate how the differences between the study and policy sites impact the performance of benefit function transfer. For this purpose, simulated data are created where all information necessary to conduct the benefit function transfer is available. We consider the six cases of difference between the study and policy sites- scale parameter, substitution possibilities, observable characteristics, population preferences, measurement error in variables, and a case of preference heterogeneity at the study site and fixed preferences at the policy site. These cases of difference were considered one at time and their impact on quality of transfer is investigated. RUM model based on reveled preference was used for this analysis. Function estimated at the study site is transferred to the policy site and willingness to pay for five different cases of policy changes are calculated at the study site. The willingness to pay so calculated is compared with true willingness to pay to evaluate the performance of benefit function transfer. When the study and policy site are different only in terms of scale parameter, equality of estimated and true expected WTP is not rejected for 89.7% or more when the sample size is 1000. Similarly, equality of estimated preference coefficients and true preference coefficients is not rejected for 88.8% or more. In this study, we find that benefit transfer performs better only in one direction. When the function is estimated at lower scale and transferred to the policy site with higher scale, the transfer error is less in magnitude than those which are estimated at higher scale and transferred to the policy site with lower scale. This study also finds that transfer error is less when the function from the study site having more site substitutes is transferred to the policy site having less site substitutes whenever there is difference in site substitution possibilities. Transfer error is magnified when measurement error is involved in any of the variables. This study do not suggest function transfer whenever the study site's model is missing one of the important variable at the policy site or whenever the data on variables included in study site's model is not available at the policy site for benefit transfer application. This study also suggests the use of large representative sample with sufficient variation to minimize transfer error in benefit transfer.

Acknowledgements

This work would not have been possible without the support and encouragement of my committee chair, Dr. Chris Parmeter. He was always there to meet and discuss about my research, guide me through the courses that I need for my research, marking up my thesis chapters, and to ask me good questions to help me think through problems. Throughout my thesis-writing period, he provided sound advice, his research expertise and insight, and technical knowhow of writing. I would also like to express my deep gratitude to my co-chair, Dr. Kevin J. Boyle for his valuable suggestions and moral support.

Dr. Nick Kuminoff has also been abundantly helpful, and has helped me in numerous ways, including helping me to understand the basics of matlab programming, coding for my research, analyzing and interpreting the results, and in particular for providing the direction to think about the problem. My thanks are due to Dr. Jaren Pope for his guidance in writing and for his comments and suggestions to improve my thesis.

I am also indebted to my many student colleagues for providing a stimulating and fun environment to learn and grow. I am especially grateful to Aissatou, Annie, Vida, Robert, Joe, Alex, Mike, Jitendra, Nabin, Achala, Ananta and Ranju. I am also thankful to the Department of Agricultural and Applied Economics for allowing me to pursue graduate studies and its tremendous support towards completion of my study. My sincere thanks go to Dr. Dan Taylor for his care and concern.

I wish to thank my entire family, on whose constant encouragement and love I have relied through out my time at Virginia Tech. My especial thanks to my husband, Arun KC, for his both physical and moral support to complete this study. Lastly and most importantly, I wish to thank my parents Shiva Prasad Prasai and Bhagawati Prasai. They have always supported and encouraged me to do the best in all matters of life. To them, I dedicate this thesis.

Table of Contents

List of Table and Figures

Chapter 1: Introduction

Nonmarket values of environmental goods and services at a given location are generally estimated by undertaking primary research at that location or by transferring economic estimates from previous studies at other locations. Primary studies are conducted for a specific purpose in a specific context and time in order to formulate and improve policy decisions. This style of research can be expensive and time consuming. When primary studies are not possible or feasible, decision makers and analysts transfer economic estimates or functions from pre-existing research studies to the new location to make informed decisions. This process is broadly known as 'benefit transfer'.

Sites or locations where primary research is undertaken are commonly referred to as 'study sites' while the new sites or locations where the benefit transfers are performed are termed the 'policy sites'. A commonly accepted definition of a benefit transfer is "The transfer of existing estimates of nonmarket values to a new study which is different from the study for which the values were originally estimated" (Boyle and Bergstrom, 1992, p. 657). In other words, this is an application of estimated values and other information from a study site to a policy site with little information or no information (Brookshire and Neill, 1992; Rosenberger and Loomis, 2000). This approach has been used extensively to estimate the benefits of environmental goods such as recreational use of national parks and coastal areas, water quality improvement, land use, and coastal land management.

1.1 Conceptual Framework on Benefit Transfer

The concept of benefit transfer came into practice during the early 1980s in an effort to reduce the cost and time of full fledged studies. Freeman (1984) started the formal process of

evaluating benefit transfers. The workshop co-sponsored by the Association of Environmental and Resource Economists (AERE) and the U.S. Environmental Protection Agency (USEPA) in 1992 focused on the issues of benefit transfer in environmental and resource management (Environment Protection Agency[EPA], 1993). The Journal of Water Resources Research (Vol. 28, issue 3), published in the same year, was mainly concentrated on the concepts, methodologies, and major issues pertaining to benefit transfer. Since then, there has been steady growth in the literature on testing the validity[1] of benefit transfer as well as the development of new methods for conducting benefit transfer.

Boyle and Bergstrom (1992) proposed a systematic conceptual foundation for conducting benefit transfer studies as follows:

- ". . . specify the theoretical definition of the value(s) to be estimated at the policy site." (p. 658)
- Study sites are identified by "conducting a thorough literature search." (p. 659)
- "Potential study site values must be examined to determine whether they are transferable." (p.659)
- ". . . evaluate the quality of these estimates in terms of their original quality." (p. 660)
- ". . . requires, or at least be improved by, simultaneous collection of primary data." (p. 661)

In the same journal, Desvousges, Naughton and Parsons (1992) also specified several conditions for valid benefit transfer.

- Original valuation studies should be carefully evaluated based on quality of data, economic methods and empirical techniques applied for the study.

[1] Validity in benefit transfer literature generally refers to external validity (extent of generalizability of the results to other locations and time).

- Changes in resources should be similar at the study and policy sites.
- Both the study and policy sites should be similar. In other word, the study site model should include variables that measures site characteristics.
- Socioeconomic characteristics of the population such as income, age, and education should be similar at the study and policy sites. Alternatively, the model at the study site should be described as a function of socioeconomic characteristics of the population at that site.
- The market at the study and policy sites should be similar. There should be sufficient information about number and types of substitutes and their prices at the study site.

Similar kinds of framework proposed by other researchers (Bergstrom and De Civita, 1999; EPA, 2000; Smith and Pattanayak, 2002) are not discussed here. These frameworks provide clear guidelines and technical criteria for valid benefit transfers. The difficulty in conducting benefit transfer within these frameworks is that the conditions specified by them are not fulfilled in most cases. Usually, some variations exist between the previous estimation and the current transfer sites unless the site under consideration is the same site. These frameworks do not specify the conditions for conducting benefit transfer when the above listed criteria are not met.

1.2 Benefit Transfer Approaches/Methods

There is no single or universally adopted classification and approach to conduct a benefit transfer. The simplest approach of transferring benefits from one site to another site is to apply the unit value estimate at the study site to the policy site. For example, one could use values per day for recreational sport fishing derived at a lake to value the sport fishing at another lake. Bergland, Magnussen, and Navrud (2002) identify three different approaches to benefit transfer:

1. Unadjusted unit values
2. Adjusted unit values
3. Valuation functions

Groothuis (2005) classified benefit transfer into four different approaches:

1. Benefit estimate transfer
2. Benefit function transfer
3. Meta-analysis
4. Preference calibration

Following Rosenberger and Loomis (2000) and Navrud and Ready (2007), benefit transfer is broadly categorized into two groups. Figure 1 shows the classification of benefit transfer.

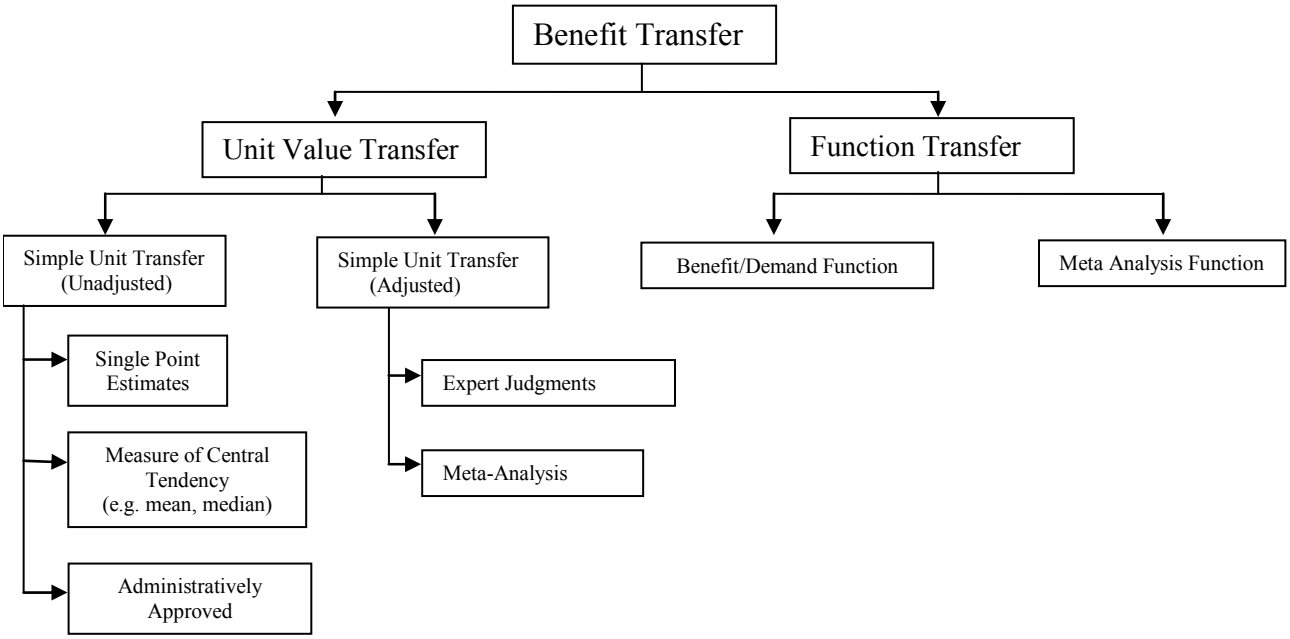

Figure 1. Benefit Transfer Methods

1.2.1 Unit Value Transfer

"Unit value transfer is the direct application of original research summary statistics (such as per unit measures of willingness to pay (WTP), measures of elasticity, or other measures of marginal effects) to a policy site" (Rosenberger and Loomis, 2003, p.448). This implicitly assumes that the marginal value to an average individual at the study site is equivalent to the marginal value to an average individual at the policy site. This approach has been used since long time to calculate benefit estimates for different resources such as whitewater rafting and bird watching (Kirchhoff, Colby, and LaFrance, 1997), wetlands (Morrison and Bennett, 2000), water quality (Bergland et al., 2002), wetlands (Morrison and Bennett, 2000), and air quality (Chattopadhyay, 2003) . There are different approaches to conduct simple unadjusted transfers. The easiest approach would be a transfer of a simple point or a range of point estimates. Another approach involves transferring similar measures of central tendency from several studies. One more approach that is widely used by organizations and administrations includes transfer of estimates derived form empirical evidence, expert judgments, and political screening. These estimates are known as administratively approved estimates. Despite the simplicity of this approach, the assumption of identical unit values across the study and policy sites may not hold in practice (Bateman, Jones, Nishikawa, and Brouwer, 2000) due to differences in socioeconomic conditions, level of changes in environmental qualities, and other factors between the sites. Even if preferences are the same between study and policy sites, opportunities in terms of substitution and activities may not be the same.

The simple unit value transfer can be adjusted to account for any possible differences in resource, population, and site characteristics between the study and policy sites. Adjustment can be made based on socioeconomic and resource characteristics data pertaining to the policy site.

Bateman et al. (2000) list three adjustment strategies: Expert judgment, re-analysis of existing study samples to identify sub-samples of data suitable for transfer, and Meta analysis of any number of previous estimates. All of these methods can be used to adjust the unit value according to the policy context.

1.2.2 Function Transfer

In the real world, it is very difficult to find the transfer sites or previous studies that meet all the criteria proposed by Boyle and Bergstrom (1992) and Desvousges et al. (1992). The study site will differ from the policy site in at least one criterion. Even if two sites are identical with respect to site, resource and user characteristics, the benefit estimates may differ if the distribution of population and their characteristics is different between the sites (Loomis, 1992). Furthermore, there exist problems with non-market valuation techniques and data for primary policy analysis which may give rise to bias in original estimates (Brookshire and Neil, 1992). The problems associated with non market valuation techniques will be exacerbated in a unit value transfer. In the quest to find a more robust benefit transfer methods, Loomis (1992) suggests transfer of the entire valuation function or the statistical model estimated at the study site to the policy site as a more rigorous benefit transfer approach. This is broadly known as 'function transfer'.

Function transfer uses more information about the differences between the policy and study sites than unit value transfer does (Navrud and Ready, 2007). For example, a study may have estimated a value function for household i, as $WTP_i = a + bX + cY + e$; where X = site/good characteristics, Y = respondent characteristics (age, income, ethnicity, etc.) and e = statistical error. Then, parameters estimated at the study site (a, b and c) are combined with the exogenous variables' values at the policy site to estimate the corresponding WTP.

Instead of transferring estimated functions from one study, the results from several studies can be combined to estimate a single function. This statistical technique of reviewing, summarizing and integrating previous studies is commonly known as Meta analysis. Since Meta analysis is a quantitative method of synthesizing results of existing studies on similar issues (Bal and Nijkamp, 2000), they are considered a very useful and refined approach to benefit transfer. This approach analyzes the effects of site and population characteristics, methodologies, and modeling assumptions on previous studies (Woodward and Wui, 2001; Rosenberger and Phipps, 2007; Navrud and Ready, 2007) and account for the effect of these factors in benefit transfer exercises.

1.3 Performing a Benefit Transfer

This section discusses the process of conducting a benefit transfer study following the stages and framework outlined in Figure 2.

1. Define the values to be estimated at the policy site

   The first step in a benefit transfer requires the definition of the values to be estimated at the policy site. This in turn requires the identification of resource, characteristics of resource for which the policy change is being considered, and the affected population by proposed policy change. The clear description of the baseline conditions of the resource, magnitude and direction of changes in the resource quality is needed in order to assess the prospect of an original valuation study. For example, the policy case can be defined as the value of sport fishing at Lake Michigan, while the policy issue can be how an increase in price of fishing accessories in the state of Michigan impacts sport fishing.

2. Conduct a literature review to identify relevant valuation data

A researcher needs to conduct a thorough literature search to identify valuation data related to a specific environmental goods and services identified in the step 1. For example, studies which have estimated WTP for sport fishing are collected. Not all identified studies are suitable for the benefit transfer process. Adequacy of original studies to the policy context can be judged based on number of factors.

- Socioeconomic and demographic characteristics of the population at the study and policy sites.
- Similarities in the characteristics of the resources and the scale of changes.
- Similarities of property right assignment and market conditions at the study and policy sites.
- Scientific and rigorous original studies (including sound econometric estimation technique, sample size, strength of explaining variation of the source, validity test).
- Age of the study.

In the above example, analysts should sort from the identified literature which measures WTP for sports fishing based on the number of substitutes available near the Michigan Lake, average household income, age, and education of the recreationists at the Lake Michigan.

3. Assess the quality of data at study sites

The quality of data used at study sites is explained in terms of:

- Scientific soundness (data collection procedures, empirical methodology).
- Richness in detail (definition of variables, treatment of site substitutes).
- Relevance to the policy site.

Since benefit transfer can only be as accurate as original studies (Brookshire et al., 1992), the quality of study sites' data plays a very important role to conduct valid transfer. Measurement errors and other uncontrolled factors in any of the initial studies will be compounded when performing benefit transfer. In the case of above example, analyst should always try to find information about sampling strategy, definition of variables, variables not included in the model, and major alternatives to sport fishing in the original study.

4. Select and summarize the data sets available from the appropriate studies

Values or data obtained from the appropriate original studies should be summarized by developing a range of parameter estimates or finding measures of central tendency such as mean and median or developing a Meta valuation function from these values.

5. Transfer benefit measures from the study to policy site

Benefit estimates or functions developed by summarizing data at the study sites are transferred to the policy sites. Benefit estimates can be adjusted or calibrated based on policy case. For example, a simple approach would be transferring the average consumer surplus for sports fishing per trip obtained from the original studies to the policy site. Another approach would be to calibrate the average consumer surplus to the policy case. Instead, the WTP function for sports fishing obtained in step 4 can be transferred in order to allow for differences in population, users, sites and resource characteristics at the study and policy sites. Meta valuation function obtained from the quantitative synthesis of several studies on sports fishing can be also transferred.

6. Determining markets over which benefit estimates are aggregated

Market is determined by geographic extent (domain), affected households within the geographic domain, and availability of substitutes (Desvousges et al., 1998). The extent of

geographical boundaries can be fixed depending upon the extent of policy effects. Another possibility is to define geographical boundaries to the extent where WTP for policy change is zero. Substitution possibilities and affected household should also be accounted while estimating aggregated benefits/costs.

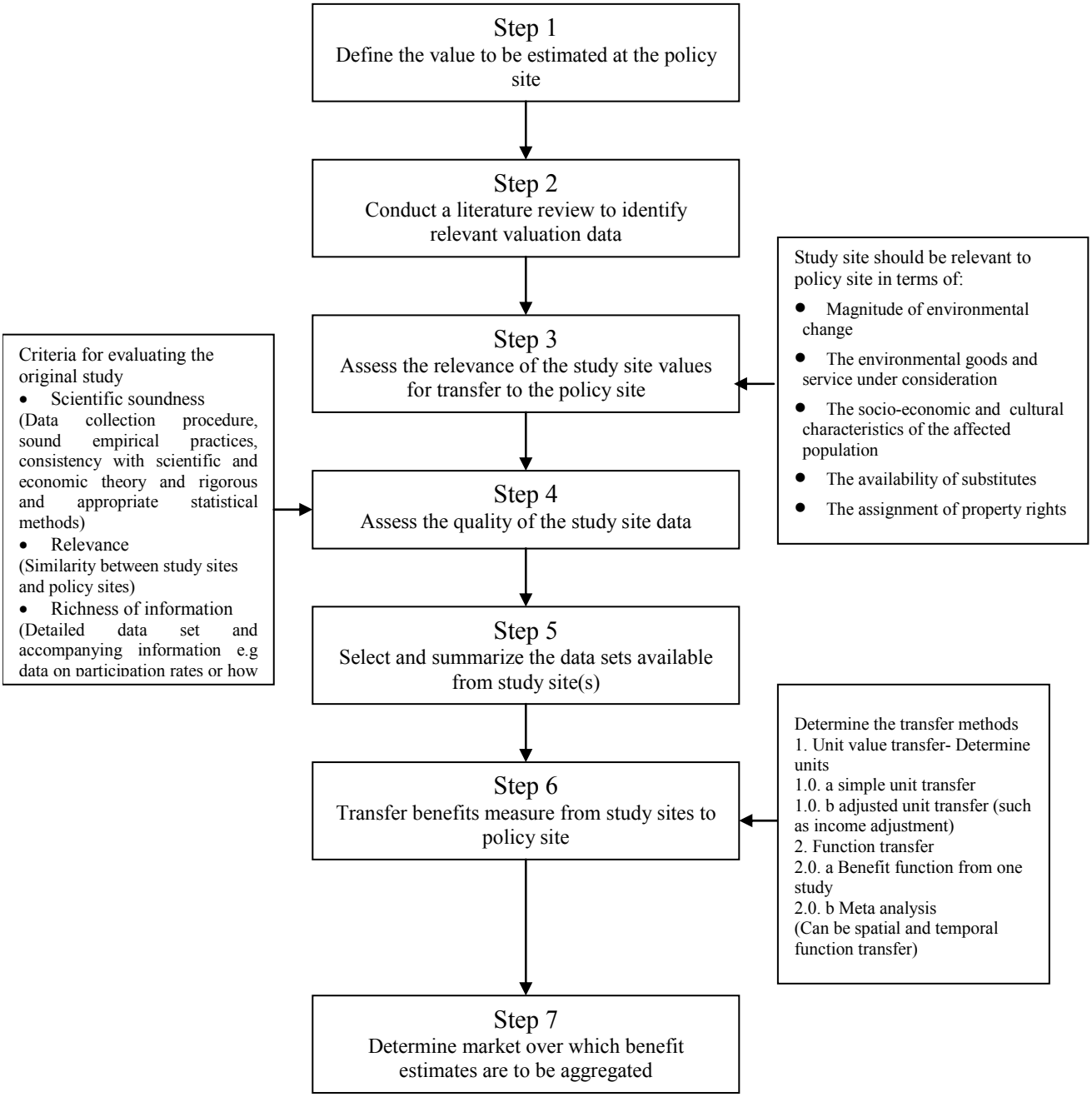

Figure 2: Framework of Benefit Transfer

1.4 Introduction to the research

Establishing the criteria of an ideal benefit transfer framework (discussed in section 1.2) imposes key challenges in performing a real world benefit transfer. Original studies that closely resemble the policy case are very difficult to find. Different valuation methods and econometric estimation techniques can lead to bias in original studies, and no original studies are likely to be free from these sources of estimation error. Benefit transfer is always associated with some magnitude of estimation error. Acknowledging this fact, an effort to reduce sources of error other than the estimation error present in original study is necessary.

Function transfer is preferred over simple unit value transfer because of the inherent ability of function to adjust with the policy context (Loomis, 1992; Kirchhoff et al., 1997; Bergstrom and De Civita, 1999; VandenBerg, Poe, and Powell, 2001; Chattopadhyay, 2003). The improvement in accuracy of function transfer is generally due to the calibration of the function with respect to differences in the socioeconomic characteristics between the study and policy sites. Most of the previous studies in benefit transfer have used the variables like household income, education level and respondents' sex; however, the variations in the study and policy sites with respect to physical characteristics are not generally calibrated to the policy context. This is mainly because physical differences- which are important for adjustments to the policy context- are not measured in the original functions (Rosenberger and Phipps, 2007). Studies having adequate variable for physical characteristics of the sites and resources become more common only after the introduction of choice modeling (CM) in nonmarket valuations. Prior to choice modeling, analysts haven't control adequately for physical characteristics. The ability of choice models to allow for both physical and socioeconomic difference between two sites make them more appropriate in the context of benefit transfer. The CM- based benefit

transfer is found to be more promising in terms of convergent validity[2] (the convergent validity evaluates whether benefit transfer estimates are consistent with known estimates for specific sites) tests of benefit transfer, and magnitude of transfer error[3] (Morrison and Bennet, 2000; Morrison, Bennett, Blamey, and Louviere, 2002; Jiang, Swallow, and McGonagle, 2005; Johnston 2007; Colombo and Hanley, 2008). However, almost all studies have implicit assumptions that:

- The study site model includes the entire important explanatory variables of the policy site.
- Data on the variables included in the study site model are easily available at the policy site.
- The same number and type of substitutes are available across the study and policy sites.
- The preferences across the study and policy sites are similar.
- There is no error on variable measurements at both the study and policy sites.

The only study which explicitly accounts for substitutes and their prices was "The evolution of a more rigorous approach to a benefit transfer: benefit function transfer", done by Loomis (1992).

This raises the concern about the performance of benefit transfer when these implicit assumptions are ruled out. In other words, how the benefit transfer results are influenced when the conditions for ideal benefit transfer do not fully hold. Rolfe (2006) points out that there are four key areas requiring attention in order to develop appropriate benefit transfer with CM data. These are site differences, framing issues (the scale and scope of the problem, number and types of substitutes), role of alternative specific constants, and population differences. This study aims

[2] Broadly, convergent validity refers to the statistical equivalence of a common parameter that is estimated using different methods. The true values of parameters are not known (Carmines and Zeller, 1979). Convergent validity in benefit transfer implies statistical equality of values/function estimated from benefit transfer exercises and the values/ functions estimated from the original study (Boyle and Bergstrom, 1992;Loomis, 1992; Bergstrom and De Civita, 1999)

[3] Transfer error is defined as percentage difference between the transferred estimates/function from the study site and the original estimates/function at the policy site.

to explore some issues- site differences, substitute differences, and population differences between the study and policy sites- as pointed by Rolfe (2006) with CM data based on revealed preference technique and goes beyond that. Furthermore, the majority of empirical literature has investigated only the convergent validity of benefit transfer. These studies fail to justify the validity of benefit transfer (Rosenberger and Phipps 2007). The major limitation in these studies is that they compare the transferred benefit estimates/functions with the estimates/functions derived at the policy site through original studies. These estimates or models developed at the policy site are taken as the true value for the comparison. However, they fail to acknowledge that these values are the estimates and are likely to suffer from estimation errors. At the same time, other forms of validity tests have been ignored in previous benefit transfer studies.

This thesis aims to add an additional dimension on validity tests of benefit transfer by exploring criterion validity[4]- corresponding to the differences in substitution possibilities, observable characteristics, and preferences of the population between the study and policy sites. Besides, this study also explores the consequence of measurement error-on variables at either the study site or the policy site-in benefit transfer.

The assumption of fixed preference parameter in a study believes that preferences are the same over individuals. In the benefit transfer context, this implies a stricter assumption than the other cases; we are assuming that preferences are fixed at both the study and policy sites. In real world scenario, one may find heterogeneity in preferences over individuals and across sites. The transfer of fixed preference parameters from the study site can produce transfer error when there

[4] Criterion validity refers to the statistical equivalence between estimates and true parameters that are known (Carmines and Zeller, 1979).Criterion validity in benefit transfer is the statistical equivalence between the transferred estimates/function from the study site and the true estimates/function known at the policy site. The simulated data will facilitate such analysis because the true parameters are specified.

is preference heterogeneity in a population of the policy site and vice versa. This issue is also addressed in this study.

A very few empirical studies have thoroughly assessed the validity of benefit transfer considering adequate information on both the study and policy sites. This study is conducted in a setting where all information necessary to conduct benefit transfer is available. This stylized analysis allows us to evaluate the performance of benefit transfer that is not compromised by data availability. This study uses random utility model (RUM) to investigate the performance of benefit function transfer when there is systematic differences between the study and policy sites. However, the coefficients of econometric model estimation from RUM are confounded with scale parameter[5]. Additionally this study also looks into the impact of the scale parameter in benefit transfer applications based on RUM estimates/functions.

1.5 Objectives

The general objective of this thesis is to test the criterion validity of the RUM (based on revealed preference data) function transfer under a controlled environment using simulated data. The simulated data sets are created where all information necessary to conduct benefit transfer is available. This stylized analysis would allow us to address the question how the results of benefit transfer are influence by the differences between the study and policy sites. In other words, the performance of benefit transfer can be analyzed when we drift away from the criteria under ideal benefit transfer. The sub-objectives of this broad question are the following:

1. To investigate the impact of a difference in scale parameter at the study and policy sites.

A well known property of RUM is that the estimated preference parameters are embedded with the information about the variance in the model. It implies that preference

[5] Scale parameter is a factor that scales the utility in RUM. In other word, a factor that normalizes the random component of RUM. It is inversely related with the variance of the random component in the RUM.

parameters are not identified without scale parameter. This means that RUM-function transfer carries information about preference parameter and variance of random component; in benefit transfer, we are implicitly assuming that preference parameters and the variance at the study and policy sites are equal. This sub-objective helps us in analyzing the performance of benefit transfer when this assumption is not met.

2. To investigate the effect of differences in substitution possibilities between the study and policy sites.

The information about substitute resources/sites can lower the preferences of resources/sites (Randal and Hoehn, 1996). In choice modeling, a different substitution possibility means difference in choice sets. Different choice sets produce different preference estimates. When we transfer a benefit function, we are implicitly assuming that the choice sets faced by individuals at the study and policy sites are the same. However, choice set misspecification can produce biased preference estimates (DeShazo, 2001). The sub-objective above focuses analysis on the performance of benefit transfer under different substitution possibilities at the study and policy sites.

3. To investigate the effect of difference in observable characteristics between the study and policy sites.

So far, almost all benefit transfer studies have assumed that the function at the study site represents and includes all the important explanatory variables at the policy site. This sub-objective aids in analyzing the performance of benefit transfer under the following two cases:

- The study site's model does not include all the important explanatory variables at the policy site.

- The variables that are included in the study site's model are important and present but they are not measurable at the policy site under benefit transfer case[6].

4. To investigate the effect of preference differences of the population at the study and policy sites

   Different populations may hold different preferences for the benefits. Majority of benefit transfer exercises has a maintained assumption of the similar preferences at the study and policy sites. The above sub-objective looks into the error arising from the differences in preference parameter at the study and policy sites.

5. To investigate the effect of measurement error in variables either at the study or policy sites

   Quality of an original study has been judged to be an important factor to improve the performance of benefit transfer. Measurement error in the variables is an important issue in benefit transfer. However, this issue has not been dealt with any of the studies in the benefit transfer literature. This sub-objective provides a basis to analyze the impact of measurement error in variables in benefit transfer.

6. To investigate the effect of fixed versus random parameter in benefit transfer.

   Fixed versus random parameter allows an investigation of the error arising from the preference parameters that varies over individuals either at the study or policy site.

1.6 Organization for the Rest of Thesis

Chapter 2 discusses the previous work in benefit transfer studies as they relate to this thesis. The descriptions of the econometric modeling framework, methodology, and necessary tests undertaken to investigate the criterion validity of benefit transfer are presented in Chapter 3. Chapter 4 details these tests and their results. The policy implications of the results in Chapter 4

[6] Benefit transfer is an alternative to the primary research. However, measurements of some variables (like attitude) require the survey and are inaccessible or immeasurable under benefit transfer applications.

are discussed in Chapter 5 and concluding remarks and directions for future research are provided in Chapter 6.

Chapter 2: Overview of the Literature

The development of several legal statutes requiring benefit cost analysis and natural resource damage assessment has led to extensive application of benefit transfers in recent years. The cost of indiscreet use of benefit transfer can be very high; at the same time if appropriate methodologies and frameworks are generated, benefit transfer can be a useful tool for saving time and resources. The development and refinement of benefit transfer methodology has been a strong concern of several researchers (e.g. Boyle and Bergstrom, 1992; Loomis, 1992; Morrison et al., 2002; Smith, Pattanayak, and Van Houtven, 2006; Johnston, 2006). The current stock of literature in benefit transfer can be broadly arranged into the following categories.

1. Unit value transfer verses function transfer
2. Stated preference technique- based studies (contingent valuation model, choice experiments) versus revealed preference technique- based studies (travel cost model, hedonic pricing model) for benefit transfer
3. Methodological development for improving benefit transfer such as Meta analysis, Bayesian analysis, Preference calibration or structural benefit transfer
4. Model specification (e.g. inclusion and exclusion of variables or Random parameter logit verses conditional logit) and their influences in benefit transfer

The current study falls into fourth category. The purpose of this chapter is to review the state of the literature on benefit transfer as it relates to this study. This review is not intended to be an exhaustive critique of the previous literature, but a synthesis of the work that informs and influences this thesis.

2.1 What do we know about the validity of benefit transfer and transfer error?

Empirical tests in benefit transfer literature can be broadly identified into two general categories: convergent validity tests and value surface tests (Boyle and Bergstrom, 1992; Bergstrom and De Civita, 1999; Chattopadhyay, 2003). A substantial volume of studies perform convergent validity tests of benefit transfer using different methods such as simple unit value transfer, function transfer, and Meta analysis (Loomis, 1992; Downing and Ozuna, 1996; Kirchhoff et.al, 1997; Rosenberger and Loomis, 2000; VandenBerg et al., 2001; Bergland et al., 2002; Shrestha and Loomis, 2002; Jiang et al., 2005; Johnston, 2007; Colombo and Hanley, 2008). In general, these studies employ tests of convergent validity in two ways: difference in means and difference in model coefficients. The former method tests for statistical difference in mean values of the parameters of interest (such as WTP or consumer surplus) between the study and policy sites. The latter determines whether there are any statistical differences between estimates at the study and policy sites (Bergstrom and De Civita, 1999). Very few studies, more specifically those involving choice modeling in the context of benefit transfer, also include comparison of correlation coefficients and implicit prices[7] between transfer and direct estimates (Morrison et al., 2002; Jiang et al., 2005; Johnston, 2007; Colombo and Hanley, 2008).

During the early development phase of benefit transfer, researchers have tested the feasibility of benefit transfer based on test of difference in model coefficients only (Loomis 1992; Parsons and Kelly, 1994; Loomis, Roach, Ward, and Ready, 1995). The general belief in these studies was if the estimated coefficients of the two-benefit functions were statistically equivalent, then the benefit estimates of the study and policy sites would also be equivalent. However, testing the equality of the models is not sufficient due to nonlinearity of the logit

[7] Implicit prices are point estimates of the value of a unit change in a nonmonetary attributes. It is calculated by taking the ratio of coefficients on non-monetary attribute variable and the coefficients on monetary cost variable.

models used to estimate the benefit functions or values themselves (Downing and Ozuna, 1995). Studies that came after Downing and Ozuna (1995) generally applied both tests-difference in mean benefit estimates and difference in model coefficients- to confirm the convergent validity of benefit estimates and functions transfer. The evidences from most of these studies fail to support the convergent validity of benefit transfer with few exceptions. Among the current pool of literature on convergent validity tests, some studies e.g. Kirchhoff et al., 1997; VandenBerg et al., 2001; Bergland et al., 2002, Jiang et al., 2005, and Johnston, 2007, tested the convergent validity in quasi-controlled conditions. Quasi-control, here, refers to the technique of concurrent estimation of benefits and function by adopting the same questionnaire formats/survey, commodity definitions, population/groups policy issues and valuation methods. As such, these studies approximate the ideal condition for assessing the validity and reliability[8] of the benefit transfer. The convergent validity of benefit transfer is not sufficiently supported by these studies too.

Kirchhoff et al. (1997) conducted the convergent validity tests with two recreational sites located at New Mexico and Arizona, respectively. Convergent validity is rejected by 90% of the tests conducted for New Mexico data set. Similarly, 55% of the tests conducted for the Arizona data set reject convergent validity. In a study by VandenBerg et al. (2001), the model equalities between the study and policy sites are rejected 36 times out of 66 possible pair wise comparisons, and the equalities of the mean WTP are rejected 44 times out of 66 possible pair wise comparisons. Johnston (2007) finds comparatively more evidence in support of validity of benefit transfer than the studies of Kirchhoff (1997) and VandenBerg et al. (2001).

[8] Validity and reliability are two different concepts. Validity implies that the unit values or functions estimated at the study site are statistically equivalent to those at the policy site (known through primary studies at that site). Reliability is related to the differences between estimated values obtained from either function transfer or values transfer from the study site and the estimated values (through original research) at the policy site (Navrud and Ready, 2007).

He tested convergent validity of benefit transfer for identical choice experiments in four different Rhode Island communities: Coventry, Exeter, West Greenwich, and Burrillville. The model equalities are not rejected for 4 of the 6 tests. Only six out of 24 tests reject the null hypothesis of equality of the implicit prices, while 17 out of 24 tests failed to reject the equality of compensating surplus.

In real benefit transfer practices, a very similar valuation scenario as used in quasi-controlled experimental test of convergent validity of benefit transfer is unlikely. Some differences between the study and policy sites are unavoidable. In the same way, statistical equalities of models and benefit estimates between the two sites are less likely. As some differences between the study and policy sites are obvious, conventional notion of test- with null hypothesis as there are no differences between the study and policy sites- is not an appropriate approach of testing validity. Kristofferson and Navrud (2005) suggest equivalence test in the context of benefit transfer. The equivalence test reverses the role of traditional null and alternate hypothesis and reduces the burden of proof by testing these hypotheses at predetermined level of accuracy or significance. The equivalence test allows focusing on the acceptable amount of error associated with transfer. Several researchers performing equivalence test do not find enough evidence to confirm convergent validity. Muthke and Holm- Mueller (2004) test the equivalence of benefit estimates for water quality improvements at tolerance limits of 20, 40 and 60% transfer error, at a significance level of $p < 0.05$. Out of 8 tests, all were rejected at 20%. However, they confirm the validity for two tests at 40 and 60% tolerance limits.

Upon failing to confirm the validity tests of benefit transfer, several researchers have examined the size of transfer error. Despite the rejection of null hypotheses, some studies find narrow margin of errors (5-15% in Loomis 1992; 4% in Parson and Kealy, ± 20% in Zanderson

et al. 2007). The error in unit value transfer ranges from a few percentage points to 1141%, while error in function transfer ranges from a few percentages to 7028 % (refer Table A1). Though the error range is high for function transfers, most of the studies claim function transfers perform better than unit value transfers (Loomis, 1992; Kirchhoff et al., 1997; Bergstrom and De Civita, 1999; VandenBerg et al., 2001; Piper and Martin, 2002; Rosenberger, 2006; Colombo et al., 2007; Colombo and Hanley, 2008).

A major limitation in all of these tests of validity of benefit transfer presides in an assumption that the values or functions estimated based on primary research is the true values or functions for the policy site. Any deviation from these estimates is characterized as bias. It is worth noting that the true values of a specific site are unobservable, and the primary research only approximates the true values. This implies that the "estimated values of a nonmarket resource are not known with certainty, even if they were from carefully performed original studies" (Rosenberger and Phipps, 2007). The size of transfer errors (biases) in these validity tests are inflated because the criterion value is not known with certainty (Ready and Navrud, 2006). The measured transfer error suffers from two sources of error: generalization error and estimation error (Ready and Navrud, 2006; Rosenberger and Stanley, 2006; Rosenberger and Phipps, 2007). Generalization error occurs when the study site's functions or value estimates are generalized to predict the estimates of the policy site (Rosenberger and Stanley, 2006). The estimation error arises mainly due to poorly conducted original research, such as poor research design, survey and respondent bias, and measurement errors. Furthermore, researchers do not know the true preference parameters or the benefits underlying the population to find the magnitude of bias. In order to assess the magnitude of estimation error, Ready and Navrud (2006) conducted sham transfers by generating two samples from the same data sets through the

Monte Carlo re-sampling procedure. The comparison of benefit estimates from one sample with the benefit estimates from the other sample of the same datasets over 1000 replications produced an average transfer error of 16%. Even the best transfer performance in their study suffered from an average transfer error of 16%.

Even though the previous literature in benefit transfer has primarily focused on convergent validity, the other approach of validity tests like face, construct, predictive, criterion, and divergent validity are also relevant (Green, 2004; Spash and Vatn, 2006). Face validity concerns whether the empirical measures confirm with common agreement and individual mental images concerning a particular concept. Construct validity is based on the logical relationships among variables within a system of theoretical foundations. Criterion validity is a measure of validity that is established by the use of a criterion measure. Predictive validity is a form of criterion validity and occurs when the empirical results predict on some criterion measure. Convergent validity is based on the extent to which the new empirical results correlate well with the measures of the same construct. On the other hand, divergent validity is based on the extent to which the new empirical results correlate badly with the measures of different unrelated constructs (Babbie, 2007). These different aspects of validity have not been explored much in the context of benefit transfer. This study seeks to investigate the criterion validity to corroborate the feasibility of benefit transfer.

2.2 What aspects of study affect validity of benefit transfer and size of transfer error?

The potential sources of error in benefit transfer practice can be classified into generalization error, measurement error (estimation error), and publication bias (Rosenberger and Stanley, 2006). The measurement error includes random error and many other factors that influence the results of the primary studies. The quality and robustness of the study site data,

valuation methods, survey designs, unit of measurements, researcher's judgment, and methods used in modeling and interpreting the data can affect the estimations (McConnell, 1992; Boyle and Bergstrom, 1992; Desvousges et al., 1998; Rosenberger and Loomis, 2001; Rosenberger and Stanley, 2006). In order to improve the performance of benefit transfer, it is important to characterize the effect of study methodology (e.g. study type, survey design and implementation, response rate, treatment of outliers, and econometric method) in the variation of WTP (Johnston et al., 2006). The second source of error is publication or selection bias. The publication bias refers to the selection bias made by publisher only publishing those researches that has statistically significant results (Rosenberger and Stanley, 2006). The publication bias affects the validity of benefit transfer by limiting the quantity of literature to conduct benefit transfer. The third source is generalization error. The generalization error arises when estimates or functions developed at a specific context is generalized to the policy site. The generalization errors are inversely related to the degree of correspondence between the study and the policy sites (Rosenberger and Stanley, 2006; Rosenberger and Phipps, 2007).

A majority of studies testing the convergent validity of benefit transfer supports the ideal that the greater the correspondence between the study and policy sites, the lesser the magnitude of the error (Loomis et al., 1995; Brouwer, 2000; VandenBerg et al., 2001; Piper and Martin, 2002; Chattopadhyay, 2003; Johnston, 2007). The correspondence between two sites includes similarity in terms of observable measures of the socioeconomic and market characteristics (age, gender, and income), and physical characteristics (topography, resource quantities and qualities, and other physical attributes of the respective sites).

Intrastate transfers yield less transfer error than cross-state transfers do. Furthermore, transfers made across states within the same region usually produce less transfer error than those

across states in different geographic regions (Loomis, 1992; Loomis et al., 1995, VandenBerg et al., 2001). Transfers within states or across same regions are believed to have less variation in socioeconomic, political, and cultural differences than across different states and regions. Piper and Martin (2002) find transfers from individual studies to similar sites produce less error (6-20%) than individual studies to dissimilar sites (89-149%). Researchers also claim that the function transfer can better adjust the differences between the sites and produce less error than the unit value transfer (Loomis, 1992; Parsons and Kealy, 1994; Kirchhoff et al., 1997; Brouwer and Spaninks, 1999; VandenBerg et al., 2001; Piper and Martin, 2001; Chattopadhyay, 2003; Johnston, 2007). Function transfer shows relatively greater accuracy than the unit value transfer when there are dissimilarities between the study and policy sites. When there is reasonable degree of similarities between the study and policy sites, unit value transfer can be as accurate as function transfer (Chattopadhyay, 2003).

Though majority of researches advocate site similarity for credible benefit transfer, it appears to be the necessary but not sufficient condition (Brouwer and Spaninks, 1999; Chattopadhyay, 2003). Even across seemingly similar sites, the error due to application of benefit transfer can be very large (Kirchhoff et al., 1997; Brouwer and Spaninks, 1999; Chattopadhyay 2003). Despite the study and policy sites being identical with respect to geography, time, resource, demography, and socioeconomic characteristics, the average transfer error can be as high as 30% for both marginal and non-marginal transfers (Chattopadhyay, 2003). Kirchhoff et al. (1997) find the magnitudes of error as large as 287% even when the two sites are nearly identical. Similarity across the sites influences the validity of benefit transfer at least in some dimension. However, site similarity should be defined very carefully. Improper measures of site similarity may produce misleading results and hence, worsen the transfer

performance. Context similarity is more important than geographical proximity for better transfer performance (Johnston, 2007). Rosenberger and Phipps (2007) mentions that the "gains in accuracy [of function transfers] may be more a function of the similarity of the sites than the calibration of site characteristics in the function transfers". This is because the adjustments of functions with respect to the policy site are done mostly for socioeconomic characteristics of the markets than the physical characteristics.

Several studies have tried to calibrate the socioeconomic characteristics of the markets through different variables like respondent age, sex, education level, and income of households. For example, Loomis et al. (1995) adjusted for the median age of the population to the policy site. Kirchhoff et al. (1997) included the variables like income, age of the respondents, education, and a dummy variable if the respondents were from foreign country in their Arizona model. However, they only included income, gender and a dummy variable if the respondent were from New Mexico, California, Arizona, Texas, or Colorado. The inclusion of more socioeconomic variables (age and education) in Arizona model than New Mexico model can be thought to produce less error in their Arizona bird watching model than for the rest of the models. Brouwer and Spaninks (1999) controls for the socioeconomic variables such as gender, education, age household income, respondent knowledge on the area and attitudinal variables such as general attitude (respondent's attitude to pay for environmental goods) and specific attitude (respondent's attitude to flowery ditch-sides and meadow birds). He found calibrating the function to the policy site with respect to these variables leads greater accuracy than simple unit value transfer does, although he did not include any variables that calibrate the physical differences between the two sites. Morrison and Bennet (2000) included the socioeconomic variables like income and age in their sample estimates but included them in their pooled model

of benefit transfer as interacting variables only. Even though, they didn't include these variables directly into the model, the error magnitude were relatively less than in other studies. This can be due to inclusion of several physical characteristics in their choice models which are calibrated to the policy context. Very few studies have calibrated function for physical characteristics to the policy context. Those studies which have calibrated functions for physical characteristics to the policy site have done so inadequately. The primary reason being physical characteristics that are important for calibrating across the sites are either unmeasured or excluded in the original functions (Brouwer, 2000; Rosenberger and Phipps, 2007). This is because these characteristics are either the same or fixed in individual site models or the researchers assumes that these differences are captured through price coefficients (Downing and Ozuna, 1996; Rosenberger and Phipps, 2007).

Studies based on the CM have advantage over others in terms of benefit transfer because it allows differences in physical (sites and resources) and socioeconomic characteristics between the study and policy sites (Morrison and Bennet, 2000; Jiang et al., 2005; Colombo and Hanley 2007). Transfer errors are comparatively lower for studies based on choice modeling. Morrison et al. (2002) find transfers of implicit prices are empirically more valid than transfers of model parameters and compensating surplus. Their estimates of compensating surplus lie within the range of 4 to 66% of true estimates. Jiang et al. (2005) emphasize on the importance of model specification in reducing transfer error. They find benefit transfer using choice modeling may be acceptable and in cases empirically valid depending upon the policy objectives and context. The model that includes the effects of respondent's environmental attitudes or that includes effects of those attitudes along with common socioeconomic attributes may help in reducing transfer error than those that omits variables identifying heterogeneity in the population at the policy site. In

contrary, Johnston (2007) only allowed for physical differences between the two sites. Despite adjusting for only physical differences between the sites, transfer performed comparatively better in his studies than in any other studies. Utility parameters and implicit prices and compensating surplus correspond to the policy site in most of the cases. Inclusion of preference heterogeneity among respondents reduces the transfer error (Hanley, Wright, and Alvarez-Farizo, 2006; Colombo et al., 2007). These studies find that random parameter model may help in reducing the transfer error in benefit transfer than the model with fixed parameters. The following key points can be made from the above review.

1. Almost all previous studies have maintained assumption that study sites' functions can represent and consists of the entire important explanatory variables at the policy site. However, there is overriding consideration on the performance of benefit transfer when the functions at the study site do not include all the variables or includes more variables than observed in policy sites.
2. All previous studies assume that the population preferences are the same between the study and policy sites. This raises the concern on performance of benefit transfer when the preferences are different between the two populations.
3. All previous studies assume that there is no measurement error in variables at both the study and policy sites. However, measurement error in variables has been the strong concern of benefit transfer practitioners and experts.
4. Studies which have controlled the differences in market condition between the study and policy sites have done so only in terms of socioeconomic characteristics like gender, age, income and household size. Majority of them failed to consider the effect of substitutes and their prices with few exceptions (Loomis et al. 1992; Parson and Kealy 1994). This can be

due to insufficient reporting of baseline conditions in original studies such as insignificant explanatory variables in the model, substitute possibilities and their prices, and complimentary effects-insufficient data/information. Loomis et al. (1992) use quality price index to control substitutes that is not sufficient to capture the substitute effects. The substitution possibilities in the context of choice modeling are more complex than any other specifications because it implies differences in choice sets an individual faces. These choice set differences have potential to cause the difference in estimated models. Therefore, it raises two principal questions. The first question asks how the performance of benefit transfer would be impacted when we have all the information that is required to conduct benefit transfer. The second question asks how benefit transfer would perform when there are different substitution possibilities between the study and policy sites.

5. A substantial amount of literature points out calibrated function transfer, specific to policy sites, seems to perform better by reducing transfer error. Evidences are in support of suitability of choice modeling in benefit transfer exercise because of their capacity to allow for adjusting the differences in physical (site and resource) as well as socioeconomic characteristics. However, most of these choice- models -based -benefit transfer are based on stated preference techniques and suffer from the criticism such as lower explanatory power and over or underestimated WTP (Brouwer and Spaninks, 1999). Models based on TCM and HPM are not without criticisms. Nevertheless, it enhances the interest to look into performance of choice modeling based - benefit transfer with revealed preference data.
6. Only three studies-Hanley et al., 2006; Colombo et al., 2007; Colombo and Hanley, 2008 have explored the potentiality of random parameter in benefit transfer. The two studies-by Hanley et al. (2006) and Colombo et al. (2007) compared the performance of fixed parameter

versus random parameter and found that random parameter model produces comparatively less error than fixed parameter model. Some more research in this area is deemed necessary to validate the scope of random parameter in benefit transfer.

The above listed point clearly indicates the missing information in benefit transfer literature.

2.3 Random Utility Model (RUM) in the framework of Benefit transfer

"The RUM provides the theoretical foundation for a class of empirical models based on consumer choices between alternatives" (Champ, Boyle and Brown, p.189). With the development of attribute based methods and choice experiments, application of the RUM has increased in non-market valuation of environmental goods and services. The widespread application of the RUM increases the likelihood of using them in benefit transfer in the future. There are a few studies which have used the RUM framework to estimate the functions and benefits at the study sites and tested their validity in the policy context (Parson and Kealy, 1994; Morrison et al., 2002; Jeong and Haab, 2004; Jiang et al., 2005; Hanley et al., 2006; Johnston, 2007; Colombo et al., 2007; Zanderson et al., 2007; Colombo and Hanley, 2008).

The RUM assumes the systematic and the random components (idiosyncratic error) affect the utility of an individual. The random component of the utility can be thought of as "unobserved variability", which is not observed by a researcher. It may include unobserved resources and individual attributes (unobserved heterogeneity), measurement errors, and error due to proxy or instrumental variables.

The unobserved variability is discussed in detail in this section than the systematic components. This part of literature review overviews the previous studies in two sections. The first part emphasizes why it is necessary to pay attention to random component of utility. The

second part concerns more on what has been done until now to take care the unobserved variability in benefit transfer and therefore reducing the magnitude of error.

2.3.1 Why should we care about unobserved variability?

The unobserved variability in choice model is assumed to play crucial role in an individual's behavioral phenomenon. The unobserved variability is assumed to follow certain distribution such as extreme value type I distribution and generalized extreme value distribution. These distributions are characterized by location and scale parameter. The scale parameter scales the utility belonging to the systematic components and is inversely proportional to the unobserved variance. The model parameters are confounded with the scale factor and cannot be identified in any particular model (Ben- Akiva and Lerman, 1985; Train, 2003). "Parameters of two identical utility specifications estimated from different data sources with unequal variances will necessarily differ in magnitude, even if the true model parameters that generated utilities are identical in both sets"(Swait and Louviere, 1993, p.305). When a parameter is estimated from the indirect utility function (includes both systematic and random components of the utility) for a discrete choice model, the implicit variance of the random component can create trouble. In essence, if there is difference in the implicit variance of the random component across two contexts, failure to allow for these differences can distort the estimates of utility parameter. If analysts wish to use the models estimated at a context to predict and forecast the choice behavior and welfare at the other context, the prediction can be highly inaccurate, given everything else similar except the implicit variance.

Methodologically, welfare estimates calculated using only systematic component from the random utility model poses no problem and is of little concern when the random component is independently and identically distributed. However, when different variances are allowed

across sample, the variance of one sub-sample is normalized. The variance for other sample is calculated with respect to the normalized sample. In this procedure, the utility parameters are kept equal across samples. Different variances in random component can lead to different distribution in predicted welfare estimates. Hence, Louviere et al. (2002) emphasize on the need to care and control for unobserved variability in order to "develop more general models of choice or demand; improve the estimates of the parameters of the systematic component; combining data source…" (p.178).

2.3.2 Role of unobserved variance (scale parameter) in benefit transfer

Generalizing the model estimated at one location to predict the choice behavior and welfare at the other location implies the implicit assumption of equality of implicit variance in random component of the utility in the RUM. However, different analysts have found the possibility of differences in implicit variances underlying the choices across different samples and populations. The recognition of different scales (inverse of unobserved variance) can lead to different results. The maintained assumption in benefit transfer practice with the RUM framework is equality of scale parameter across two samples and populations. Parson and Kealy (1994) used the RUM framework for the first time to confirm the validity of benefit transfer. They tested only the equality of the parameters across the study and policy sites. However, studies after late 1990's have conducted the econometric tests of the equality of the scale parameters across samples (Swait and Louiviere, 1993). Many of these studies reject the equality of scale parameter across the study and policy sites. Morrison et al. (2002) tested the equivalence of preference parameter except for the differences in variance across study and policy sites by rescaling the data. Jiang et al. (2005) reject both the tests: the first test-equality of parameters and variances and second test- equality of parameters except for the variances

between the study and policy sites. Hanley et al. (2006) rejects the chow test of structural break for no difference between the study and policy sites. Colombo et al. (2007) also reject the parameter equality and confirm that the models at the study site and policy site are different even after considering the scale differences. None of these studies has adjusted the scale difference between study and policy site to predict welfare estimates at the policy site after finding the difference in variance between the two sites. Neither these studies attempts to measure the contribution of scale parameter in transfer error. Furthermore, there is a practice of pooling the data from different samples, which may have different magnitude of unobserved variance. Many of these studies have only used systematic component of utility to predict welfare estimates at the policy site and therefore to measure transfer error. However, as mentioned in section 2.3.1, the different unobserved variance may lead to different distribution of welfare estimates.

This study aims to address this gap in previous researches addressing these two issues of scale parameter in benefit transfer. First, the contribution of scale parameter in the magnitude of transfer error is analyzed by systematically allowing differences in scale across sites. This helps to address the issue: Do we really need to worry about scale parameter in benefit transfer? This study also looks at the transfer error after including the random component of utility in measuring the welfare estimates at the policy site.

In summary, this study tries to test the criterion validity of benefit transfer in a controlled environment where we have all the information about the study and policy sites needed to conduct benefit transfer. This research intends to investigate the performance of benefit transfer when there are differences (described in Table 1) between the study and the policy. All these issues are addressed through choice modeling based on RUM- revealed preference techniques. Only 9 studies have used RUM model out of 25 studies listed in Table A1. Among these studies,

only three (Parson and Kealy, 1994, Jeong and Haab, 2004, and Zanderson, 2007) have used RUM model with travel cost data.

Table 1. Differences between the study and policy sites

Case No.	*Case*	*Extent of differences*
1.	Difference in scale parameter	• Low scale at the study site vs. high scale at the policy site. • High scale at the study site vs. low scale at the policy site.
2.	Substitution possibilities (Difference in choice sets)	Difference in choice sets • More choice alternatives available to the study site population than the policy site population. • Less choice alternatives available to the study site population than the policy site.
3.	Site differences	• The policy site does not have all the key attributes of the study site • Though the policy sites actually have these key attributes of the study site, they are not accessible for benefit transfer application.
4.	Preference difference	• Different preference for some attributes between the study and policy sites.
5.	Measurement error in variables	• Either the study or the policy sites variables contain measurement error.
6.	Fixed versus random parameter	• The study site has different preferences over individual while the policy site has no preference heterogeneity.
7.	All difference	• The study and policy sites are different in scale parameter, choice set, observed variables and even in terms of populations' preference.

Chapter 3: Theoretical Model and Research Methodology

3.1 Theoretical Framework and Econometric Model

3.1.1 Discrete Choice Modeling

The discrete choice model is used to analyze the choice behavior of an individual. The RUM provides the theoretical framework to integrate choice behavior with economic valuation. In RUM model, the probability of an individual i choosing an alternative j from the choice set C containing J alternatives depends on the utility of the alternative j relative to the utility of other alternatives. Hence, an individual will choose the alternative j if and only if it provides the maximum utility.

The RUM assumes that utility is composed of the systematic and idiosyncratic components (unobserved heterogeneity). The equation below establishes the basic relationship where V_{ij} represents the measurable component of utility (systematic component) and ε_{ij} represents the idiosyncratic components. ε_{ij} controls the effect of unobserved and omitted influences on choice variable and is only known to the individual but not to the researcher. ε_{ij} is distributed Extreme Value (Type I) with mean zero and variance σ.

$$U_{ij} = V_{ij}(X_{jk}, Y_i - P_{ij} : \beta_k) + \varepsilon_{ij} \tag{3.1}$$

where, $j = 1, 2, 3\ldots, m, \ldots\ldots J \quad \forall \; j \in C$ (C is a choice set which is assumed to be the same for each individual)

The systematic component of the utility is the function of attributes of the alternatives (represented by X_{jk}), the travel cost (represented by P_{ij}), and the income of each individual (represented by Y_i).This function is characterized as:

$$V_{ij} = \left[\sum_{k=1}^{9} \beta_k X_{jk} + \beta_p (Y_i - P_{ij}) \right] \tag{3.2}$$

V_{ij} is further decomposed into the constant and the variable utility. The constant utility $\left(\sum_{k=1}^{9} \beta_k X_{jk} \right)$ represents the component of V_{ij} which does not vary among individuals. The variable utility $(\beta_p P_{ij})$ represents the component of V_{ij} which differs from individual to individual. Each of the explanatory variables included in the model 3.2 are described in Table 2. β_k is a vector of preference coefficients of attributes. β_p is a coefficient of the difference between income and travel cost.

Table 2: Description of Explanatory Variables

Variable	*Description*
P_{ij}	P_{ij} is the travel cost for each individual i to site j. P_{ij} =Opportunity cost of time+ trip cost: [1/3× (yearly income/2080] × time spent on traveling+ 0.0963 × distances to and from the site(s). Assumption: there are 2080 working days in a year and 0.0963 is the cost of gasoline ($) per mile as of 2008.
Y_i	Y_i is an annual income of individual i
x_{j1}	x_{j1} =1, if a site j has X_1 characteristic and 0 otherwise
x_{j2}	x_{j2} =1, if a site j has X_2 characteristic and 0 otherwise
x_{j3}	x_{j3} =1, if a site j has X_3 characteristic and 0 otherwise
x_{j4}	x_{j4} =1, if a site j has X_4 characteristic and 0 otherwise
x_{j5}	x_{j5} =1, if a site j has X_5 characteristic and 0 otherwise
x_{j6}	x_{j6} is a continuous characteristic X_6 for site j
x_{j7}	x_{j7} is a continuous characteristic X_7 for site j
x_{j8}	x_{j8} is a continuous characteristic X_8 for site j
x_{j9}	x_{j9} is a continuous characteristic X_9 for site j

Note: X is defined as a set of characteristics. Binary variables can be presence or absence of bathroom, pier, shower, etc. Continuous variables can be water quality, catch rate, income, travel cost, etc.

According to the random utility theory, an individual will choose an alternative *m* if and only if the utility derived from the alternative *m* is greater than the utility derived from choosing alternative *j*. Thus the probability of choosing the alternative *m* by an individual *i* is:

$$p_{im} = prob(V_{im} + \varepsilon_{im} > V_{ij} + \varepsilon_{ij}) \qquad \forall j \neq i$$

$$= prob(\varepsilon_{ij} < \varepsilon_{im} + V_{im} - V_{ij}) \qquad \forall j \neq i$$

$$= prob\left(\varepsilon_{ij} < \varepsilon_{im} + \sum_{k=1}^{9} \beta_k X_{mk} + \beta_p (Y_i - P_{im}) - \sum_{k=1}^{9} \beta_k X_{jk} + \beta_p (Y_i - P_{ij}\right) \qquad \forall j \neq i$$

$$= prob\left(\varepsilon_{ij} < \varepsilon_{im} + (\sum_{k=1}^{9} \beta_k X_{mk} - \sum_{k=1}^{9} \beta_k X_{jk}) - \beta_p (P_{im} - P_{ij})\right) \qquad \forall j \neq i \qquad (3.3)$$

Since ε_{im} is independent and identically distributed and follow a standard type I extreme-value distribution, the probability of choosing the alternative *m* by individual *i* becomes:

$$p_{im}\Big|\varepsilon_{im} = \prod_{m \neq j} e^{-e^{-\left[\varepsilon_{im} + \left(\sum_{k=1}^{9} \beta_k X_{mk} - \sum_{k=1}^{9} \beta_k X_{jk}\right) - \beta_p \left(P_{im} - P_{ij}\right)\right]}} \qquad (3.4)$$

However, researchers do not know ε_{im} so the choice probability will be the integral of $p_{im}\Big|\varepsilon_{im}$ over all the values of ε_{im} weighted by its density:

$$p_{im} = \int \left[\prod_{m \neq j} e^{-e^{-[\varepsilon_{im} + (\sum_{k=1}^{9} \beta_k X_{mk} - \sum_{k=1}^{9} \beta_k X_{jk}) - \beta_p (P_{im} - P_{ij})]}} \right] e^{-\varepsilon_{im}} e^{-e^{-\varepsilon_{im}}} \partial \varepsilon_{im} \qquad (3.5)$$

After some algebra, the logit choice probability becomes:

$$p_{im} = \frac{e^{(\sum_{k=1}^{9} \beta_k X_{mk}) - \beta_p P_{im}}}{\sum_{j=1}^{J} e^{(\sum_{k=1}^{9} \beta_k X_{jk}) - \beta_p P_{ij}}} \qquad (3.6)$$

When equation (3.6) is substituted to the likelihood function in a conditional logit model, the equation below is obtained:

$$LL(\beta) = \sum_{i=1}^{I}\sum_{m=1}^{J} y_{im}\left(\sum_{k=1}^{9}\beta_k X_{mk} - \beta_p P_{im} - \ln\sum_{j=1}^{J} e^{(\sum_{k=1}^{9}\beta_k X_{jk} - \beta_p P_{ij})}\right) \tag{3.7}$$

The estimators (β) are the values that maximizes the log-likelihood function. y_{im} is a binary dependent variable that takes on a value of 1 if the respondent selects alternative *m,* and 0 otherwise.

3.1.1.1 Scale Parameter in the logit model

The unobserved heterogeneity ε_{ij}^{*} [9] in the expression (3.1) has variance of $\sigma^2 \times \frac{\Pi^2}{6}$. Since the scale of utility is irrelevant to behavior, we can normalize the utility dividing by σ. This implies each of the coefficients is rescaled by 1/σ. σ is called the scale parameter because it scales the coefficients to reflect the variance of unobserved heterogeneity. From the standard logit expression, only the ratio β^{*} / σ can be estimated. β_k can be more accurately represented as $\beta_k \sigma$. The coefficients that are estimated using the logit expression are thus said to be confounded with the variance of unobserved heterogeneity. A larger variance in an unobserved heterogeneity leads to smaller coefficients even if the observed factors have the same effect on utility and vice-versa. The presence of the scale parameter directly affects the estimation of the model parameters making the results of different models directly incomparable. However, the scale parameter does not affect the ratio of any two coefficients since it drops out

[9] Represents idiosyncratic error component with zero mean and variance of $\sigma^2 \times \frac{\Pi^2}{6}$. ε_{ij} represents normalized idiosyncratic error component ($\varepsilon^{*}_{ij} = \sigma\varepsilon_{ij}$).

$\left(\frac{\beta^*_k / \sigma}{\beta^*_{k-1} / \sigma} = \frac{\beta^*_k}{\beta^*_{k-1}} \right)$ thus making them comparable. Estimates of WTP and marginal rate of substitution (MRS) are not affected by the scale parameter.

3.1.1.2 Welfare Measure in Discrete Choice Model

The preference coefficients obtained from the multinomial logit model can be used to estimate welfare measures such as compensating variation (CV) and WTP. How much would an individual be willing to pay to prevent a site loss or for a change in an attribute? An individual is willing to pay the amount of money represented by the difference between consumer surplus before and after the change holding utility constant. Depending on the individual's utility based on the original utility level or proposed utility level, compensating or equivalent variation can be applied to measure the WTP.

CV can be defined as the amount of money taken from income to allow the individual to remain at the same level of utility of the preferred choice before the change. Equivalent variation (EV) can be defined as the amount of money that is given to the individual to keep him at the proposed level of utility of the preferred choice after the change. This study uses compensating variation to measure WTP.

The compensating variation of the change in the attributes of one or more alternatives is implicitly defined as (Bockstael and Connell, 2007):

$$\max_{j \in C} \left[V_{ij}^{0}(Y_i - P_{ij}, X_{jk}^{0}; \beta_k) + \varepsilon_{ij} \right] = \max_{j \in C} \left[V_{ij}^{1}(Y_i - CV - P_{ij}, X_{jk}^{1}; \beta_k) + \varepsilon_{ij} \right] \tag{3.8}$$

where X_{jk}^{0} is the set of attributes of alternative *j* before any change is made in the attribute(s).

V_{ij}^{0} is the observed utility derived by an individual i from the alternative j before any change is made in the attribute(s).

X_{jk}^{1} is the set of attributes for alternative j after the proposed change in attribute(s). V_{ij}^{1} represents the observed utility of an individual i from the alternative j after the proposed change in attribute(s). CV stands for the compensating variation. Replacing V_{ij} by equation (3.2) in equation (3.8) gives:

$$\max_{j \in C}\left[\left(\sum_{k=1}^{9}\beta_k X_{jk}^{0} - \beta_p P_{ij}\right) + \varepsilon_{ij}\right] = \max_{j \in C}\left[\left(\sum_{k=1}^{9}\beta_k X_{jk}^{1} - \beta_p P_{ij} - \beta_p CV\right) + \varepsilon_{ij}\right] \tag{3.9}$$

$$CV = \left[\max_{j \in C}\left\{\left(\sum_{k=1}^{9}\beta_k X_{jk}^{1} - \beta_p P_{ij}\right) + \varepsilon_{ij}\right\} - \max\left\{\left(\sum_{k=1}^{9}\beta_k X_{jk}^{0} - \beta_p P_{ij}\right) + \varepsilon_{ij}\right\}\right] \Big/ \beta_p \tag{3.10}$$

Substituting $\beta\left(\beta = \dfrac{\beta^*}{\sigma}\right)$ in equation (3.10), for normalizing the idiosyncratic component if the variance is different than $\Pi^2/6$, gives:

$$CV = \left[\max_{j \in C}\left\{\left(\sum_{k=1}^{9}\beta_k^* X_{jk}^{1} - \beta_p^* P_{ij}\right) + \sigma\varepsilon_{ij}\right\} - \max\left\{\left(\sum_{k=1}^{9}\beta_k^* X_{jk}^{0} - \beta_p^* P_{ij}\right) + \sigma\varepsilon_{ij}\right\}\right] \Big/ \beta_p^* \tag{3.11}$$

CV in equation (3.11) is known as actual compensating variation faced by an individual. Since the idiosyncratic error component is unobserved by the researcher, the researcher tries to locate the CV that equates the expected maximum utility before the change with expected maximum utility after the change. Equation below establishes the relationship to calculate the expected CV:

$$E\left[\max_{j \in C} V_{ij}^{0}(X^{0}{}_{jk}, Y_i - P_{ij}; \beta_k) + \varepsilon_{ij}\right] = E\left[\max_{j \in C} V_{ij}^{1}(X^{1}{}_{jk}, Y_i - CV - P_{ij}; \beta_k) + \varepsilon_{ij}\right] \tag{3.12}$$

For the type-I extreme value distribution, equation (3.12) becomes

$$\ln\left(\sum_{j=1}^{J} \exp V_{ij}^{0}(X^{0}{}_{jk}, Y_i - P_{ij}; \beta_k)\right) + \overline{C} = \ln\left(\sum_{j=1}^{J} \exp V_{ij}^{1}(X_{jk}, Y_i - CV - P_{ij}; \beta_k)\right) + \overline{C} \tag{3.13}$$

where $\overline{C}$ is an unrecoverable constant. Replacing V_{ij} in the equation (3.2) to equation (3.13) gives expected CV as below:

$$E(CV) = \left[\ln\left\{\sum_{j=1}^{J} \exp\left(\sum_{k=1}^{9} \beta_k X_{jk}^{1} - \beta_p P_{ij}\right)\right\} - \ln\left\{\sum_{j=1}^{J} \exp\left(\sum_{k=1}^{9} \beta_k X_{jk}^{0} - \beta_p P_{ij}\right)\right\}\right] \Big/ \beta_p \tag{3.14}$$

$$j \in C$$

or

$$E(CV) = \left[\ln\left\{\sum_{j=1}^{J} \exp\left(\sum_{k=1}^{9} \beta_k^{*} X_{jk}^{1} - \beta_p^{*} P_{ij}\right)\right\} - \ln\left\{\sum_{j=1}^{J} \exp\left(\sum_{k=1}^{9} \beta_k^{*} X_{jk}^{0} - \beta_p^{*} P_{ij}\right)\right\}\right] \Big/ \beta_p^{*} \tag{3.15}$$

$$j \in C$$

where $E(CV)$ represents expected compensating variation. Equations (3.11) and (3.15) illustrate that CV or WTP are directly comparable across models. Estimation of CV gives the measure of actual WTP and estimation of $E(CV)$ gives the measure of expected WTP.

3.1.2 Benefit Transfer of the Discrete Choice Modeling Functions

The benefit estimates (WTP_{ij_s}) for a study site, *j* and individual, *i* can be represented as a function of the site attributes X_{jk_s} and the travel cost to this site P_{ij_s}:

$$WTP_{ij_s} = f(X_{jk_s}, P_{ij_s}; \beta_{k_s})$$

where the symbols have the same meaning as defined in section 3.1 except that the sub-subscript 's' implies the measure is related to the study site(s). CV calculated from equation (3.11) gives the measure of actual WTP. WTP_s is defined as the actual WTP estimated at the study site. $E(CV)$ calculated from equation (3.15) is the measure of expected WTP. $E(WTP_s)$ is used to represent the expected WTP for the study site.

The underlying principle of benefit transfer is that the benefit functions or estimates of study site(s) can well represent the benefit function or estimates of the policy site(s). Thus, the transfer of a benefit function can be defined as:

$$WTP_{ij_{p/s}} = f(X_{jk_p}, P_{ij_p}; \beta_{k_s})$$

where the symbols have the usual meaning as defined in section 3.1. The sub-subscript 'p/s' implies that the value is estimated at the policy site by using the functions estimated at the study site(s) and the data of the policy site. The sub-subscript 'p' means that the measures correspond to the policy site(s) and are estimated through the original research at the policy site.

CV and $E(CV)$ calculated from equations (3.11) and (3.15) give the measure of actual WTP and expected WTP respectively for the policy site(s) also. Equations (3.16) and (3.17) below illustrate WTP calculated at the policy site(s) by using the function of the study site(s).

$$WTP_{p/s} = \left[\max \left\{ \left(\sum_{k=1}^{9} \beta^*_{k_s} X^1_{jk_p} - \beta^*_{p_s} P_{ij_p} \right) + \sigma\varepsilon_{ij_p} \right\} - \max \left\{ \left(\sum_{k=1}^{9} \beta^*_{k_s} X^0_{jk_p} - \beta^*_{p_s} P_{ij_p} \right) + \sigma\varepsilon_{ij_p} \right\} \right] \Big/ \beta^*_{p_s} \qquad (3.16)$$

$$j \in C$$

$$E(WTP)_{p/s} = \left[\ln \left\{ \sum_{j=1}^{J} \exp \left(\sum_{k=1}^{9} \beta^*_{k_s} X^1_{jk_p} - \beta^*_{p_s} P_{ij_p} \right) + \sigma\varepsilon_{ij_p} \right\} - \ln \left\{ \sum_{j=1}^{J} \exp \left(\sum_{k=1}^{9} \beta^*_{k_s} X^0_{jk_p} - \beta^*_{p_s} P_{ij_p} \right) + \sigma\varepsilon_{ij_p} \right\} \right] \Big/ \beta^*_{p_s} \qquad (3.17)$$

$$j \in C$$

WTP_p and $E(WTP_p)$ stands for the WTP estimated at the policy site from the original function estimated at the policy site using equations (3.11) and (3.15) respectively.

3.1.3 Validity Tests of Benefit Transfer

Criterion validity is a measure of validity that is established by using a criterion measure. The simulated data facilitates such analysis because the true parameter values are specified. The criterion measure that we will be using in our simulation exercises is defined as the difference between the true preference parameter of the population and our estimates. Then we use the ratio of any two true preference parameters (MRS) and true WTP for the population as our criterion measure to test validity. In the context of benefit transfer, this either refers to the comparison between the WTP estimates-obtained from function transfer or from direct value transfer- and the actual estimates of WTP calculated from the known parameters at the policy site. We will investigate the criterion validity by conducting the following statistical tests.

The first involves the test of equality of the actual and transferred benefit functions. This is done by comparing the MRSs between the two (Morrison et al., 2002; Jiang et al., 2005; Johnston, 2007). The equality of MRSs gives valid benefit function transfer. Hence the relevant test hypothesis would be:

$$H_0: \frac{\beta_{k_s}}{\beta_{l_s}} = \frac{\beta_{k_{p-true}}}{\beta_{l_{p-true}}} \qquad \text{(H1)}$$

where $\beta_{k_{p-true}}$ is the vector of true preference coefficient for the policy site. $\beta_{l_{p-true}}$ is any one true preference coefficient from the vector of the true preference coefficients. The standard error for MRSs is generated by Krinsky and Robb's bootstrapping procedure. This

procedure involves randomly drawing a large number of parameter vector estimates from a multivariate normal distribution with mean and variance equal to preference vector and a variance-covariance matrix from the estimated multinomial logit model. For each draw, $\frac{\beta_k}{\beta_l}$ is computed. The calculated standard deviation is then assumed to be the standard error around the estimate (Alberni, Longo, Veronesi, 2007). The equality of estimated MRSs at the study site and true MRSs at the policy site is formally tested by Wald's test (Green, 2007).

$$W=\left(\frac{\beta_{k_s}}{\beta_{l_s}}-\frac{\beta_{k_{p-true}}}{\beta_{l_{p-true}}}\right)' Var\left(\frac{\beta_{k_s}}{\beta_{l_s}}-\frac{\beta_{k_{p-true}}}{\beta_{l_{p-true}}}\right)^{-1}\left(\frac{\beta_{k_s}}{\beta_{l_s}}-\frac{\beta_{k_{p-true}}}{\beta_{l_{p-true}}}\right)$$

where W is a Wald's score with k-1 degree of freedom at 5% level of significance.

The second involves the test of equality of the estimated mean WTP (from equation (3.16) and (3.17)) and the true WTP at the policy site (from equation (3.11) and (3.15)) can be tested with the help of t-test. The t-test assumes that the amounts of WTP are drawn from a normal distribution. Formally the test hypothesis would be:

H_0: $WTP_{p/s} = WTP_{true/p}$ (H2)

where $WTP_{true/p}$ is the mean WTP at the policy site. $WTP_{true/p}$ is calculated by using the true population parameters for the policy site. $WTP_{p/s}$ is the mean WTP at the policy site provided by the benefit transfer. The above hypothesis is tested by bootstrap t-test. The procedure of bootstrap t test is discussed in section 3.2.2.

$$\text{t-value}=\frac{\hat{WTP}_{p/s}-WTP_{true/p}}{se_{WTPp/s}}$$

3.1.4 Transfer Error

Transfer error is calculated as the absolute percentage difference between the mean WTP calculated from the function estimated at the study site and data at the policy site and the mean WTP calculated from the true function for the policy site's population.

$$\text{Transfer error}=\left|\frac{(WTP_{p/s}-WTP_{true/p})}{WTP_{true/p}}\times 100\right|$$

The transfer error is calculated for both estimated WTP and actual WTP in each replication. The average transfer error is calculated by taking the mean of transfer error in each replication. We use the absolute percentage difference between the means because benefit transfer may overestimate or underestimate the value in each iteration. Thus, it is important to caution against overestimates cancelling out underestimates when these transfer errors are aggregated.

3.2 Research Methodology

In order to address the empirical investigation on criterion validity of benefit transfer, a controlled experiment is needed in which statistical properties of data can be easily manipulated to figure out how these anomalies affect the ability of benefit transfer to predict the true values. We assume both the study and policy sites are identical in all respects except for the distribution of error terms under the best case scenario (ideal benefit transfer). Therefore, we use the same data for attributes, choice set and sample for both the study and policy sites. This implies that the estimated functions and values at the study and policy sites are the same under ideal conditions. However, the actual benefit estimates are different between the study and the policy sites (Equation 3.11). Then, we systematically alter the distribution of error terms (scale parameter) between the study site and the policy site. Similarly, choice set (substitution possibilities), attributes and preference parameter are also altered systematically to evaluate their impact on benefit transfer. We also consider a case of measurement error in the variables and a

case of preference heterogeneity at one of either the study or the policy sites They are discussed in detail in section 3.2.4.

3.2.1 Data Description

For transfer exercise, this study uses two types of data: real world and simulated data. Our real world data was first used by Legget et al. (2002) to estimate user day values on Padre Island National Seashore. The survey was conducted on Jan/Feb 2001 that includes 1501 individuals. The Padre Island Gulf Coast Data is a very rich data source. This study only uses respondent's income, distance and time taken to travel to the selected beach to calculate travel cost and a data on beach characteristics.

There are 64 beaches on the bay side and the gulf side of the barrier islands of the Texas gulf coast. Nine attributes were identified for these beaches- restrooms, showers, piers, concession stands, access to state park, access to gulf beach, lifeguard services, regular cleaning, and presence of vehicle free area. All of them are binary variables. The travel cost is an only continuous variable.

Brouwer (2002) mentions that the continuous variables can reflect the strength of people's preferences and WTP for a specific change in the level of environmental goods and services. Including only binary variables may not reveal preference strength and WTP of an individual for potential changes. Therefore, we feel it is necessary to investigate the effect of the binary and continuous variables on the estimation of preference coefficients and WTP. For this purpose, we conduct three different experiments by applying the Monte Carlo Simulation Technique, which is discussed in detail in section 3.2.2. The first experiment is defined by setting all the 9 variables as binary. The only continuous variable is travel cost. For this purpose, we use Padre Island Texas Gulf Coast data. The second experiment is defined by

setting all the 9 variables as continuous. We generate these data randomly from uniform distribution. To simulate the similar scenario as in Padre Island Gulf Coast data we generate these data for 64 sites. We use the same travel cost from Padre Island Gulf Coast data for our convenience. The third experiment is defined by setting half the variables as binary and the other half as continuous. The five binary variables and the travel cost are simply taken from Padre Island Gulf Coast data. The remaining four are generated randomly from uniform distribution. These different experiments lead us to conclude that preference coefficients are rapidly recovered when all the variables are continuous than when they are binary. In order to recover preference coefficients for binary variables, relatively large sample size is needed than for continuous variables. The recovery rate of preference coefficients for half discrete and half continuous variables lies between all binary and all continuous variables. Thus, for this study, we use a mixed data set with half binary and half continuous variables. This data set includes 64 beach sites, 9 attributes that defines each beach and samples with specified number of observations (discussed in section 3.2.2). We use Padre Island Gulf Coast data as the reference to generate these data. The use of mixed data set can avoid the possibility of not reflecting the strength of people's preferences and WTP for a potential change in the good's attributes. Use of such type of data sets is also common in many empirical studies.

Calculation of travel cost

Travel cost is the price of trips to the chosen site. Incorporating trip cost to other sites allows us to account for the substitutes. Travel cost is calculated as the sum of opportunity cost of time and a trip cost.

P_{ij} = [1/3× (yearly income/2080] × time spent on traveling+ 0.0963 × distances to and from the site.

We assume an individual works eight hours a day for five days in a week. The price of gas per gallon is assumed $2.89 and the vehicle gives thirty miles a gallon.

3.2.2 Description of the Monte Carlo Simulation Technique and Benefit transfer

The Monte Carlo Simulation is a method for iteratively evaluating a deterministic model using sets of randomly generated data. This technique is used to imitate the real life or make predictions. In order to replicate the real world choice making process in our study, we face two major sources of data limitations. First, we do not know the true preferences underlying the population. Second, we have no idea about idiosyncratic error component. Therefore, we apply the Monte Carlo Simulation technique to imitate the choice making process. Based on the RUM, we simulate the choice of an individual. The data generating process is a conditional logit model, with binary dependent variable (y_i). We proceed the simulation through following steps.

1. As explained in section 3.2.1, we use the mixed data set for independent variables. The four continuous variables are first obtained as a random draw from the uniform distribution for each 64 sites. Remaining variables are taken from the Padre Island Gulf Coast data.
2. We specify a vector of preference parameters for the population- which is considered as the true preference parameter for the population through out the experiments. True preference parameter for the population is defined the same for both the study and policy sites.
3. Observed utility (V_{ij}) derived by each individual from each site is calculated (equation 3.2).
4. Idiosyncratic error component (ε_{ij}) is drawn randomly from type I extreme value distribution with mean zero and specified variance for each individual and alternative (discussed in section 3.2.4). Then U_{ij} is calculated (equation 3.1) for each individual and each alternative.

5. The RUM suggests that an individual chooses the alternative that provides the maximum utility (U_{ij}). Therefore, we set choice (y_{ij})-dependent variable-equal to 1 for alternative *j* when U_{ij} is the maximum and zero for the alternatives other than *j*.
6. The maximum likelihood estimates from the conditional logit regression of *y* on independent variables (refer Table 2) is obtained. After estimating logit model, we calculate expected mean WTP (equations 3.14 and 3.17) for four different scenarios- that will be described in section 3.2.3- for the policy site by using the function estimated from the multinomial logit model.
7. We generate a new set of idiosyncratic error component for the policy site with the different variance than for the study site. We estimate mean actual WTP (equation 3.10 and 3.16) and for the policy site under four different scenarios -explained in section 3.2.3- by using the function estimated at step 6. Estimated mean expected WTP is the same for both the study and policy sites under ideal conditions.
8. Both the mean expected and actual WTP is calculated for the same sample using the true preference parameter, which is regarded as the true mean expected and actual WTP for the sample at the policy site. True mean expected WTP is the same for both the study and policy sites under ideal conditions, while actual mean WTP are different between these two sites.
9. We conduct validity test as described in section 3.1.3 for both hypothesis H1 and H2. The Wald test and original t-statistic is calculated as described in 3.1.3. In order to conduct bootstrap t-test, the model is re-estimated by drawing a new set of errors with the same scale of the study site to get a new set of preference coefficients and is used to estimate the WTP at the policy site. Then t value is calculated as described in 3.1.3. This process is repeated 399 times. These t values are sorted in ascending order. This will create a t-distribution. For two

sided test, the bootstrap critical values at level α=0.05, are the lower $\alpha/2$ and upper $\alpha/2$ quantiles of the ordered test statistics. Null hypothesis is rejected at level α if the original t-statistic lies outside the range.

10. We calculate transfer error as described in section 3.1.4
11. We repeat this simulation for 2000 times. In each new simulation, we repeat steps 4 to 10.
12. We calculate the probability of rejecting null hypothesis out of total 2000 simulations.
13. We calculate the mean transfer error out of total 2000 simulations.

Sample size sensitivity and Monte Carlo Simulation

Monte Carlo studies often consider the range of sample sizes. We simply set sample size of 100, 500 and 1000 for our study. In this study, the 9 attributes of alternatives are independent of sample size (see section 3.1.1). The only attribute that varies with sample size is travel cost. We draw income randomly from the uniform distribution. Then we calculate travel cost for each sample size. For each sample size, we conduct simulation steps 1 to 10.

3.2.3 Scenarios for Calculating WTP

Changes in welfare under this study are considered from four different perspectives. WTP are calculated for following changes.

I. Change in binary attribute(s) of alternatives

Under this scenario, we consider two cases. First, we choose the binary attribute with the largest coefficients. The change is allowed in such a way that all the alternatives are assumed to have that particular attribute. This implies that the values for that attribute become 1 for all the alternatives after the change. WTP (equation 3.10 and 3.14) is calculated for this change. Second, we change all the five binary attributes (X_{j1}, X_{j2}, X_{j3}, X_{j4} and X_{j5}). We assume that all

the alternatives possess these attributes. It means that all these variables become 1 for all the alternatives when the change is allowed. Both the actual and expected WTP is calculated.

II. Change in continuous attribute(s) of alternatives

We consider this change in two ways. First, we choose the continuous attribute with the highest coefficient. We increase that particular attribute by 20% in all the alternatives. Second, we increase all the continuous attributes (X_{j6}, X_{j7}, X_{j8}, X_{j9} and P_{ij}) by 20%. We assume that all the alternatives have this increase. We consider travel cost is also increased by 20% for each individual to all the beach alternatives. Both the actual and expected WTP is calculated for these changes.

III. Site lost based on the highest number of visitors

We drop the alternative that has highest number of visitors. Both the actual and expected WTPs are calculated for the loss of alternative. We calculate the sites with the greatest number of visitors.

3.2.4 Experiment Set Up for Benefit Transfer Simulations

To meet the objectives of the study, we design seven experiments in which data are created and changed systematically to analyze the effects of these changes in the quality of transfer. These experiments are designed in such a way to allow the study and policy sites to differ in specified ways to get the magnitude of error involved in benefit transfers.

The simulation and transfer exercise are conducted as described in section 3.2.2. These experiments are discussed below.

3.2.4.1 Scale differences

This experiment assumes that the study and policy sites are identical. We use the same mixed data sets for both the sites. The only thing that is different between them is the scale

parameter. This implies that we are drawing idiosyncratic error component with different variance for the study and policy sites. We change the scale parameter systematically between the study and policy sites, and run the simulation and transfer exercise with different scale parameters. We change the scale so that it comes from the set of {2, 5, and 10}. The study site can have the scale of 2, 5, or 10 and the policy site has the scale of {5, 10}, {2, 10} or {2, 5} respectively. This allows us to evaluate the effect of transfer from higher scale to lower scale and vice-versa.

3.2.4.2 Substitution difference

This experiment assumes that the study and policy sites have the same scale. The attributes that are observed for the alternatives at the study site are also observed at the policy site and these attributes are sufficient to define the alternative. There is no case of omitted variable bias. The only thing that is different between the study and policy sites is the number of choice alternatives. We define the choice sets of 16, 32 and 64 alternatives. If the study site has 16 site alternatives, then the policy site will have either 32 or 64 site alternatives. Alternatively, we will allow the study site to have 32 site alternatives and the policy site to have 16 or 64 site alternatives. We also take the case where the study site has 64 site alternatives and the policy site has 16 or 32 site alternatives. We take the same mixed data sets that we defined in section 3.2.2 step 1. But this data set consists of 64 site alternatives. We draw randomly the choice set of 16, 32 and 64 alternatives from that choice set.

3.2.4.3 Attribute(s) differences

This experiment investigates the effect of unobserved characteristics at both the study and policy sites. In this set up, we have the same scale and the choice set between the study and policy sites. We drop one of the observable characteristic from the study site. It means we

cannot access this variable at the policy site that may influence the choice of an individual. We drop one variable at a time that has the largest impact on observed utility (calculated from true preference parameter), among all other variables. Thus, the variable that is dropped can be either continuous or binary depending upon its impact on utility. Similarly, we conduct the same routine to the policy site. In this case, we assume the study site has more observable variables than at the policy site. These variables are actually present at the policy site but are immeasurable under benefit transfer practice. It refers that we cannot access this particular variable in benefit transfer application. We set the level of that particular variable as 'zero' at policy site under benefit transfer application.

3.2.4.4 Preference differences

In all of the above set up, we have specified the same population parameter for the study and policy sites. Here, we have different population parameters for the study and policy sites. Consequently, the preference coefficient estimates from the maximum likelihood estimation will be different. This allows us to investigate the effect of preference difference on transfer error. We conduct this experiment twice. First, we change the true coefficients on one binary and one continuous variable with the highest true preference coefficient. We choose a binary variable with the least true preference coefficient among the entire binary variables. Similarly, a continuous variable is also chosen. Then, they are changed with the highest preference coefficient. Second, we change the coefficients on one binary and one continuous variable with the minimum true preference coefficient. The binary variable that has the highest preference coefficient among all binary variables is chosen. Similarly, a continuous variable is also chosen. Then they are replaced with the minimum of true preference coefficients. In this set up, the study and the policy sites have the same mixed data set and the scale parameter.

3.2.4.5 Measurement error in variables

This experiment allows us to evaluate the effects of measurement error in benefit transfer applications. The variable of interest is the one with the highest preference coefficient, and it is measured with error at the study site. We introduce measurement error in the variable of interest by generating data randomly from normal distribution with the same mean and one third of the variance of the variable when it is measured with no measurement error. The preference coefficients are estimated for the variables with measurement and no measurement error. The coefficients so estimated are transferred to estimate the WTP at the policy site. The magnitude of transfer error in function transfer estimated from variables with measurement error is compared to the transfer error in function transfer estimated from variables with no measurement error.

3.2.4.6 Fixed versus Random parameter

This experiment assumes the presence of preference heterogeneity in one of the site. For the sake of convenience, preference heterogeneity is assumed at the study site while they are fixed at the policy site. The Random Parameter Logit Model (RPLM) is estimated at the study site and transferred to the policy site. WTP are measured from both the coefficients estimated from the RPLM and the true fixed preference parameter.

Chapter 4: Results and Discussion

4.1 Benefit Transfer under Scale Differences

Variables included in the choice model discussed below are defined in Chapter 3: Section 3.1.1. The results of multinomial logit estimation show that almost all of the site characteristics are highly significant in influencing site choice except X_8 and X_9 for a representative sample of 1000 observations and a error variance of 2(Table B1). Those variables that are significant at the scale of 2 have no significant influence in site choice upon increasing the scale. Only travel cost (P) and X_6 are significant for the same representative sample of 1000 observations and the scale of 10. The models shown in Table B1 clearly show that the scale parameter influences the mean preference coefficients while true preference parameters for generating the choices are the same. Results show the magnitude of preference coefficients decreases with the increasing the scale parameter. The sole reason for insignificancy of characteristics as the scale increases, which are significant at lower scale, can be taken as the effect of scale. The scale parameter is inversely related with the error variance. This implies that we are dividing each of the preference coefficients with the magnitude of the scale. Therefore, higher the magnitude of the scale, lower will be the preference coefficients.

The magnitude of bias increases with the increase in scale for the same representative sample (Table B2). Even though MRS is not confounded with scale, the effect of scale (error variance) is still prominent which is observed through increasing bias as the scale increases. The larger variance in the idiosyncratic error component influences the individual choice and thus influences the preferences for the site characteristics. This can lead to imprecise estimates of preference coefficients.

The functions which are estimated at the study site is transferred to the policy site and the WTP at the policy site is calculated based on this function. In order to assess the magnitude of error, transfer is done by keeping everything same between the study and policy sites except the scale parameter. The scenario is described in Table 3 below.

Table 3: Scenario of benefit transfer under scale difference

Variable	Study site	Policy site
P_{ij}	Same	Same
Y_i	Same	Same
x_{j1}	Same	Same
x_{j2}	Same	Same
x_{j3}	Same	Same
x_{j4}	Same	Same
x_{j5}	Same	Same
x_{j6}	Same	Same
x_{j7}	Same	Same
x_{j8}	Same	Same
x_{j9}	Same	Same
Scale	2	5 or 10
	5	2 or 10
	10	2 or 5
Choice set (J)	64	64
True preference parameter	Same	Same

88.80% or more out of 2000 tests suggest the transferability of the benefit function (Table B3). Hypothesis test on equality between estimated and true WTP suggest that WTP are statistically equivalent 89.7% or more. The probability of rejecting null hypothesis comes out to be comparatively less (1.9-3.5%) for the case of site loss at the policy site. However, the probability of rejecting null hypothesis changes dramatically with the increase in scale. A very important point to note here (Table B4) is that the probability of rejecting null hypothesis is relatively less for policy option 2 and 4 than the policy option 1 and 2 (described in Table 4). In majority of the cases, the rejection rate goes down with the increasing scale. This can be

explained as the effect of scale. As the scale increases, the effect of random component increases. Consequently, the variance on WTP also increases.

Table 4: Scenario for welfare measurement

Policy option No	Policy options used in welfare estimation (WTP)
1	All site are assumed to have X_5 attribute
2	All site are assumed to have all binary attributes
3	20% increase in X_6
4	20% increase in all continuous attributes including travel cost
5	Loss of a site

As shown in chapter 3, we find that transfer error on expected WTP is insensitive to the scale parameter (Table B5). However, transfer error on actual WTP is sensitive to the scale parameter. Effect of the scale while measuring expected WTP is exhibited through its effect on the estimation of preference coefficients on the study site. Actual WTP is influenced by the scale parameter through both additive error component and its effect on the estimation of the preference coefficients.

For almost all cases, transfer error is higher for a function transferred from higher scale than from lower scale. Site closure is very sensitive to the magnitude of scale at which the function is estimated. Site closure leads to least bias if a function is transferred carefully from the lower scale (8.15% for study site scale 2 to policy site scale of 5). On the other hand, site closure leads to the largest bias with a wrong estimate of as high as 49.85% (for study site scale 10 to policy site scale 5) when a function from the higher scale is transferred. For a policy option1, error on WTP is higher in magnitude when a function is transferred from the higher scale (as high as 32.04%) than the one transferred from the lower scale (as high as 20.19%).

However, for policy option 2, the range of error is 9.86-17.24% which is comparatively less than the change in just one binary attribute (policy option 1).

Transfer error on WTP is high for a change in continuous attribute than for the change in binary. While transfer error on WTP shows rapid improvement over sample size for continuous attributes. The magnitude of error on WTP for a policy option 3 is 21.53% when a function estimated at scale of 2 is transferred. While for the same case, the error is as high as 59.26% when a function is estimated at scale of 10. The transfer error on WTP for policy option 4 lies in the range of 1.9-14.65%. The lowest is for the scale parameter 2 and the highest is for the scale parameter 10. The effect of scale parameter is less when there is proportionate change in all attributes in each site. This can be due to the offsetting effect[10] followed by the changes in the attributes. The results also help us to conclude that offsetting effect is more prominent for the change in all-continuous attributes (policy option 4) than all-binary attributes (policy option 2). The magnitude of transfer error shows comparatively less improvement with the increase in sample size for the change in binary attribute than for the change in continuous attributes. The magnitude of transfer error on expected WTP is approximately similar in magnitude of transfer error in actual WTP. The error magnitude is very less when a function estimated at scale of 2 is transferred to policy site having scale of 10. In contrary, error magnitude is very high in the reverse cases. We can infer that benefit transfer is more appropriate in one direction only as mentioned by Kirchhoff et al (1997). In general, we find the tendency of underestimating the true WTP when a function is estimated at small scale and overestimating the true WTP when a function is estimated at high scale.

From this experiment, the following point can be summarized.

[10]When the effect of a change in one attribute is counteracted or compensated by the change in other attribute.

1. Effect of scale parameter in the RUM depends upon objectives of the estimation at the policy site. If we want to estimate MRS and WTP at the policy site, we do not need to worry about scale parameter. Nevertheless, scale parameter influences the estimation of preference coefficients. The higher the magnitude of the scale parameter, the lesser will be the precision of preference vector. Table B2 shows the increase in bias as the scale increases for the same data. In essence, error associated in transfer with scale difference is from imprecise estimation of preference coefficients. Researcher should be very careful in defining the utility function so that very less or no information is embodied in error component. More specifically, primary research should do a good job in defining the utility function. Benefit transfer practitioner should be able to judge whether the function chosen in the study site represents the function at the policy site. If there is no other source of error while conducting transfer exercise, the error due to scale differences can be minimized by taking large representative sample.

2. Both MRS, estimated WTP and true WTP are found to be statistically equal majority of the times at both the sites.

3. The scale parameters play a large role in biasing the welfare estimates when the scale parameter is appreciably different between the study and policy sites. As mentioned above, impact of scale parameter on WTP is through its effect in estimation of preference coefficients. This analysis indicates that it is always better to transfer the function estimated at lower scale parameter.

4.2 Benefit Transfer under Substitution Possibilities

The results from multinomial logit estimation are presented in Table C1 (Appendix C). The preferences for attributes change largely with the choice sets. Most of the attributes defined for a site are not significant when choice set consists of just 16 alternatives for the same sample.

However, the same attributes are highly significant when the choice set consists of 64 alternatives. In all different choice sets, travel cost is the only variable that is highly and consistently significant. For a sample of 1000 observations, travel cost and X_5 are significant at 1% while X_4 is significant at 5% when an individual faces only 16 site alternatives. All variables except X_2 are significant at 5% level of significance when the same individual faces 32 site alternatives. Similarly, when an individual faces 64 site alternatives, all variables are significant at 5% except X_2.

The magnitude of bias decreases comparatively with the increase in choice alternatives for the same representative sample (Table C2). The decrease in bias with the increase in choice alternatives can be thought as more site alternatives may introduce higher variation in the sample. In revealed preference setting (TCM), our estimates are consistent with the findings of Parson and Hauber (1998), DeShazo (2001), and Whitehead and Hoevanagel (2002) who mention that parameter estimates may be affected by the researchers choice set specifications.

Table 5: Scenario of benefit transfer under different substitution possibilities

Variable	Study site	Policy site
P_{ij}	Different (based on site alternatives)	Different(Based on site alternatives)
Y_i	Same	Same
x_{j1}	Same	Same
x_{j2}	Same	Same
x_{j3}	Same	Same
x_{j4}	Same	Same
x_{j5}	Same	Same
x_{j6}	Same	Same
x_{j7}	Same	Same
x_{j8}	Same	Same
x_{j9}	Same	Same
Choice set	16 32 64	{32, 64} {16,64} {16,32}
Scale	2	2
True preference coefficients	Same	Same

Among all the transfer errors calculated under five different scenarios (described in Table C3), transfer error on WTP for policy option 1 is the highest for a case of function transfer from the study site having 16 alternatives to the policy site having 64 alternatives (43% for sample size of 1000). The error magnitude is drastically lowered in the reverse case (6%). We consistently find (except in few cases) that function transfer made from sites with higher alternatives than policy site produce less transfer error than function transfer made from site with lesser alternatives than policy sites. In Table C3(Column 4), though the function transfer from the study site having 32 alternatives to the policy site having 64 alternatives produces less error (2.9%) than the function transfer from the study site having 64 alternatives to the policy site having 32 alternatives (4.98%), the difference is not large. The result follow the similar trend of decreasing transfer error on WTP when the transfer is made from the sites with the higher choice alternatives than the policy sites for policy option 2 as well (Table C3, Column C6 & C7). However, error range is very narrow (0.010-12.87%) for policy option 2 than policy option 1 (5-43%) when there is different substitution possibilities between the study and policy sites. This can be due to offsetting effect because of changes in all the binary variables.

Transfer error shows sample size sensitivity more in the context of changes made in continuous attributes than in the context of binary attributes (Table C3, Column C8, C9, C10, & C11). The error range is comparatively narrow and lies in the range of 15-27% for policy option 3 than the error range (5-43%) of policy option 1. The Table C3 (Column 8) indicates deviation from the trend we find in the case of policy option 1 and 2. Transfer error is less when the function transfer is made from the study site with 32 alternatives to the policy site with 16 alternatives (15.78%) than from the study site with 16 alternatives to the policy site with 32

alternatives (15.87%); the reduction in transfer error is very less. For the rest of the transfers, we do not observe this trend. Thus this can be explained as simulation error.

In policy option 4 (Table C3, Column 10 &11), we find that transfer error is comparatively less when a function transfer is made from the study site with large number of substitutes than the policy site in majority of the cases. The error magnitude is significantly reduced to 5.4% for function transfer with the study site having 64 choice alternatives to the policy site having 16 choice alternatives. On the other hand, we can see the error magnitude of 13% if the transfer is conducted the other way. We observe similar case when a transfer is made from the study site having 64 choice alternatives to the policy site having 32 choice alternatives and vice versa. However, the transfer made from the study site with 32 choice alternatives to the policy site with 16 alternatives and vice versa do not confirm to this common observation. This can be due to some extreme random draws.

We observe this common observation for all transfer cases in policy option 5(Table C3-Column 12 &13). The error magnitudes are comparatively less for all the transfers than are those in policy option 1 and 4. The error magnitude for the function transfer from the study site having 16 alternatives to the policy site with 64 alternatives is 34.4%. On the other hand the error magnitude is lowered to 20.41% when they are reversed. Likewise, we can observe the error of 11.6% for a function transfer from the study site having 32 alternatives to the policy site having 64 alternatives while there is an error of 2.96% only when the function transfer is carried in reverse order. The Table C3 on transfer error in appendix C show that in majority of cases, transfer error in expected WTP are similar(in magnitude) as the actual WTP. In this study, we also note that in majority of the cases, errors due to estimation and different substitution possibilities travel in different direction. Consequently, the transfer error is lowered when a

transfer exercise is taken between the sites with different substitution possibilities than those between the sites with similar substitutes.

In summary, results indicate that the availability of choice sets in the study and policy sites plays important role to decide the transferability of benefit functions. We can infer that transfer error can be drastically reduced if a function transfer is done from the study site with large number of substitutes (choice alternatives) to the policy site with less number of substitutes than the function transfer from the study site having less number of substitutes to the policy site having a large number of substitutes. The transfer error is comparatively less for the changes in all attributes in this experiment than for the change in just one attribute. From our results, it can be inferred that the transfer error on actual WTP is similar in magnitude as expected WTP.

4.3 Benefit Transfer under Characteristics Differences

The principal consideration in a function transfer is whether the transferred function consists of all the variables that are important in the context of the policy site. Moreover, how does the function transfer perform when data on some variables in the function transfer are not available in the policy context? This experimental set up tries to address these two questions. In the first case, we assume one of the important variables in policy context is absent in the study site. This means an important variable in the policy site is missing in the transferred function from the study site. This type of transfer will be called as case I. In the second case, we assume the variable is present in both the study and policy sites but the variable can not be measured easily in the policy context. This case will be called as case II. For example, variable like respondent's attitude cannot be directly accessed in the policy context. The preference coefficients for the first and ideal cases (the study and policy sites are exactly same) from multinomial logit estimation are presented in Table D1. The preference coefficients and their

significance changes with the exclusion or inclusion of a variable (X_5) at the study site. For the sample of 1000, all the variables except X_9 are significant at 1% level for the first case of characteristic difference. However, all variables are significant at 1% level except X_9 and X_8 for the ideal case. The preference coefficients are the same for the second case of benefit transfer (Table D4).

Table 6: Scenario of benefit transfer under characteristics differences

Variable		Study site	Policy site
P_{ij}		Same	Same
Y_i		Same	Same
x_{j1}		Same	Same
x_{j2}		Same	Same
x_{j3}		Same	Same
x_{j4}		Same	Same
x_{j5}	Case1	Unobserved at study site	Observed in policy site
	Case 2	Observed at study site	Observed in policy site but not accessible under benefit transfer case (variables such as attitude, ability, skill)
x_{j6}		Same	Same
x_{j7}		Same	Same
x_{j8}		Same	Same
x_{j9}		Same	Same
Choice set		64	64
Scale		2	2
True preference coefficients		Same	

Case I type of benefit function transfer

Result shows transfer error for both expected WTP and actual WTP decreases consistently with the increasing sample size in ideal transfer for policy option 1 (Table D3, column 3 &4). However, transfer error decreases drastically from 115.05% to 7.26% as sample increases from 100 to 500. Error increases as the sample size reaches to1000, though the increase in the magnitude of error is very less (0.7%). Similarly, we find inconsistent transfer error for the actual WTP in the policy option 1. The error magnitude decreases from 118% to

5.6% and then increases again to 19.48%. In policy option 2, we find consistently decreasing error for both ideal and case 1 type of transfers (Table D3, Column 5). A very important point to note here is that error magnitude is larger for case 1 type of function transfer (15.04%) than the ideal case (10.12%) for the sample size 1000. The result for policy option 3 follows similar trend as the policy option 2. The error magnitude decreases from 78% to 21% as the sample increases from 100 to 1000 while the error magnitude decreases from 73% to 24.87% as the sample increases from 100 to 1000 (Table D3, Column 7).

We do not find consistent results for policy option 4 with the increase in sample size. The transfer error is very low (1.6%) for sample size of 100 which then increases to 61.97% as sample increases to 500; the magnitude of error then decreases to 43.18% as the sample increases to 1000 (Table D3,Column 9). For the policy option 5, the transfer error decreases from 17% to 1.3% which then increases to 30.98% as the sample goes from 100, 500 and 1000 respectively (Table D3, Column 11). The magnitude of transfer error on expected WTP is approximately equal to actual WTP in the cases of policy option 2 and policy option 3. We can not make similar comments for policy option 1 and 4 as the results are deviated for some of the transfers.

Summary for Case I type of transfer:

Out of five different type of policy option, this study finds inconsistent result for three different policy option i.e. policy option 1, 4 and 5. These evidences limit the generalization of the results. At this point, we can only say that we obtain mixed type of results. There can be either the tendency of improving or worsening the transfer error with the increasing sample size. The effect of such type of transfer is more on site lost than in any other cases of policy scenario. In this study, we observe serious inconsistency for such policy change. We also find seven cases out of 15 in which transfer error on expected WTP are sufficiently different from transfer error

on actual WTP- the magnitude of difference is greater than 3. This also demands an analysis whether the difference is just a white noise or an issue to worry.

Case II type of benefit transfer

The magnitude of transfer error is inconsistent with the increasing sample size for the entire policy scenario. The results are difficult to compare with the ideal case because of inconsistency across different sample sizes. For policy option 1, we find that transfer error on expected WTP first decreases from 44.75% to 3.61 % which then increases to 20% for different sample size of 100, 500 and 1000 respectively (TableD6, Column 3). On the other hand, transfer error on actual WTP follow similar trend but the magnitudes are quite different from expected WTP. The error magnitude decreases from 29.89% to 4.53% and then increases again to 32.53% for the sample of 100, 500 and 1000 respectively (Table D6, Column 4). Similarly for policy option 2, error first decreases and then increases as the sample change from 100, 500, and 1000. However, for policy option 3, 4 and 5, error first increases and then decreases as the sample changes from 100, 500, and 1000 (Table D6). This type of pattern needs further analysis which is not done here because of limited resource and time.

Summary for case II type of transfer

In this study we find that benefit function transfer is inconsistent majority of times in case II than case I type of transfer. The magnitude of transfer error on expected WTP is different from actual WTP in many cases. In 3 different types of policy option (3, 4, and 5), we find that transfer error first increases and then decreases while for the remaining policy options, they first decreases and then increases. This raises the further concern about the change in magnitude of error with increasing sample size.

In brief, there is lot of uncertainty associated in the magnitude of transfer error when the benefit transfer is performed for either Case I or Case II type of function transfer. This uncertainty demands more thorough analysis with advance simulation. We can not come into any conclusion from this experiment. However, this study indicates some general and unpredictable patterns in the errors while practicing benefit transfer under characteristics differences. In broad terms, we find the magnitude of transfer error comparatively less for policy option 5 than any other policy options.

4.4 Benefit Transfer under Preference Differences

This experiment is set up in such a way that population at the study and policy sites are different with respect to their preferences. To allow this difference, we generate the data for the study site based on different true preference parameter than the policy site. While in our previous experiments, we have the same true preference parameter underlying the populations at the study and policy sites. We changed the true preference coefficients at the study sites from the policy site in two ways. For the first case, we replace the minimum coefficients among binary variables with the highest coefficient among all the variables. We also replace the minimum coefficient among continuous variables with the highest among all the variables. This case from this point onward will be called as case I. For the second case, we replace the highest coefficient among binary variables with the smallest coefficients among all variables. We do the same for continuous variables too. This case will be called as case II. The results of the multinomial estimation are listed in Table E1. From this table, we can see the different magnitude of estimated preference coefficients generated from different true preference coefficients of the population for each different sample size. For all the cases including the ideal

case of benefit transfer, (everything is identical between the study and policy sites), all variables are significant except X_8 at 5% level.

Table 7: Scenario of benefit transfer under preference differences

Variable	Study site	Policy site
P_{ij}	Same	Same
Y_i	Same	Same
x_{j1}	Same	Same
x_{j2}	Same	Same
x_{j3}	Same	Same
x_{j4}	Same	Same
x_{j5}	Same	Same
x_{j6}	Same	Same
x_{j7}	Same	Same
x_{j8}	Same	Same
x_{j9}	Same	Same
Choice set	64	64
Scale	2	2
Preference	Different	Different
Case 1.High	[0.5 0.4 0.3 **0.6** 0.5 0.6 0.5 0.4 0.3 **0.6**]	[0.5 0.4 0.3 0.3 0.5 0.6 0.5 0.4 0.3 0.2]
Case 2.Low	[0.5 0.4 0.3 0.3 0.5 **0.2** **0.2** 0.4 0.3 0.2]	[0.5 0.4 0.3 0.3 0.5 0.6 0.5 0.4 0.3 0.2]

Unlike the ideal case of transfer, we find the transfer error first increasing and then decreasing as the sample goes from 100, 500 and 1000. Our result on transfer error for different policy option indicates that they are inconsistent with the increasing sample size. The increase in magnitude of error as the sample changes from 100 to 500 is large in majority of the cases. This inconsistency also creates difficulty in comparing the result with the ideal case. For example, the error magnitude increases from 4% to 36% as the sample size increases from 100 to 500 observations in the case of policy option 1(Table E3, Column 3). For the same case, the error magnitude decreases to 16% when the observation reaches to 1000. The same pattern will follow the case II type of function transfer. For policy option 2, error magnitudes are inconsistent for case I but they are consistent with the case II. The error magnitudes are comparatively lower for the case I

than the case II even though they are inconsistent. For the case I, error range is 6-8% with the changing sample size. While the error range for the second case is 11-36% with the changing sample size (Table E3, Column 5).

For policy option 3 and 4, the error magnitude follows the similar pattern as in the case of policy option 1. They are highly inconsistent and error magnitudes are comparatively greater than in the policy option 1 and 2. For example, the error range for the case I under policy option 3 is 3-90%. The error range for case II under same policy option is 37-139 % (Table E3- column 7). A very important point to note from table E3 is that the error magnitude is comparatively less for policy option 2 and 4 than policy option 1 and 3. This indicates the error magnitudes are less for the change in all attributes than the change in one attribute. In the case of policy option 5, we find appreciable amount of difference in transfer error between expected and actual WTP (Table E3, Column 11 & 12). These errors are highly inconsistent too.

Broadly, this experiment shows the transfer error is inconsistent under preference difference. The uncertainty associated with the preference difference between the study and policy sites limits us from getting into any conclusion. This experiment point us the need of more rigorous simulation experiments with large sample and large number of simulations than we have in this experiment.

4.5 Benefit Transfer under Measurement Errors

This experiment is set up in such a way to investigate the impact of measurement error in variable in benefit transfer. We introduce the measurement error in the variable with the largest impact on systematic component of utility. In this simulation, the variable of interest is X_5. The measurement error is introduced in this study by generating data randomly from normal distribution with the same mean but one third of the variance of X_5.

Table 8: Scenario of benefit transfer under measurement error in variable

Variable	Study site	Policy site
P_{ij}	Different (based on site alternatives)	Different(based on site alternatives)
Y_i	Same	Same
x_{j1}	Same	Same
x_{j2}	Same	Same
x_{j3}	Same	Same
x_{j4}	Same	Same
x_{j5}	Different $x_{j5} \sim (\overline{X}_{j5}, \frac{1}{3} \text{ of var } (X_{j5})$	Different
x_{j6}	Same	Same
x_{j7}	Same	Same
x_{j8}	Same	Same
x_{j9}	Same	Same
Choice set	64	64
Scale	2	2
True preference coefficients	Same	Same

The results from multinomial logit estimation are presented in Table F1. We find the significance and magnitude of coefficients of variable changes with the introduction of the measurement error in X_5. The result also indicates that the magnitude of bias is not consistent for several variables with increasing sample size, unlike the case of ideal transfer. The magnitude of bias for X_1, X_2, X_3, X_4, X_5, X_6, and X_8 first decrease and then increases as the sample increases from 100,500, and 1000 respectively.

We find transfer error is comparatively higher for transfer of function estimated with measurement error in one variable than the ideal function transfer; these errors tend to decrease as the sample size increases. In policy option 2, we find that transfer error decreases consistently with increasing sample sizes. However, the magnitude of transfer error is greater for the function transfer involved with measurement error than the ideal function transfer. In ideal function transfer, the error magnitudes are 18.6%, 12.2%, and 10.1% for the sample sizes of 100,

500, and 1000 respectively while it is 20.74%, 9.2%, and 5.2 % for the sample sizes of 100, 500 and 1000 under function transfer with measurement error in a variable (Table F3, Column 5). The result is completely different for policy option 1 (Table F3, Column 3 & 4). The magnitude of transfer error increases consistently with the increasing sample size unlike the case of ideal transfer. This can be due to the measurement error involved in a binary variable. In policy option 3, error magnitude for ideal case of transfer is 78.6%, 28.9%, and 21.7% for sample size of 100, 500, and 1000 respectively; the magnitude of error for the same policy option under function transfer estimated at the study site involving measurement error in the variable are 670.47, 297.837, and 238.327% respectively (Table F3, Column 7 & 8) . Similar is true for policy option 4 though the magnitude of error involved in this transfer is comparatively less than policy option 3. We find similar result for policy option 5. The transfer error is only 6% for ideal function transfer while it is 14.40% for function transfer involving measurement error in one of the variables (Table F3, Column 11). We also find that transfer error is less for policy option 5 than any other policy options.

Broadly, this experiment finds that transfer error increases with the introduction of measurement error in the variable. However, these errors decrease with the increasing sample size. We find transfer error is very less in the case of site closure than in any other case. This experiment also indicates that there is less transfer error involved for a change in all attributes than the change in one attribute.

4.6 Benefit Transfer under Preference Heterogeneity versus Fixed Preferences

This experiment allows us to evaluate the performance of benefit transfer in the case of preference heterogeneity at the study site and the fixed preference at the policy site. The scenario of benefit transfer is presented in Table 9.

Table 9: Scenario of benefit transfer under preference heterogeneity versus fixed preferences

Variable	Study site	Policy site
P_{ij}	Different (based on site alternatives)	Different(based on site alternatives)
Y_i	Same	Same
x_{j1}	Same	Same
x_{j2}	Same	Same
x_{j3}	Same	Same
x_{j4}	Same	Same
x_{j5}	Same	Same
x_{j6}	Same	Same
x_{j7}	Same	Same
x_{j8}	Same	Same
x_{j9}	Same	Same
Choice set	64	64
Scale	2	2
Preference coefficient (true)	B~(mean(preference coefficients at study policy site), 1)	[0.5 0.4 0.3 0.3 0.5 0.6 0.5 0.4 0.3 0.2]

The results from the multinomial logit model are presented in Table G1. The magnitudes of estimated preference coefficient in the case of preference heterogeneity are comparatively less than those estimated for the case of fixed preferences. Those variables which are found to be significant in the case of fixed preferences are insignificant in the case of preference heterogeneity. In the case of fixed true preferences, we find eight variables significant except X_8 and X_9 while only four variables i.e. P, X_1, X_5, and X6 are significant at 5% level when preference heterogeneity is assumed. We will call the case of same fixed true preference parameter as the ideal case since everything is identical between the study and policy sites. Likewise, we will call the case of preference heterogeneity as the case I. Preference heterogeneity in this study implies that we estimate the choice of an individual from the set of alternatives is estimated by using the preference coefficients that comes from normal distribution and are different for each individual.

In all the different cases of policy changes, we find transfer error to be very large in magnitude for case I than the ideal function transfer. These errors are inconsistent (policy option 1, 2, and 4) in majority of the cases with the increasing sample size. In the case of policy option 2, error magnitudes are 5117.28%, 1794.23%, and 2127.91 for the sample size 100, 500 and 1000 respectively (Table G3, column 5). The magnitude of error first decreases and then increases again with these different sample sizes. On the other hand, the errors involved in the case of ideal transfer are only 18.64%, 12.20%, and 10.12% in the sample size of 100, 500, and 1000. We find similar observation in the case of policy option 1 but the error magnitude for both the ideal and case I type of function transfer is higher than the policy option 2. The error magnitudes are comparatively very high for the policy option 3. These errors are consistent than others. The error decreases from 6920.85% to 479.06% as the sample size increases from 100 to 1000 (Table G3, 7). Error magnitudes are comparatively less in the case of policy option 4 than those in the case of policy option 3. We find very less transfer error in the case of policy option 5 than others do and they are consistent too.

In summary, transfer error is very large in the case of preference heterogeneity at one site and fixed preferences over other site. In this study, we find transfer errors are relatively less in magnitude for the change in all attribute (policy option 2 and 4) than the change in one attribute (policy option 1 and 3). Transfer error in the case of site closure is the least among all the different cases of policy options.

4.7 Benefit Transfer under more than One Difference between the Study and Policy Sites

This experiment is conducted by allowing the study and policy sites to differ in more than one factor at the same time. The scenarios of difference between study and policy sites are described in Table 10.

Table 10: Scenario of benefit transfer under more than one difference between the study and policy sites

Variable	Study site	Policy site
P_{ij}	Different (based on site alternatives)	Different(based on site alternatives)
Y_i	Same	Same
x_{j1}	Same	Same
x_{j2}	Same	Same
x_{j3}	Same	Same
x_{j4}	Same	Same
x_{j5}	Observed	Inaccessible at policy site
x_{j6}	Measure with error $X_{j6} \sim (\overline{X}_{j6}, \frac{1}{3} \text{ of var } (X_{j6})$	Measured with no error
x_{j7}	Same	Same
x_{j8}	Same	Same
x_{j9}	Same	Same
Choice set	32	16
Scale	2	5

The error magnitude we find here is greater in magnitude in the case of multiple differences than the case of single difference between the study and policy sites (the random versus fixed preference is not considered here). The error magnitude can be as large as 811% for the sample of 1000 (Table H3). The error magnitudes are inconsistent for policy option 2, 3 and 5 while they are consistent for the rest of the policy options. For the policy option 1, error decreases from 540.80% to 281.89% when sample increases from 100 to 1000 (Table H3, Column 3). However, for the policy option 2, the error first increases from 37.13% to 137%, and then increases again to 100.64% as the sample changes from 100, 500, and 1000 respectively (Table H3, Column 5). The magnitude of error is less for policy option 2 than policy option 1 though they are inconsistent. We find similar result for the policy option 3 as policy option 2. The error magnitudes are less for policy option 4 than policy option 3. For the sample size of 100, the error is 811.59% in the case of policy option 3 while it is 281.89% for policy option 4 (Table H3,

Column 7 & 9). The error is comparatively less for site loss though they are inconsistent. The error first decreases from 49.45% to 7 % and again increases to 87.45%. (Table H3, Column 11).

In summary, when the study and policy sites differ in more than one characters or criteria (specified in chapter one) benefit transfer suffers from large magnitude of error. From the above experiments, we can infer that this large magnitude of error is due to transfer made from the study site with preference heterogeneity to the policy site with fixed preferences. We also find that transfer error is the least for the site closure. Transfer error for the changes in all attributes is less than the change in one attribute.

Chapter 5: Policy Implications of Benefit Transfer

This chapter presents policy implications of the findings from the different experiment that we conducted in this study.

In most policy settings that call for benefit transfer estimation, an analyst may not have any idea about the unobserved variances in the random component and the other baseline conditions. The unobserved error or random component plays significant role in the determination of choice and consequently in the estimates of preference coefficients and the magnitude of errors in welfare estimate. In real analysis, the analyst may not have any idea about this unobserved heterogeneity. This implies that performance of benefit transfer is always limited by real world data availability. The researchers' judgment of choosing the study site plays an important role in reducing transfer errors in benefit transfer practice. Therefore, researcher should carefully choose the original studies which have carefully specified the WTP or utility functions such that the unobserved component do not include any important information. If utility/WTP functions are specified correctly, the unobserved component will be just white noise (Train, 2003).

When the numbers of available substitutes between the study and policy sites are different, it is suggested that original studies should be so chosen that they have larger number of site substitutes than the policy sites. Involving original studies with larger number of substitutes includes more variance (variance in the systematic component) which can thus reduce the transfer error. Many uncertainties are associated with the function transfer when a variable of interest is not included in the model of the study sites. Similar is true when important variables at the study sites' model are missing or data on these variables are unavailable at the policy site. This study does not suggest function transfer under such difference between the study and policy

sites. Function transfer under preference difference between the study and policy sites' populations should not be conducted as well. This study indicates inconsistencies in such kind of transfer.

It is suggested that there be no measurement error in the variables of interest either at the study or policy sites. The magnitude of error is amplified when there is measurement error in any one or more variables of interest than when there is no error in these variables. However, if analysts suspect the presence of measurement error in the variables, the magnitude of error can be lowered by taking sufficiently large samples. This study finds that preference heterogeneity should be carefully considered in benefit transfer and plays a significant role in determining the magnitude of transfer error. In real analysis, it is less likely to find the same preferences for different attributes among different people. An important finding is that, large sample size sufficiently representing the population is one solution for majority of the problems investigated in this study. The effect of the random component, different site substitutes and measurement errors can be addressed by taking large sample sizes. They are all sensitive to the sample sizes. Majority of research suggest the greater the similarity between the study and policy sites, the lesser will be the magnitude of error. This study does not suggest benefit function transfer when the study and policy sites are different in several important factors; the quality of function transfer can be drastically reduced.

Chapter 6: Conclusions and Directions for Future Research

This chapter presents general conclusions from this study. The final section briefly considers limitations in the study and issues for further research.

6.1 Conclusion

This study conducts an empirical investigation in the transfer of choice model and valuation function from the study site to the policy site when they are different from each other. Broadly, this study addresses the question of how the benefit transfer is influenced by the difference between the study and policy sites. Six different cases of difference between the study and policy sites are studied:

- Differences in the scale parameter
- Differences in the substitution possibilities
- Differences in the observable characteristics
- Differences in the preferences
- Measurement errors in the variable of interest
- Preference heterogeneity in one site and fixed preferences at the other site

This study investigates the criterion validity of benefit transfer. Based on the experiment we conducted, the criterion validity is not rejected most of the times when the study and policy site are different only in terms of scale parameter. The findings in this study may not generalize easily to other context. However the result in this study shows that lower the magnitude of error variance, lower will be the transfer error and vice versa. The transfer error magnitude on expected WTP lies in the range of 6.72 - 21.53% for the scale of 2, while it ranges from 17.24 - 59.26 % for the scale of 10 (for sample size 1000) for the different policy changes. In the case of

different substitution possibilities between the study and the policy site, we find the transfer error will be lowered if the transfer is taken from the study site with greater substitution possibilities than the policy site. The transfer error on expected WTP falls in the range of 11.40 - 43.03% (sample 1000) when a function is transfer from study sites with 16 substitutes to the policy site with 64 substitutes. Similarly, the error on expected WTP for the reverse case of transfer is in the range of 0.10 - 23.77% (sample 1000) for the different policy changes.

In this study, benefit function transfer is inconsistent with the increasing sample size and performs poorly when the observed characteristics at the study site are not accessible at the policy sites. Similarly, they are infeasible and inconsistent when the transfer is made under the condition of differences in preferences of the populations between the study and policy sites. Measurement error in any of the variables increases the magnitude of transfer error. The transfer error on expected WTP lies in the range of 5.92 - 238.82% (sample size 1000) for the different proposed policy changes at the policy site. Benefit function transfer exhibits the worst performance when one of the sites has preference heterogeneity while the other has fixed preference. In this study, the magnitude of error involved in this case is the largest among all the different experiments conducted. The error magnitude as large as 2127.91% is observed on expected WTP for the sample of 1000. Benefit function transfer perform poorly with an error magnitude in the range of 87.45 - 811.59 % (in sample size of 1000) for different policy changes when the study site is different to the policy site in several factors. Benefit transfer should not be undertaken when the study and policy sites are very different. If benefit transfer is only the available alternatives, original studies with large number of substitutes than policy site, large sample size, well defined utility function which has all the important characteristics at the study

site should be chosen. Researcher should be careful to choose the representative sample with the sufficient variation in the amenity.

6.2 Directions for Future research

This study is conducted by simulating data where we assume we have all the information required to conduct benefit transfer. In order to confirm the findings from this study, we feel that it is necessary for conducting similar analysis with the real world data. Due to limitation of time, we are not able to test the criterion validity for our other sub objectives. We are not able to derive any specific conclusion for the case of characteristic and preference differences between the study and policy sites. A more detailed analysis with large number of simulation and sample size than this study may produce consistent results and help to derive conclusion. Due to limitation of time and resource, we are not able to conduct this research with large number of simulations and sample size. We have conducted the transfer of choice model and valuation function based on revealed preference data. In order to investigate how this study differs from the choice model based on stated preference data and their implications to benefit transfer, a similar study with stated preference technique can be conducted and compared with this study.

References

Alberini, A., Longo, A., and Veronesi, M. (2007). Basic statistical models for stated choice studies. In K. Barbara (Eds.) *Valuing environmental amenities using choice experiments: A common sense guide to theory and practice*. Springer.

Babbie, E. (2007). *The practice of social research*. California: Thomson Learning Inc.

Bateman, I. J., Jones, A. P., Nishikawa, N., and Brouwer, R. (2000). Benefit transfer in theory and practice: A review. London: Center for Social and Economic Research on the Global Environment.

Bergland, O., Magnussen, K., and Navrud S. (2002). Benefit transfer: Testing for accuracy and reliability. In R. J. G. M. Florax, P. Nijkamp, and K. G. Willis (Ed.). Comparative environmental economic assessment (117-132). UK: Edward Elgar Publishing Limited.

Ben-Akiva, M. E., and Lerma, S.R. (1985). *Discrete choice analysis: Theory and application to travel demand*. Cambridge: MIT Press.

Bergstrom, J. C., and De Civita P. (1999). Status of benefits transfer in the United States and Canada: A review. *Canadian Journal of Agricultural Economics*, 47 (1), 79-87.

Bockstael, N. E., and Mc Connell, K. E. (2007). *Environmental and resource valuation with revealed preferences: A theoretical guide to empirical models*. Dordrecht, The Netherlands: Springer.

Boyle, K. J., and Bergstrom, J.C. (1992). Benefit transfer studies: Myths, pragmatism, and idealism. *Water Resources Research*, 28(3), 657-663.

Brookshire, D. S., and Neill, H. R. (1992). Benefit transfers: Conceptual and empirical issues. *Water Resources Research*, 28(3), 651-655.

Brouwer, R., and Spaninks, F. A. (1999). The validity of environmental benefits transfer: Further empirical testing. *Environmental and Resource Economics*, 14(1), 95-117.

Brouwer, R. (2000). Environmental value transfer: State of the art and future prospects. *Ecological economics*, 32(1), 137-152.

Champ, P. A., Boyle, K. J., and Brown, T. C. (Eds.). (2003). *A primer on nonmarket valuation*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Chattopadhyay, S. (2003). A repeated sampling technique in assessing the validity of benefit transfer in valuing non-market goods. *Land Economics*, 79 (4), 576-596.

Colombo, S., and Hanley, N. (2008). How can we reduce the errors from benefits transfer? An investigation using the choice experiment method. *Land Economics*, 84 (1), 128-147.

Colombo, S., Calatrava-Requens, J., and Hanley, N. (2007). Testing choice experiment for benefit transfer with preference heterogeneity. *American Journal of Agricultural Economics*, 89 (1): 135-151.

Desvousges, W. H., Johnson, F. R., Banzhaf, H. S. (1998). *Environmental policy analysis with limited information: Principles and applications of the transfer method*. Northampton, MA: Edward Elgar.

Desvousges, W. H., Naughton, M. C. and Parsons, G. R. (1992). Benefit transfer: Conceptual problems in estimating water quality benefits using existing studies. Water *Resources Research*, 28 (3), 675-683.

Downing, M., and Ozuna, J. T. (1996). Testing the reliability of the benefit function transfer approach. *Journal of Environmental Economics and Management*, 30 (3), 316-322.

Groothuis, P. A. (2005). Benefit transfer: A comparison of approaches. *Growth and Change* 36, 551-564.

Hanley, N., Wright, R. E., and Alvarez-Farizo, B. (2006). Estimating the economic value of improvements in river ecology using choice experiments: An application to the water framework directive. *Journal of Environmental Management*, 78(2), 183-193.

Jiang, Y., Swallow S. K., and Mcgonagle, M. P. (2005). Context-sensitive benefit transfer using stated choice models: Specification and convergent validity for policy analysis. *Journal of Environmental Management*, 31(4), 477–499.

Johnston, R. J., Besedin, E. Y., and Ranson, M. H. (2006). Characterizing the effects of valuation methodology in function-based benefits transfer. *Ecological Economics*, 60 (2), 407-419.

Johnston, R. (2007). Choice experiments, site similarity and benefits transfer. *Environmental and Resource Economics*, 38 (3), 331-351.

Kirchhoff, S., Colby, B. G. and LaFrance, J. T. (1997). Evaluating the performance of benefit transfer: An empirical inquiry. *Journal of Environmental Economics and Management*, 33 (1), 75-93.

Legget, C., Markowski, M., Unsworth, R., Sigman B., Parsons, G., and Boyle K. (2002). Padre Island National Seashore Visitor Day Values. U.S. Department of Interior National Park Service, Environmental Quality Division, USA.

Loomis, J.B. (1992). The evolution of a more rigorous approach to benefit transfer: benefit function transfer. *Water Resources Research*, 28 (3), 701–705.

Loomis, J., Roach, B., Ward F., and Ready, R. (1995). Testing transferability of recreation demand models across regions: A study of corps of engineer reservoirs. *Water Resources Research*, 31 (3), 721–730.

Louviere, J., Street, D., Carson, R., Ainslie, A., Deshazo, J.R., Cameron, T., Hensher, D., Kohn, R., and Marley, T. (2002). Dissecting the random component of utility. *Marketing Letters*, 13 (3), 177-193.

McConnell, K. E. (1992). Model building and judgment: Implications for benefit transfers with travel cost models. *Water Resources Research*, 28 (3), 695-700.

Morrison, M., Bennett, J., Blamey, R., and Louivere, J. (2002). Choice modelling and tests of benefit transfer. *American Journal of Agricultural Economics*, 84(1), 161-70.

Navrud, S., and Ready, R. (Eds.) (2007). *Environmental value transfer: Issues and methods*. Dordrect, The Netherlands: Springer,

Parsons, G. R., and Hauber, A. B. (1998). Spatial boundaries and choice set definition in a random utility model of recreation demand. *Land Economics*, 74, 32−48.

Parsons, G.R., and Kealy, M.J. (1994). Benefits transfer in a random utility model of recreation. *Water Resource Research*, 30 (8), 2477-2484.

Piper, S., and Martin, W. E. (2001). Evaluating the accuracy of the benefit transfer method: A rural water supply application in the USA. *Journal of Environmental Management*, 63 (3), 223-235.

Randall, A., and Hoehn, J. P. (1996). Embedding in market demand systems. *Journal of Environmental Economics and Management,* 30, 369–380.

Rosenberger, R. S., and Loomis, J. B. (2000). Using meta-analysis for benefit transfer: In-sample convergent validity tests of an outdoor recreation database. *Water Resources Research*, 36 (4), 1097-1107.

Rosenberger, R.S. and Loomis, J. B. (2001). Benefit transfer of outdoor recreation use studies: A technical document supporting the Forest Service Strategic Plan (2000 revision). General Technical Report RMRS-GTR-72. Fort Collins, CO: USDA Forest Service, Rocky Mountain Research Station.

Rosenberger, R.S. and Loomis, J.B. (2003). Benefit transfer. In P.A. Champ, K.J. Boyle and T.C. Brown (Eds.), A Primer on Nonmarket Valuation. Dordrecht: Kluwer Academic Publishers.

Rosenberger, R.S., and Phipps, T. T (2007). Correspondence and convergence in benefit transfer accuracy: A meta-analytic review of the literature. In S. Navrud and R. Ready (Eds.), Environmental Values Transfer: Issues and Methods. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Shrestha, R. K., and Loomis, J. B. (2001). Testing a meta-analysis model for benefit transfer in international outdoor recreation. *Environmental and Resource Economics*, 39 (1), 67-83.

Smith, V. K., and Pattanayak, S. K. (2002). Is Meta-analysis a Noah's ark for non-market valuation? *Environmental and Resource Economics*, 22(1-2), 271-96.

Smith, V. K., Pattanayak, S. K., and Van Houtven, G. L. (2006). Structural benefit transfer: An example using VSL estimates. *Ecological Economics*, 60 (2), 361-371.

Spash, C. L., and Vatn, A. (2006). Transferring environmental value estimates: Issues and alternatives. *Ecological Economics*, 60 (2), 379-388.

Swait, J., and Louviere, J. (1993). The role of the scale parameter in the estimation and comparison of Multinomial Logit Models. *Journal of Marketing Research*, 30 (3), 305-314.

Train, K. (2003). *Discrete choice methods with simulation*. New York: Cambridge University press.

VandenBerg, T. P., Poe, G. L., and Powell, J. R. (2001). Assessing the accuracy of benefits transfers: Evidence from a multi-site contingent valuation study of groundwater quality. In J.C. Bergstrom, K.J. Boyle, and G.L. Poe (Eds.), The Economic Value of Water Quality. Northampton, MA: Edward Elgar.

Woodward, R. T., and Wui, Y.S. (2001). The economic value of wetland services: A meta-analysis. *Ecological Economics*, 37 (2), 257-270.

Zandersen, M., Termansen, M., and Jensen, F. S. (2007). Evaluating approaches to predict recreation values of new forest sites. *Journal of Forest Economics*, 13 (2-3), 103-128.

Appendix A: Summary of Benefit Transfer Validity Tests

Reference	Resource/activity	Valuation Methods used	Value transfer percent error	Function transfer percent error
Loomis (1992)	Recreation	Zonal TCM	4–39	1–18
Parsons and Kealy (1994)	Water/recreation	RUM TCM	4–34	1–75
Loomis et al. (1995) Nonlinear least-squares model Heckman model	Recreation	Zonal TCM	-	 1–475 1–113
Downing and Ozuna (1996)	Fishing	Dichotomous CVM Individual TCM	0–577	57-64 14-29
Kirchhoff et al. (1997)	Whitewater rafting Bird watching	Non-dichotomous payment card CVM	36–56 35–69	87–210 2–35
Bowker et al. (1997) Pooled data (n−1) Pooled data (all)	Whitewater rafting	Count data TCM	-	 14–160 16–57
Brouwer and Spaninks (1999)	Biodiversity	CVM	27–36	22–40
Morrison and Bennett (2000)	Wetlands	CE	4–191	-
Rosenberger and Loomis (2000) Meta analysis	Recreation	CVM RUM TCM Ind. TCM Zonal TCM	-	0–319
Piper and Martin (2001) Individual sites (similar) Individual sites (dissimilar) Pooled data	Rural water supply	CVM	 - - -	 6–20 89–149 3–23
VandenBerg et al. (2001) Individual sites Pooled data (multi-state) Pooled data	Water quality	CVM	 1–239 0–105 3–57	 0–298 1–56 0–39

(state-level) Pooled data (contaminated sites)			3–100	2–50
Shrestha and Loomis (2001)	International recreation	CVM and/or TCM	-	1–81
Bergland et al. (2002)	Water quality	Iterative Bidding CVM	25–45	18–41
Engel (2002) Benefit function transfer Meta-analysis transfer	Recreation/habitat	CVM TCM	- -	2–475 3–7028
Morrison et al.(2002)	Wetlands	Discrete -CM Stated Preference	4-66	-
Chattopadhyay (2003) N=304 (similar subgroups) N=609 (similar subgroups) N=913 (similar subgroups) N=1218 (similar subgroups) N=1522 (similar subgroups) N=913 (dissimilar subgroups)	Air quality	Hedonic Method	106–429 57–150 42–82 36–67 32–58 89–128	104–486 57–153 42–82 36–67 32–58 65–110
Ready et al. (2004)	International air and water quality (health benefits)	CVM	20–81	20–83
Jeong and Haab (2004) Access per trip Per one fish increase	Marine recreational fishing	Two Stage Nested RUM	- -	4–230 2–457
Rozan (2004)	International air quality (health benefits)	CVM	-	16–30
Jiang et al. (2005)	Coastal land protection	CE	-	53–85
Groothuis (2005)	Deer Hunting for	CVM	0.7-223	1-205

	outdoor recreation	TCM	2-301	2-292
Hanley et al 2006 WTP-implicit price	River ecology	CE	- -	87-88 87-88
Colombo et al. 2007	Off site impact of soil erosion	CE	0.13-20 4-28	8-1401 8-4574
Zanderson et al.(2007)	Forest Recreation	TC RUM	1.3-1110.6	-
Colombo and Hanley (2008)	Farming	CE	30-1141	2-2423

Source: Adapted from and expanded on Rosenberger and Stanley (2006)

Appendix B: Benefit Transfer under Scale Difference

Table B1: Preference coefficient with increasing sample size

True Preference		P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
		0.5	**0.4**	**0.3**	**0.3**	**0.5**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**
Estimated Preference Coefficients											
N	σ_s										
100	2	0.581*	0.439	0.389	0.283	0.565	0.667**	0.430	0.383	0.090	0.184
		(.064)	(.296)	(.392)	(.33)	(.288)	(.33)	(.528)	(.503)	(.488)	(.421)
1000	2	0.580*	0.469*	0.355*	0.306*	0.530*	0.553*	0.632*	0.476*	0.152	0.189
		(.02)	(.085)	(.115)	(.095)	(.102)	(.088)	(.14)	(.134)	(.144)	(.12)
100	5	0.293*	0.218	0.132	0.183	0.249	0.329	0.270	0.212	0.057	0.087
		(.036)	(.265)	(.341)	(.295)	(.27)	(.266)	(.472)	(.425)	(.417)	(.378)
1000	5	0.291*	0.227*	0.151	0.171	0.249*	0.304*	0.318*	0.244**	0.110	0.112
		(.011)	(.082)	(.102)	(.09)	(.094)	(.079)	(.121)	(.12)	(.129)	(.113)
100	10	0.171*	0.115	0.050	0.117	0.127	0.191	0.142	0.153	0.015	0.053
		(.022)	(.261)	(.326)	(.293)	(.273)	(.247)	(.449)	(.405)	(.393)	(.369)
1000	10	0.170*	0.119	0.077	0.100	0.136	0.187**	0.189	0.124	0.087	0.068
		(.007)	(.082)	(.098)	(.09)	(.093)	(.076)	(.116)	(.117)	(.124)	(.112)

Notes:
Standard errors are shown in brackets.
Preference coefficients and standard errors are the mean of coefficients and standard errors estimated over 2000 replications.
*: denotes significance at the 1% level.
**: denotes significance at the 5% level.
$\sigma_{s:}$ scale at the study site.
N: sample size

Table B2: Bias (Measured as the difference between true MRS* and estimated MRS)**

N	σ_s	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
100	2	0	0.415	0.552	0.464	0.417	0.472	0.786	0.724	0.776	0.615
1000		0	0.121	0.164	0.145	0.162	0.255	0.201	0.186	0.357	0.177
100	5	0	0.732	0.991	0.845	0.768	0.756	1.317	1.192	1.207	1.074
1000		0	0.222	0.292	0.257	0.286	0.252	0.345	0.331	0.401	0.312
100	10	0	1.243	1.628	1.423	1.320	1.169	2.102	1.952	1.896	1.755
1000		0	0.391	0.473	0.414	0.467	0.354	0.564	0.562	0.567	0.529

Notes:

1. *: Ratio of any true preference coefficients on any two attributes. In this study, true MRS is taken as the ratio of coefficient on any attributes with respect to coefficient on P.

2. **: Ratio of any estimated preference coefficients on any two attributes. Estimated MRS is taken as the ratio of estimated coefficients on any attributes with respect to coefficient on P.

Bias= $\frac{\beta}{\beta_1} - \frac{\hat{\beta}}{\hat{\beta}_1}$ where β is a vector of true preference coefficients. β_1 is a true preference coefficient on P. $\hat{\beta}$ is a vector of estimated preference coefficients. $\hat{\beta}_1$ is an estimated coefficient of attribute P.

Table B3: Probability of rejecting null hypothesis (Measured as no difference between true MRS and estimated MRS)

N	σ_s	Probability of rejecting null hypothesis
100	2	0.099
000		0.112
100	5	0.032
1000		0.082
100	10	0.014
1000		0.079

Table B4: Probability of rejecting null hypothesis of WTP (measured as no difference between estimated WTP obtained by transferring function the study site and true WTP at the policy site)

Policy Option → / Study & Policy Context			Change in a binary attribute		Change in all binary attributes		Change in a continuous attribute		Change in all continuous attributes		Loss of site	
N	σ_s	↓ σ_p	$pstat^{1}_{E(CV)}$	$pstat^{1}_{A(CV)}$	$pstat^{all}_{E(CV)}$	$pstat^{all}_{A(CV)}$	$pstat^{1}_{E(CV)}$	$pstat^{1}_{A(CV)}$	$pstat^{all}_{E(CV)}$	$pstat^{all}_{A(CV)}$	$pstat^{Loss}_{E(CV)}$	$pstat^{Loss}_{A(CV)}$
100	2	2	0.086	0.074	0.072	0.066	0.086	0.070	0.060	0.068	0.021	0.022
		5	0.086	0.088	0.072	0.068	0.086	0.078	0.060	0.064	0.023	0.027
		10	0.086	0.096	0.072	0.074	0.086	0.070	0.060	0.068	0.024	0.035
1000	2	2	0.089	0.091	0.085	0.081	0.090	0.089	0.081	0.085	0.023	0.029
		5	0.089	0.096	0.085	0.088	0.090	0.093	0.081	0.087	0.027	0.035
		10	0.089	0.103	0.085	0.101	0.090	0.097	0.081	0.090	0.032	0.039
100	5	2	0.062	0.066	0.060	0.062	0.066	0.064	0.062	0.060	0.019	0.021
		5	0.062	0.078	0.060	0.067	0.066	0.074	0.062	0.062	0.020	0.022
		10	0.062	0.068	0.060	0.074	0.066	0.074	0.062	0.064	0.027	0.028
1000	5	2	0.078	0.076	0.073	0.076	0.078	0.074	0.072	0.077	0.020	0.029
		5	0.078	0.081	0.073	0.082	0.078	0.083	0.072	0.080	0.022	0.032
		10	0.078	0.087	0.073	0.087	0.078	0.089	0.072	0.083	0.028	0.033
100	10	2	0.059	0.061	0.058	0.059	0.058	0.066	0.057	0.062	0.018	0.021
		5	0.059	0.069	0.058	0.062	0.058	0.064	0.057	0.066	0.020	0.023
		10	0.059	0.075	0.058	0.070	0.058	0.078	0.057	0.075	0.023	0.026
1000	10	2	0.066	0.069	0.061	0.067	0.069	0.072	0.067	0.070	0.019	0.025
		5	0.066	0.074	0.061	0.069	0.069	0.075	0.067	0.074	0.024	0.029
		10	0.066	0.078	0.061	0.074	0.069	0.079	0.067	0.077	0.025	0.033

Notes:

$pstat^{1}_{E(CV)}$: probability of rejecting null hypothesis of equality of estimated expected WTP (calculated by transferring function from the study site) and true expected WTP for a change in one binary and one continuous attribute respectively at the policy site.

$pstat^{1}_{A(CV)}$: probability of rejecting null hypothesis of equality of estimated actual WTP (calculated by transferring function from the study site) and true actual WTP for a change in one binary and one continuous attribute respectively at the policy site.

$pstat_{E(CV)}^{all}$: probability of rejecting null hypothesis of equality of estimated expected WTP (calculated by transferring function from the study site) and true expected WTP for a change in all binary and continuous attributes respectively at the policy site.

$pstat_{A(CV)}^{all}$: probability of rejecting null hypothesis of equality of estimated actual WTP (calculated by transferring function from the study site) and true actual WTP for a change in all binary and continuous attributes respectively at the policy site.

$pstat_{E(CV)}^{Loss}$: probability of rejecting null hypothesis of equality of estimated expected WTP (calculated by transferring function from the study site) and true expected WTP for a loss of site at the policy site.

$pstat_{A(CV)}^{Loss}$: probability of rejecting null hypothesis of equality of estimated actual WTP (calculated by transferring function from the study site) and true actual WTP for a loss of site at the policy site.

Table B5: Transfer error on WTP

Policy Option → / Study & Policy Context ↓			Change in a binary attribute		Change in all binary attributes		Change in a continuous attribute		Change in all continuous attributes		Loss of site	
N	σ_s	σ_p	$TE^1_{E(CV)}$	$TE^1_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^1_{E(CV)}$	$TE^1_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{loss}_{E(CV)}$	$TE^{loss}_{A(CV)}$
100	2	2	39.874	40.023	18.695	19.686	81.083	80.280	57.557	56.611	25.928	26.072
		5	39.874	38.380	18.695	18.405	81.083	80.150	57.557	57.467	23.922	17.874
		10	39.874	38.685	18.695	18.448	81.083	79.732	57.557	57.585	26.454	15.079
1000	2	2	20.919	20.580	10.007	10.151	21.530	21.382	14.513	14.253	6.762	7.338
		5	20.919	20.097	10.007	9.454	21.530	21.068	14.513	14.247	8.153	6.528
		10	20.919	20.583	10.007	9.273	21.530	20.667	14.513	14.132	10.296	8.733
100	5	2	69.101	67.294	32.652	32.355	136.451	137.653	93.570	94.456	28.769	34.626
		5	69.101	64.809	32.652	32.722	136.451	132.357	93.570	93.851	72.452	50.802
		10	69.101	63.779	32.652	31.795	136.451	132.325	93.570	95.343	78.381	45.618
1000	5	2	19.417	19.227	9.886	9.873	36.515	36.574	23.364	23.501	20.553	18.393
		5	19.417	20.444	9.886	10.851	36.515	35.776	23.364	23.030	25.014	22.591
		10	19.417	20.355	9.886	11.107	36.515	35.115	23.364	22.841	31.933	34.428
100	10	2	112.937	110.017	54.450	54.034	213.950	216.333	142.444	144.658	44.436	50.947
		5	112.937	102.680	54.450	51.894	213.950	209.493	142.444	145.062	88.231	72.026
		10	112.937	102.313	54.450	52.506	213.950	208.044	142.444	146.096	278.256	123.560
1000	10	2	32.042	30.550	17.243	17.140	59.266	59.380	40.381	40.660	23.068	26.946
		5	32.042	29.498	17.243	17.611	59.266	58.273	40.381	40.033	49.855	48.793
		10	32.042	30.594	17.243	18.516	59.266	56.066	40.381	39.261	95.986	88.356

Notes:

$TE^1_{E(CV)}$: mean transfer error of 2000 iterations measured as the absolute percentage difference between estimated expected WTP (calculated by transferring the function from the study site) and true expected WTP for a change in one binary and continuous attribute respectively at the policy site.

$TE^1_{A(CV)}$: mean transfer error of 2000 iterations measured as the absolute percentage difference between estimated actual WTP (calculated by transferring the function from the study site) and true actual WTP for a change in one binary and one continuous attribute respectively at the policy site.

$TE^{all}_{E(CV)}$: mean transfer error of 2000 iterations measured as the absolute percentage difference between estimated expected WTP (calculated by transferring the function from the study site) and true expected WTP for a change in all binary and continuous attributes respectively at the policy site.

$TE_{A(CV)}^{all}$: mean transfer error of 2000 iterations measured as the absolute percentage difference between estimated actual WTP (calculated by transferring the function from the study site) and true actual WTP for a change in all binary and continuous attributes respectively at the policy site.
$TE_{E(CV)}^{loss}$: mean transfer error of 2000 iterations measured as the absolute percentage difference between estimated expected WTP (calculated by transferring the function from the study site) and true expected WTP for a loss of site at the policy site.
$TE_{A(CV)}^{loss}$: mean transfer error of 2000 iterations measured as the absolute percentage difference between estimated actual WTP (calculated by transferring the function from the study site) and true actual WTP for a loss of site at the policy site.

Appendix C: Benefit Transfer under Substitution Possibilities

Table C1: Preference coefficient with increasing sample size

True Preference		P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
		0.5	**0.4**	**0.3**	**0.3**	**0.5**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**
N	C_s	Estimated Preference Coefficients									
100	16	0.363*	0.200	-0.480	0.570	-0.228	0.832	1.873	1.621	0.393	2.602**
		(.053)	(.682)	(.67)	(.996)	(.582)	(.82)	(.969)	(1.116)	(.93)	(1.293)
1000	16	0.387*	-0.058	0.118	0.283	0.487**	0.618*	0.309	-0.230	0.732	0.454
		(.017)	(.183)	(.17)	(.241)	(.237)	(.172)	(.195)	(.327)	(.444)	(.264)
100	32	0.633*	0.445	-0.476	0.999**	0.811**	0.948	1.400**	0.667	0.700	0.791
		(.078)	(.404)	(.573)	(.45)	(.339)	(.527)	(.632)	(.793)	(.695)	(.467)
1000	32	0.540*	0.654*	-0.166	0.504*	0.775*	0.410*	0.461*	0.470**	0.631*	0.419**
		(.021)	(.149)	(.197)	(.118)	(.128)	(.156)	(.17)	(.197)	(.222)	(.167)
100	64	0.593*	0.220	0.276	0.664	0.398	1.042*	0.583	-0.051	1.270**	0.255
		(.067)	(.313)	(.411)	(.362)	(.291)	(.346)	(.55)	(.48)	(.506)	(.425)
1000	64	0.627*	0.509*	0.295**	0.472*	0.587*	0.708*	0.787*	0.505*	0.043	0.299**
		(.022)	(.088)	(.121)	(.097)	(.106)	(.092)	(.143)	(.135)	(.148)	(.123)

Notes:
Standard errors are shown in brackets.
Preference coefficients and standard errors are the mean of coefficients and standard errors estimated over 2000 replications.
*: denotes significance at the 1% level.
**: denotes significance at the 5% level.
C_s: Choice set at the study site.
N: Sample size.

Table C2: Bias (Measured as the difference between true MRS and estimated MRS)

N	C_s	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
100	16	0	0.249	1.923	0.971	1.629	1.095	4.165	3.671	0.483	6.775
1000	16	0	0.949	0.295	0.131	0.256	0.396	0.202	1.394	1.290	0.771
100	32	0	0.097	1.352	0.978	0.281	0.298	1.210	0.253	0.505	0.849
1000	32	0	0.410	0.907	0.333	0.436	0.441	0.147	0.070	0.568	0.377
100	64	0	0.430	0.136	0.519	0.330	0.556	0.017	0.886	1.541	0.030
1000	64	0	0.012	0.130	0.154	0.064	0.071	0.256	0.006	0.532	0.077

Note: Bias is measured similarly as explained in notes of Table B2.

Table C3: Transfer error on WTP

Policy Option → / Study & Policy Context ↓			Change in a binary attribute		Change in all binary attributes		Change in a continuous attribute		Change in all continuous attributes		Loss of a site	
N	C_s	C_p	$TE^1_{E(CV)}$	$TE^1_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^1_{E(CV)}$	$TE^1_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{loss}_{E(CV)}$	$TE^{loss}_{A(CV)}$
100	16	16	91.185	90.350	30.328	35.758	478.614	481.732	538.545	534.463	29.716	26.9719
		32	103.647	116.501	22.899	25.638	528.395	502.124	595.517	579.400	7.8695	3.5762
		64	110.347	165.167	23.722	20.484	520.639	544.844	628.937	627.448	24.567	5.825
1000	16	16	32.979	32.269	15.901	16.864	11.943	13.627	20.286	22.046	8.184	13.794
		32	37.139	39.511	12.909	13.727	15.870	15.968	24.519	24.469	25.589	17.162
		64	43.038	48.424	11.402	12.219	19.102	20.331	13.091	10.495	34.461	33.572
100	32	16	40.908	40.738	13.037	8.426	139.685	130.998	96.700	96.833	32.403	39.131
		32	41.566	41.898	18.007	13.636	131.586	128.435	96.863	94.882	48.223	18.097
		64	37.137	36.153	13.300	14.179	134.345	131.430	102.160	95.787	61.793	27.670
1000	32	16	11.732	9.740	12.875	12.657	15.785	16.464	37.963	37.050	8.407	8.058
		32	9.198	5.441	14.338	14.322	14.169	14.374	32.441	32.536	11.188	4.074
		64	2.957	0.973	12.971	14.571	13.495	12.234	30.027	29.394	11.615	9.776
100	64	16	39.344	34.506	0.371	2.463	6.049	7.695	27.932	25.695	21.475	29.175
		32	32.103	34.502	3.828	5.392	8.151	8.730	24.775	24.893	33.257	14.526
		64	22.536	23.701	5.658	9.904	6.352	5.165	27.985	32.056	42.079	17.796
1000	64	16	6.369	5.288	0.010	0.291	23.775	24.207	5.423	3.995	20.411	21.274
		32	4.981	5.049	0.194	0.336	27.141	25.733	8.609	8.787	2.961	4.616
		64	12.648	11.676	3.547	3.104	28.370	28.284	2.998	1.933	0.189	0.356

Notes: Symbols have usual meaning as explained in notes of Table B5.
C_p: Choice set at the policy site.

Appendix D: Benefit Transfer under Characteristics Difference

Case I: Function at the study site is missing one important variable at the policy site

Table D1: Preference coefficient with increasing sample size

True Preference	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
	0.5	**0.4**	**0.3**	**0.3**	**0.5**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**
N	Estimated preference coefficients in Ideal benefit transfer									
100	0.5812	0.438	0.387	0.275	0.562	0.668	0.424	0.3939	0.0626	0.1834
	(.064)	(.296)	(.391)	(.33)	(.288)	(.33)	(.528)	(.503)	(.487)	(.421)
500	0.5791	0.4234	0.4051	0.3002	0.5292	0.5574	0.5778	0.2592	0.1985	0.3303
	(.029)	(.123)	(.182)	(.134)	(.138)	(.141)	(.189)	(.198)	(.213)	(.206)
1000	0.5807	0.4667	0.3589	0.3092	0.5255	0.5544	0.6272	0.4714	0.1509	0.1837
	(.02)	(.085)	(.116)	(.095)	(.102)	(.089)	(.14)	(.134)	(.144)	(.12)
	Estimated preference coefficients in Case I of benefit transfer									
100	0.610*	-0.060	0.340	0.714**	0.387	-	1.004**	0.133	1.326**	0.110
	(.069)	(.278)	(.351)	(.277)	(.271)	-	(.503)	(.456)	(.462)	(.392)
500	0.642*	0.394*	0.317	0.458*	0.904*	-	-0.097	0.219	0.319	0.191
	(.032)	(.11)	(.174)	(.114)	(.136)	-	(.18)	(.196)	(.205)	(.21)
1000	0.627*	0.458*	0.317*	0.520*	0.770*	-	0.744*	0.495*	0.149	0.371*
	(.022)	(.079)	(.109)	(.085)	(.097)	-	(.145)	(.134)	(.136)	(.116)

Notes:
Standard errors are shown in brackets.
Preference coefficients and standard errors are the mean of coefficients and standard errors estimated over 2000 replications.
*: denotes significance at the 1% level.
**: denotes significance at the 5% level.
Ideal: No difference between the study and policy site.
Case I: X_5 is present at policy site but not observed in the study site.

Table D2: Bias (Measured as the difference between true MRS and estimated MRS)

Scenario	*N*	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Ideal	100	0	0.423	0.577	0.464	0.415	0.477	0.764	0.719	0.784	0.608
	500	0	0.185	0.260	0.195	0.211	0.280	0.270	0.410	0.361	0.312
	1000	0	0.122	0.166	0.144	0.164	0.253	0.202	0.193	0.360	0.176
Case I	100	0	0.899	0.042	0.571	0.565	-	0.647	0.581	1.574	0.219
	500	0	0.186	0.106	0.113	0.208	-	1.151	0.458	0.103	0.103
	1000	0	0.069	0.095	0.228	0.028	-	0.187	0.011	0.363	0.191

Note: Bias is measured similarly as explained in notes of Table B2.

Table D3: Transfer error on WTP

Policy Option → / Study & Policy Context ↓		Change in a binary attribute		Change in all binary attributes		Change in a continuous attribute		Change in all continuous attributes		Loss of site	
Scenario	*N*	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{loss}_{E(CV)}$	$TE^{loss}_{A(CV)}$
Ideal	100	40.283	40.287	18.638	19.621	78.622	77.872	57.080	56.114	26.196	26.418
	500	23.267	23.187	12.208	12.560	28.959	28.600	24.697	24.360	8.605	9.053
	1000	20.851	20.502	10.125	10.277	21.715	21.582	15.111	14.825	6.523	7.040
Case I	100	115.049	118.922	45.964	49.388	73.564	71.794	1.644	5.380	17.845	2.311
	500	7.262	5.640	15.802	15.721	70.766	70.475	61.971	61.216	1.340	7.395
	1000	7.962	19.485	15.057	14.420	24.870	25.596	43.189	43.768	30.908	30.894

Note: Symbols have the usual meaning as described in notes of Table B5.

Case II: Data is not available for one of the important variables contained in function transferred from the study site

Table D4: Preference coefficient with increasing sample size

True Preference	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
	0.5	**0.4**	**0.3**	**0.3**	**0.5**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**
N	Estimated Preference Coefficients in Ideal Case of benefit transfer									
100	0.593	0.220	0.276	0.664	0.398	1.042	0.583	-0.051	1.270	0.255
	(.067)	(.313)	(.411)	(.362)	(.291)	(.346)	(.55)	(.48)	(.506)	(.425)
500	0.619	0.426	0.289	0.410	0.787	0.528	0.061	0.133	0.267	-0.042
	(.031)	(.12)	(.186)	(.137)	(.14)	(.14)	(.19)	(.198)	(.212)	(.213)
1000	0.627	0.509	0.295	0.472	0.587	0.708	0.787	0.505	0.043	0.299
	(.022)	(.088)	(.121)	(.097)	(.106)	(.092)	(.143)	(.135)	(.148)	(.123)
	Estimated Preference Coefficients in Case II of benefit transfer									
100	0.593*	0.220	0.276	0.664	0.398	1.042*	0.583	-0.051	1.270*	0.255
	(.067)	(.313)	(.411)	(.362)	(.291)	(.346)	(.55)	(.48)	(.506)	(.425)
500	0.619*	0.426*	0.289**	0.410**	0.787*	0.528*	0.061	0.133	0.267	0.030
	(.032)	(.11)	(.174)	(.114)	(.136)	(.18)	(.196)	(.205)	(.21)	(.21)
1000	0.627*	0.509*	0.295**	0.472*	0.587*	0.708*	0.787*	0.505*	0.043	0.299**
	(.022)	(.088)	(.121)	(.097)	(.106)	(.092)	(.143)	(.135)	(.148)	(.123)

Notes:
Standard errors are shown in brackets.
Preference coefficients and standard errors are the mean of 2000 iterations.
* denotes significance at the 1% level.
** denotes significance at the 5% level.
Case II: is present in both study and policy site but data on a X_5 is not available at the policy site under benefit transfer application.

Table D5: Bias (Measured as the difference between true MRS and estimated MRS)

Scenario	*N*	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Ideal	100	0	0.430	0.136	0.519	0.330	0.556	0.017	0.886	1.541	0.030
	500	0	0.112	0.134	0.062	0.270	0.348	0.901	0.585	0.168	0.468
	1000	0	0.012	0.130	0.154	0.064	0.071	0.256	0.006	0.532	0.077
Case II	100	0	0.430	0.136	0.519	0.330	0.556	0.017	0.886	1.541	0.030
	500	0	0.112	0.134	0.062	0.270	0.348	0.901	0.585	0.168	0.468
	1000	0	0.012	0.130	0.154	0.064	0.071	0.256	0.006	0.532	0.077

Note: Bias is measured similarly as described in notes of Table B2.

Table D6: Transfer error on WTP

Policy Option → / Study & Policy Context ↓		Change in a binary attribute		Change in all binary attributes		Change in a continuous attribute		Change in all continuous attributes		Loss of site	
Scenario	*N*	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{loss}_{E(CV)}$	$TE^{loss}_{A(CV)}$
Ideal	100	22.536	23.701	5.658	9.904	90.495	90.525	76.282	76.116	42.079	17.796
	500	28.975	25.682	3.503	2.239	28.370	28.284	27.985	32.056	12.135	0.775
	1000	12.648	11.676	3.547	3.104	6.352	5.165	2.998	1.933	0.189	0.356
Case II	100	44.751	29.890	31.459	32.785	4.926	0.436	26.543	29.701	4.044	6.636
	500	3.617	4.533	18.281	18.285	90.739	90.922	76.478	76.184	12.023	13.332
	1000	20.013	32.534	21.734	20.906	29.021	28.332	2.713	1.703	30.550	30.060

Note: Symbols have usual meaning as described in notes of Table B5.

Appendix E: Benefit Transfer under Preference Difference

Table E1: Preference coefficient with increasing sample size

True preferences at policy site	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
	0.5	**0.4**	**0.3**	**0.3**	**0.5**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**
N	Estimated Preference Coefficients Ideal Case									
100	0.593*	0.220	0.276	0.664	0.398	1.042*	0.583	-0.051	1.270**	0.255
	(.067)	(.313)	(.411)	(.362)	(.291)	(.346)	(.55)	(.48)	(.506)	(.425)
500	0.619*	0.426*	0.289	0.410*	0.787*	0.528*	0.061	0.133	0.267	-0.042
	(.031)	(.12)	(.186)	(.137)	(.14)	(.14)	(.19)	(.198)	(.212)	(.213)
1000	0.627*	0.509*	0.295**	0.472*	0.587*	0.708*	0.787*	0.505*	0.043	0.299**
	(.022)	(.088)	(.121)	(.097)	(.106)	(.092)	(.143)	(.135)	(.148)	(.123)
True preferences[c] at study site (Case I)	**0.5**	**0.4**	**0.3**	**0.6**	**0.5**	**0.6**	**0.5**	**0.4**	**0.3**	**0.6**
100	0.572*	0.207	0.106	1.136*	0.335	0.890*	0.578	0.115	1.222**	0.592
	(.065)	(.305)	(.421)	(.361)	(.287)	(.332)	(.537)	(.479)	(.506)	(.421)
500	0.608*	0.408*	0.282	0.755*	0.785*	0.519*	0.060	0.181	0.302	0.453**
	(.03)	(.12)	(.192)	(.141)	(.141)	(.141)	(.184)	(.2)	(.212)	(.22)
1000	0.632*	0.488*	0.241	0.847*	0.585*	0.704*	0.733*	0.413*	0.089	0.784*
	(.022)	(.088)	(.125)	(.1)	(.107)	(.093)	(.14)	(.135)	(.148)	(.126)
True preferences[d] at study site (Case II)	**0.5**	**0.4**	**0.3**	**0.3**	**0.5**	**0.2**	**0.2**	**0.4**	**0.3**	**0.2**
100	0.586*	-0.044	0.675	0.421	0.227	0.689**	0.735	0.128	1.065**	0.166
	(.066)	(.306)	(.39)	(.335)	(.276)	(.324)	(.524)	(.465)	(.465)	(.406)
500	0.664*	0.372*	0.245	0.523*	0.803*	0.076	-0.295	0.199	0.251	0.245
	(.033)	(.116)	(.182)	(.139)	(.139)	(.133)	(.189)	(.199)	(.209)	(.212)
1000	0.626*	0.481*	0.261**	0.531*	0.614*	0.297*	0.403*	0.467*	0.063	0.276**
	(.022)	(.085)	(.116)	(.096)	(.101)	(.087)	(.14)	(.133)	(.144)	(.12)

Notes:
Standard errors are shown in brackets.
Preference coefficients and standard errors are the mean of 2000 iterations.
[c]: X_3 and X_9 are replaced by the highest preference coefficients of X_5 at the study site.
[d]: X_5 and X_6 are replaced by the lowest preference coefficients of X_9 at the study site.

Case I: True preferences at the study and policy site are different. True preferences on variables X_3 and X_9 are replaced by the highest preference coefficients of X_5 at the study site.
Case II: True preferences at the study and policy site are different. True preferences on variables X_5 and X_6 are replaced by the lowest preference coefficients of X_9 at the study site.
* denotes significance at the 1% level.
** denotes significance at the 5% level.

Table E2: Bias (Measured as the difference between true MRS and estimated MRS)

Scenario	*N*	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Ideal	100	0	0.430	0.136	0.519	0.330	0.556	0.017	0.886	1.541	0.030
	500	0	0.112	0.134	0.062	0.270	0.348	0.901	0.585	0.168	0.468
	1000	0	0.012	0.130	0.154	0.064	0.071	0.256	0.006	0.532	0.077
Case I	100	0	0.438	0.415	1.385	0.415	0.355	0.009	0.600	1.536	0.634
	500	0	0.129	0.136	0.641	0.290	0.348	0.901	0.503	0.103	0.345
	1000	0	0.028	0.218	0.740	0.075	0.085	0.161	0.147	0.460	0.841
Case II	100	0	0.874	0.550	0.118	0.613	0.025	0.253	0.581	1.215	0.117
	500	0	0.240	0.231	0.187	0.209	1.085	1.445	0.501	0.222	0.031
	1000	0	0.032	0.183	0.247	0.020	0.725	0.356	0.054	0.500	0.041

Note: Bias is measured similarly as described in Table B2.

Table E3: Transfer error on WTP

Policy Option → / Study & Policy Context ↓		Change in a binary attribute		Change in all binary attributes		Change in a continuous attribute		Change in all continuous attributes		Loss of site	
Scenario	N	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{loss}_{E(CV)}$	$TE^{loss}_{A(CV)}$
Ideal	100	28.975	25.682	5.658	9.904	90.495	90.525	76.282	76.116	42.0792	17.7957
	500	22.536	23.701	3.503	3.104	28.370	28.284	27.985	32.056	12.135	0.7753
	1000	12.648	11.676	3.547	2.239	6.352	5.165	2.998	1.933	0.1892	0.3564
Case I	100	4.348	4.711	6.152	0.021	3.536	2.757	56.265	57.444	59.816	29.154
	500	35.992	32.134	8.924	10.051	90.528	90.516	44.583	43.771	35.810	14.493
	1000	16.054	14.162	8.482	9.140	19.529	19.772	18.831	19.524	4.290	3.514
Case II	100	0.230	8.388	36.115	40.683	23.131	22.405	30.553	32.810	16.387	3.457
	500	89.269	88.771	18.780	18.549	139.653	138.809	80.909	79.663	2.107	8.952
	1000	59.700	58.688	11.478	10.818	37.064	36.571	30.529	30.254	13.182	14.292

Note: Symbols have usual meaning as described in notes of Table B5.

Appendix F: Benefit Transfer under Measurement error

Table F1: Preference coefficient with increasing sample size

True pref. at policy site	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
	0.5	**0.4**	**0.3**	**0.3**	**0.5**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**
N	Estimated Preference Coefficients in Ideal Case									
100	0.581*	0.438	0.387	0.275	0.562	0.668	0.424	0.394	0.063	0.183
	(.064)	(.296)	(.391)	(.33)	(.288)	(.33)	(.528)	(.503)	(.487)	(.421)
500	0.579*	0.423*	0.405**	0.300**	0.529*	0.557*	0.578*	0.259	0.199	0.330
	(.029)	(.123)	(.182)	(.134)	(.138)	(.141)	(.189)	(.198)	(.213)	(.206)
1000	0.581*	0.467*	0.359*	0.309*	0.526*	0.554*	0.627*	0.471*	0.151	0.184
	(.02)	(.085)	(.116)	(.095)	(.102)	(.089)	(.14)	(.134)	(.144)	(.12)
	Estimated Preference Coefficients in Case I									
100	0.575*	0.689**	0.075**	0.608**	0.709	0.188	-1.998	0.497	0.027	0.266
	(.065)	(.266)	(.354)	(.284)	(.278)	(.247)	(6.056)	(.524)	(.467)	(.411)
500	0.584*	0.622*	0.229*	0.524*	0.713	0.066	0.695	0.149	0.369	0.377
	(.029)	(.113)	(.17)	(.122)	(.134)	(.114)	(1.953)	(.205)	(.211)	(.223)
1000	0.565*	0.617*	0.057*	0.571*	0.774	-0.042	-0.950	0.512*	0.098	0.131
	(.02)	(.081)	(.106)	(.089)	(.096)	(.075)	(1.51)	(.147)	(.15)	(.121)

Notes:
Standard errors are shown in brackets.
Preference coefficients and standard errors are the mean of 2000 iterations.
* denotes significance at the 1% level.
** denotes significance at the 5% level.
Case I: Variable X_5 is measured with error.

Table F2: Bias (Measured as the difference between true MRS and estimated MRS)

Scenario	*N*	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Ideal Case	100	0	0.423	0.577	0.464	0.415	0.477	0.764	0.719	0.784	0.608
	500	0	0.185	0.260	0.195	0.211	0.280	0.270	0.410	0.361	0.312
	1000	0	0.122	0.166	0.144	0.164	0.253	0.202	0.193	0.360	0.176
Case I	100	0	0.522	0.655	0.548	0.450	0.879	9.312	0.793	0.830	0.592
	500	0	0.281	0.292	0.310	0.266	1.085	2.815	0.568	0.289	0.363
	1000	0	0.296	0.498	0.411	0.374	1.274	3.155	0.233	0.439	0.217

Note: Bias is calculated similarly as described in Table B2.

Table F3: Transfer error on WTP

Policy Option → / Study & Policy Context ↓		Change in a binary attribute		Change in all binary attributes		Change in a continuous attribute		Change in all continuous attributes		Loss of site	
Scenario	*N*	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{loss}_{E(CV)}$	$TE^{loss}_{A(CV)}$
Ideal Case	100	40.283	40.287	18.638	19.621	78.622	77.872	57.080	56.114	26.196	26.418
	500	23.267	23.187	12.208	12.560	28.959	28.600	24.697	24.360	8.605	9.053
	1000	20.851	20.502	10.125	10.277	21.715	21.582	15.111	14.825	6.523	7.040
Case I	100	71.046	71.325	20.744	21.951	670.407	683.256	233.350	238.830	90.838	81.328
	500	90.285	90.209	9.236	9.531	297.837	295.433	122.666	122.402	24.542	26.230
	1000	106.729	106.783	5.922	6.474	238.827	235.041	106.602	105.987	14.402	15.224

Note: Symbols have usual meaning as described in notes of Table B5

Appendix G: Benefit Transfer under Random Parameter versus Fixed Parameter

Table G1: Preference coefficient with increasing sample size

True preferences at policy site	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
	0.5	**0.4**	**0.3**	**0.3**	**0.5**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**
N	Estimated Preference Coefficients in Ideal Case of benefit transfer									
100	0.581*	0.438	0.387	0.275	0.562	0.668	0.424	0.394	0.063	0.183
	(.064)	(.296)	(.391)	(.33)	(.288)	(.33)	(.528)	(.503)	(.487)	(.421)
500	0.579*	0.423*	0.405**	0.300**	0.529*	0.557*	0.578*	0.259	0.199	0.330
	(.029)	(.123)	(.182)	(.134)	(.138)	(.141)	(.189)	(.198)	(.213)	(.206)
1000	0.581*	0.467*	0.359*	0.309*	0.526*	0.554*	0.627*	0.471*	0.151	0.184
	(.02)	(.085)	(.116)	(.095)	(.102)	(.089)	(.14)	(.134)	(.144)	(.12)
	Estimated Preference Coefficients in Case I of benefit transfer									
100	0.005	0.304	0.252	-0.412	0.447	0.449	-0.069	0.210	0.792**	-0.142
	(.006)	(.245)	(.33)	(.3)	(.27)	(.236)	(.439)	(.393)	(.369)	(.376)
500	0.020*	0.466*	-0.152	-0.146	0.699*	0.318*	0.566*	0.001	0.259	-0.715*
	(.003)	(.104)	(.149)	(.117)	(.122)	(.108)	(.175)	(.179)	(.163)	(.169)
1000	0.015*	0.403*	-0.151	0.040	0.499*	0.265*	0.001	-0.140	-0.018	-0.092
	(.002)	(.075)	(.099)	(.089)	(.09)	(.072)	(.107)	(.115)	(.123)	(.111)

Notes:
Standard errors are shown in brackets.
Preference coefficients and standard errors are the mean of 2000 iterations.
* denotes significance at the 1% level.
** denotes significance at the 5% level.
N: sample size.
Case I : Preference Heterogeneity at the study site but they are fixed at the policy site.

Table G2: Bias (Measured as the difference between true MRS and estimated MRS)

Scenario	*N*	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Ideal Case	100	0	0.423	0.577	0.464	0.415	0.477	0.764	0.719	0.784	0.608
	500	0	0.185	0.260	0.195	0.211	0.280	0.270	0.410	0.361	0.312
	1000	0	0.122	0.166	0.144	0.164	0.253	0.202	0.193	0.360	0.176

Note: Bias is measured similarly as described in Table B2.

Table G3: Transfer error on WTP

Policy Option → / Study & Policy Context ↓		Change in a binary attribute		Change in all binary attributes		Change in a continuous attribute		Change in all continuous attributes		Loss of site	
Scenario	*N*	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{loss}_{E(CV)}$	$TE^{loss}_{A(CV)}$
Ideal Case	100	40.283	40.287	18.638	19.621	78.622	77.872	57.080	56.114	26.196	26.418
	500	23.267	23.187	12.208	12.560	28.959	28.600	24.697	24.360	8.605	9.053
	1000	20.851	20.502	10.125	10.277	21.715	21.582	15.111	14.825	6.523	7.040
Case I	100	8599.47	7395.14	5117.28	4491.23	6920.85	6805.99	7214.179	8095.927	4342.37	3171.99
	500	1300.58	1019.78	1794.23	1634.18	2760.39	2978.27	466.6861	576.782	709.64	544.77
	1000	1504.57	1234.74	2127.91	1974.06	479.06	480.70	685.6136	659.3915	554.77	317.28

Appendix H: Benefit Transfer under more than one Difference between the Study and Policy Sites

Table H1: Preference coefficient with increasing sample size

True preferences at policy site	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
	0.5	**0.4**	**0.3**	**0.3**	**0.5**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**
N	Estimated Preference Coefficients in Ideal Case of benefit transfer									
100	0.581*	0.438	0.387	0.275	0.562	0.668	0.424	0.394	0.063	0.183
	(.064)	(.296)	(.391)	(.33)	(.288)	(.33)	(.528)	(.503)	(.487)	(.421)
500	0.579*	0.423*	0.405**	0.300**	0.529*	0.557*	0.578*	0.259	0.199	0.330
	(.029)	(.123)	(.182)	(.134)	(.138)	(.141)	(.189)	(.198)	(.213)	(.206)
1000	0.581*	0.467*	0.359*	0.309*	0.526*	0.554*	0.627*	0.471*	0.151	0.184
	(.02)	(.085)	(.116)	(.095)	(.102)	(.089)	(.14)	(.134)	(.144)	(.12)
	Estimated Preference Coefficients in Case I									
100	0.594	-0.636	1.005	1.694	-0.620	1.249	5.938	2.324	1.875	-0.282
	(.074)	(.383)	(.483)	(.373)	(.326)	(.365)	(9.203)	(.597)	(.526)	(.49)
500	0.565	1.375	1.020	0.395	0.504	1.518	-3.648	-1.572	0.458	-0.468
	(.03)	(.134)	(.201)	(.12)	(.162)	(.167)	(2.334)	(.274)	(.25)	(.267)
1000	0.566	0.523	-0.002	1.774	0.002	-1.164	5.086	-1.080	2.047	-1.231
	(.022)	(.1)	(.158)	(.122)	(.118)	(.096)	(1.591)	(.174)	(.179)	(.154)
	0.015	0.403	-0.151	0.040	0.499	0.265	0.001	-0.140	-0.018	-0.092
	(.002)	(.075)	(.099)	(.089)	(.09)	(.072)	(.107)	(.115)	(.123)	(.111)

Notes:

Standard errors are shown in brackets

Preference coefficients and standard errors are the mean of 2000 iterations

* denotes significance at the 1% level

** denotes significance at the 5% level

Case I: Benefit transfer under more than one difference such as difference in scale parameter, choice set between the study and policy sites. One of important variable is measured with error in study site but none is measured with error in the policy site. Preference heterogeneity is assumed in the study context but they are assumed to be fixed in the policy context.

Table H2: Bias (Measured as the difference between true MRS and estimated MRS)

Scenario	*N*	P	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Ideal Case	100	0	0.423	0.577	0.464	0.415	0.477	0.764	0.719	0.784	0.608
	500	0	0.185	0.260	0.195	0.211	0.280	0.270	0.410	0.361	0.312
	1000	0	0.122	0.166	0.144	0.164	0.253	0.202	0.193	0.360	0.176

Note: Bias is measured as described in Table B2.

Table H3: Transfer error on WTP

Policy Option → / Study & Policy Context ↓		Change in a binary attribute		Change in all binary attributes		Change in a continuous attribute		Change in all continuous attributes		Loss of site	
Scenario	*N*	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{1}_{E(CV)}$	$TE^{1}_{A(CV)}$	$TE^{all}_{E(CV)}$	$TE^{all}_{A(CV)}$	$TE^{loss}_{E(CV)}$	$TE^{loss}_{A(CV)}$
Ideal Case	100	40.283	40.287	18.638	19.621	78.622	77.872	57.080	56.114	26.196	26.418
	500	23.267	23.187	12.208	12.560	28.959	28.600	24.697	24.360	8.605	9.053
	1000	20.851	20.502	10.125	10.277	21.715	21.582	15.111	14.825	6.523	7.040
Case I	100	753.069	784.039	37.131	31.333	900.842	906.0885	540.8017	544.994	49.451	32.090
	500	478.243	399.290	137.083	142.291	733.4071	738.6107	378.283	397.812	7.113	1.026
	1000	468.255	485.962	100.644	88.102	811.598	804.2729	281.8913	259.171	87.457	51.534

Note: Symbols have usual meaning as described in Table B5.