Benefit Transfer: Choice Experiment Results

Geoffrey N. Kerr
Lincoln University, New Zealand

Basil M.H. Sharp
University of Auckland, New Zealand

Paper presented at the 2004 NZARES Conference
Blenheim Country Hotel, Blenheim, New Zealand. June 25-26, 2004.

Benefit Transfer: Choice Experiment Results

Geoffrey N. Kerr, Lincoln University
Basil M.H. Sharp, University of Auckland

Abstract
Benefit transfer entails using estimates of non-market values derived at one site as approximations to benefits at other sites. The method finds favour because it can be applied quickly and cheaply, however the validity of benefit transfer is frequently questioned. Published studies generally indicate that errors from the approach can be extremely large and could result in significant resource misallocations. Assessing the validity of benefit transfer is complicated by differences in the nature of study and policy sites, the changes being valued, valuation methods, time of study, availability of substitutes and complements, and demographic, social and cultural differences. A choice experiment was used to evaluate the transferability of benefit estimates for identical goods between two different populations. The study design allowed most of the confounding factors to be controlled, so provides a strong test of benefit transfer validity. Several different tests were applied to evaluate benefit transfer validity, with conflicting results. The paper investigates the merits of the alternative tests and concludes that utility functions were different for the two populations.

Key Words: Choice model, Choice experiment, Benefit transfer, Mitigation

The Problem

Choice experiments have the flexibility to value a wide range of potential outcomes. A novel use of this attribute of choice experiments is provided in a study undertaken to evaluate the adequacy of off-site mitigation. The expense, skills and time involved in undertaking choice studies provide ample motivation for benefit transfer. Consequently the study was designed to provide a test of benefit transfer, using two separate populations within a region and two types of stream protection sites to test the possibility of benefit transfer, with the view to using study results across the region should the outcome be favourable.

Development Impacts on Waterways

Every year hundreds of hectares of land in New Zealand's Auckland region are disturbed for transportation, housing, industrial, commercial and community amenity purposes. On-site activities commonly involve construction site earthworks, such as site contouring for residential subdivision development, stream channelisation, armouring and culverting. Impacts include complete loss of waterways (for example, when a stream is piped), and modifications to wildlife habitat, visual amenity and other waterway attributes, as well as off-site impacts, such as sedimentation.

Projects in the Auckland Region involving land disturbance are required to incorporate best practice erosion and sediment controls. Best management practices are not 100% effective and even with appropriately designed and maintained systems in place significant environmental impacts occur.

Mitigation

In addition to requiring best management practices, the Auckland Regional Council has the ability to place conditions on resource consents, including specific offsetting mitigation requirements. Offsetting mitigation may augment stream quality at one site to compensate for the adverse environmental effects associated with development at other sites. Enhancement could occur within the catchment undergoing development and/or possibly in other catchments. The idea is to use mitigation to achieve and sustain desired environmental outcomes.

Requiring the consent holder to provide offset mitigation for the unavoidable damage caused by an activity is well established internationally. Typical examples of offset mitigation include riparian planting and stream bank retirement to offset water quality degradation, planting forests to offset greenhouse gas emissions, and creating or enhancing wetlands or indigenous bush to offset impacts of land drainage.

The method for establishing "appropriate mitigation" in Auckland and generally relies on a "best professional judgement" approach based on ecological indicators. In order for the offset mitigation to function effectively as required by the Resource Management (1991) and Local Government Act (2002) the community needs to have confidence in the mitigation process. However, very little is known about community preferences regarding alternative states of Auckland streams. Without information on community preferences it is not possible for the Auckland Regional Council to identify mitigation that reflects the environmental outcomes the community desires. Consequently, it is highly desirable either to quantify in dollar terms the costs of both the adverse effects at the site of development and the benefits of the offset mitigation, or to identify how the community is willing to trade off site attributes.

Transparent quantification of costs and benefits ensures that the mitigation proposed offers the potential to offset, from both the ecological and the economic perspectives, the adverse effects generated. A choice experiment was employed to identify and evaluate important Auckland stream quality attributes. The following section provides a brief description of study design. It is followed by results and an evaluation of benefits transfer, which leads to discussion and conclusions. Kerr and Sharp (2003) provide full details of study design and results.

Choice Experiment

Choice experiments entail several key steps:

1. Salient attribute identification
2. Choice model design
3. Data collection
4. Data analysis
5. Application to policy

Salient Attribute Identification

Salient attribute identification was undertaken using discussions with Auckland Regional Council personnel, and using focus groups conducted in the two case study communities (South Auckland and North Shore). Focus groups were important to get direct input from the community about their concerns over stream management, their salient attributes, and their willingness to undertake and ability to complete choice questions about stream management. Details of the procedure followed at each focus group meeting are reported in Kerr and Sharp (2002). The focus groups identified the following salient stream attributes.

- Water clarity
- Flow of water
- Quality of the stream bank
- Access
- Safety
- Surrounding land use
- Habitat for wildlife
- Natural shape of the stream

Focus group studies indicated that stream attributes could be described in relatively simple terms that could be understood by the general population. Participants understood the idea of a choice game and were prepared and able to carefully consider the tradeoffs and make meaningful choices. The choice game format used in the focus groups provided the basis for developing the survey questionnaire.

Choice Model Design

Choice models typically employ a linear utility function of the form[1]:

$$V_k = V(Z_k, Y_k) = \beta_0 + \beta_1 Z_{1,k} + \beta_2 Z_{2,k} + \ldots + \beta_n Z_{n,k} + \beta_Y Y_k = \boldsymbol{\beta Z'}$$

Where V is the observable component of utility and the Z_i are choice attributes (or transformations of choice attributes) at the study site. Y is the cost to the individual. The subscript k indexes the choice. Attributes differ between choices, but coefficients in the utility function do not. Data analysis entails selection of the vector of coefficients that maximises the probability of obtaining the observed choices. This

[1] In order to clarify the nature of the changes involved in using a choice experiment to evaluate off-site mitigation, socio-economic effects have been suppressed.

model allows evaluation of on-site mitigation. The overall impact of site changes that degrade some attributes and improve others (including monetary compensation, Y_k) can be identified by comparing prior and posterior utility (V_k). Alternatively, the utility function can be used to identify mitigation that leaves utility unchanged.

The primary study objective was identification of whether off-site attributes could be used as mitigation for specified on-site environmental changes. Consequently, attributes needed to vary simultaneously at two sites. Extending the utility function to incorporate two sites (suppressing $_k$ for clarity) yields:

$$V = \beta_0 + [\beta_{11}Z_{11} + \ldots + \beta_{n1} Z_{n1}] + [\beta_{12}Z_{12} + \ldots + \beta_{n2} Z_{n2}] + \beta_Y Y$$

Where β_{ij} is marginal utility of attribute i at site j and Z_{ij} is the level of attribute i at site j. On-site mitigation requires that a change in an attribute at site 1 (say Z_{11}) is offset by changes in other attributes at site 1 (i.e. by changing attributes Z_{m1} where $m \neq 1$). Off-site mitigation entails changing attributes at the other site. A change in an attribute at site 1 (say Z_{11}) is offset by changes in attributes at site 2 (i.e. by changing attributes Z_{j2} where j includes all attributes at site 2). To identify willingness to trade-off attributes between sites the utility function must include attributes at both sites, effectively doubling the number of attributes in the utility function compared with single site models. While this model form allows identification of off-site mitigation, an extremely useful by-product is the ability to evaluate the adequacy of on-site mitigation (or a mixture of on-site and off-site mitigation) using the same model. Inclusion of the cost attribute (Y) allows monetary measurement of the non-market costs of development impacts and the non-market benefits of stream enhancements. Auckland Regional Council wanted to measure monetary values of impacts, so it was necessary to include a cost attribute in the choices presented to citizens.

Recent choice studies typically incorporate 4-6 attributes. With these numbers of attributes, survey designs are available to estimate interaction effects between the attributes. For example, willingness to pay for additional fish species might be expected to depend upon the amount of fish habitat available, suggesting an interaction between number of fish species and available fish habitat. This study did not allow the possibility of interaction effects of this type. The requirement for attributes to vary at two sites, along with the number of attributes that were identified in the focus groups as being potentially significant, and the requirement for a money cost attribute to allow assessment of money values for site attributes, resulted in selection of the ten choice attributes in Table 1.

Table 1: Choice Attributes

Attribute	**Values: Natural Stream**	**Values: Degraded stream**
Water clarity	Clear, Muddy	Clear, Muddy
Native fish species	1,3,5	2, 3, 4
Fish habitat	2km, 3km, 4km	1km, 2km, 3km
Native streamside vegetation ♦	Little or none, Moderate, Plentiful	Little or none, Moderate, Plentiful
Channel form	Natural	Straightened, Natural
Cost to household	$0/year, $20/year, $50/year	

♦ Dummy-coded, with *Little or none* as the base.

Because of the large number of attributes in the choice sets, the number of choice events faced by each individual was limited to five to reduce fatigue. In each choice event survey participants were able to choose between the status quo (clearly labeled as such) and two unlabeled alternatives. The first choice option presented in each case was the status quo, forming the base. The first alternative to the base in each choice event was developed from the statistical design plan. The second alternative to the base was the fold over of the first alternative (Louviere *et al*., 2000).

The payment vehicle was regional council rates[2]. Justification for this vehicle was provided with the following introduction, which was read out by the interviewer. The statement was designed to ensure that survey participants were aware that it is not always possible to identify the people responsible for environmental degradation, yet the community may benefit from improving damaged environments. It also sought to introduce the concepts of opportunity costs through environmental trade-offs.

> *Stream restoration and management can be expensive. Sometimes it is obvious who has caused stream changes and they can be made to pay to restore the condition of the stream. In other cases, the changes occurred a long time ago or have been caused by things done for the whole community. In these cases the condition of streams is a community responsibility. Regional Council rates could be raised to allow extra stream restoration activities to be undertaken. If this happened then costs to your household would increase through your rates bill or, if you are renting your house, through having to pay higher rent to your landlord.*
>
> *While the condition of some streams continues to decline because of new and ongoing activities, other streams are getting better because of management actions. Stream managers have to decide whether it is better to try to protect streams that have not been changed much, or to restore streams that have already been degraded. Sometimes it is much easier and cheaper to restore streams that have already been degraded. Restoring degraded streams can mean there is less money available to manage other streams, so their condition can decline.*

Data Collection

Data were collected in personal interviews conducted at the respondent's own home by a professional research agency. The sample was obtained by randomly drawing individual names and addresses from registered voters in South Auckland and North Shore. Response rates were 44% in North Shore and 40% in South Auckland, with 308 interviews completed on the North Shore and 311 completed in South Auckland. Surveying was undertaken in January and February 2003.

The survey drew heavily on design parameters that have proved to be successful in similar Australian studies (Whitten & Bennett, 2001). Attribute levels were communicated wherever possible by the use of icons to allow visual identification of the trade-offs being made. In order to ensure that all respondents were reacting to the same stimuli a two-sided A4 glossy brochure was given to each survey participant to read at their own pace before commencement of choice questioning. The brochure provided photographs of representative stream conditions alongside labelled icons. Large, coloured show cards were used to present the choice questions. The interviewer described the items on the card and explained the choices that were available to the respondent.

Sample Characteristics

Differences between population and sample distributions were tested using population data from the 2001 census for people 20 years of age or older. The

[2] The study preceded the Auckland rates revolt.

sampling frame was a specific address and the participant was randomly selected from people 20 years or older resident at that address. Consequently, the sample should ideally conform to household level census data. The two surveys obtained responses that are representative of home ownership rates and the sex and age distributions within the populations. People with a university degree were more likely to respond than others. The South Auckland sample was over-representative of people from households with incomes less than $50,000 per year, whereas on the North Shore, the sample closely matched population incomes. Large households were over-represented in both samples, possibly a result of the higher probability of finding someone at home in a larger household.

Results

Site-specific models are reported for the two population groups in Table 2. Where possible, the Heteroscedastic Extreme Value model (HEV) was fitted to avoid potential independence of irrelevant alternatives problems. However, the HEV offered no improvement over the standard Multinomial Logit model (MNL) for North Shore, so the MNL is reported in Table 2. Scale parameters[3] are reported for the South Auckland HEV model, but these are not significantly different to the scale parameter for the third option, which is identically set to unity. The models forced inclusion of all stream attributes and the money attribute, but each model includes different interaction effects[4]. While all possible interactions of attributes and socio-economic variables were tested for each model, only significant effects have been retained in the models presented in Table 2.

Personal attributes that significantly affect choices are:

- Age — Respondent's age in years
- People — Number of people in the household
- Degree — 0,1 Dummy: 1 if respondent has a university degree
- Homeowner — 0,1 Dummy: 1 if residence is owned by the inhabitants
- High Income — 0,1 Dummy: 1 if household income exceeds $50,000
- Very High Income — 0,1 Dummy: 1 if household income exceeds $100,000

The coefficients on *Money* are highly significant and of the expected negative sign, indicating that any particular option is less likely to be selected if it costs more. While the low rho-square statistics indicate relatively poor model fits, the significance of stream attribute coefficients is generally strong, with only three of 22 stream attribute coefficients not being significant. The relatively low goodness of fit for these models indicates that there are explanatory factors that have not been included in the models, or that there is considerable underlying inter-personal variance (or both).

Alternative Specific Constants (ASCs) are significant when factors other than independent variables in the model are important determinants of choice[5]. In each

[3] The scale parameter for alternative i is an inverse function of the standard deviation of the unobserved effects for alternative i (Louviere *et al.*, 2000: 139).
[4] While the experimental design precludes interactions between site attributes, it does not curtail interactions between site attributes and individual respondent characteristics.
[5] ASCs are analogous to the constant in a linear regression model, they account for unexplained components of choices. A positive ASC indicates that alternative is preferred to the base alternative for reasons that are not explained by the utility function used to model choices.

choice situation the first option was labelled as the status quo, while the other two options were unlabelled. The choice models used here arbitrarily set the ASC for the third option to zero. Second-option ASCs are not significant, indicating that there were no perceived differences between the unlabelled alternatives apart from the attributes used to describe them. Status quo ASCs are positive, and significant, indicating a preference for the status quo[6].

Interaction effects allow detection of the influence of individual or household-specific characteristics (such as respondent age and household income) on the probability of selecting a particular option. Interaction effects were tested in several ways.

- Firstly, interactions of the variables High Income and Very High Income with the variable MONEY tested income effects. The effects were significant in all cases and supported prior beliefs that wealthier respondents would be prepared to pay more for any given environmental enhancement.
- Secondly, independent variables were interacted with ASCs to test whether personal characteristics influenced choice between the options, particularly between the status quo and either of the two change options. None of these interactions was significant.
- Thirdly, personal characteristics were interacted with each of the site attributes to identify whether particular groups of individuals valued attributes differently. Significant interactions are reported in Table 2. Interaction effects vary significantly between models.

[6] The hypothesis that the status quo is preferred to either of the options entailing change was tested by utilisation of models that included an ASC on the status quo and no ASC on either of the other options. Results mirrored those in Table 2, indicating a significant preference for the status quo, with no significant change to other coefficients. Since these alternative models contain less information, the more general models that allow detection of all order effects are presented in Table 2.

Table 2: Site-Specific Models

	Attribute	**North Shore**	**South Auckland**
Natural Stream Attributes	Water Clarity (N1)	0.6509***	0.6420***
	Fish Species (N2)	0.1082***	0.04667**
	Fish Habitat (N3)	-0.3969***	-0.001452
	Moderate Vegetation (N4A)	0.2759**	0.1567
	Plentiful Vegetation (N4B)	0.2105**	0.5116***
Degraded Stream Attributes	Water Clarity (D1)	0.7706***	0.5996***
	Fish Species (D2)	0.2640**	0.09391*
	Fish Habitat (D3)	0.1315***	0.2098***
	Moderate Vegetation (D4A)	0.2110	0.3447**
	Plentiful Vegetation (D4B)	0.1977**	0.5258***
	Channel (D5)	0.3213***	0.3042***
	Money	-0.009828***	-0.009545***
Personal Attributes	Age x D2	-0.004970**	
	Age x N3	0.007976***	
	Degree x N3	0.1548*	-0.3144***
	Degree x D1	0.3798**	
	Degree x D5	-0.4428***	
	People x D1	-0.1188**	
	People x N4B		-0.08021**
	Homeowner x D3		-0.2394***
	High Income x D5	0.5985***	
	Very High Income x N4B		0.8449**
	Very High Income x D1		0.6737**
	Very High Income x D2		-0.6100***
	Very High Income x D5		0.6585**
Alternative-specific constants	Status Quo	0.2984*	0.5740**
	Second option	0.01845	-0.0955
HEV Scale Parameters	Status Quo	na	1.473
	Second option	na	0.867
N		1331	1281
LL_R		-1433.81	-1388.87
LL_{UR}		-1305.79	-1273.40
Rho^2		0.089	0.083

Significance levels * (10%), ** (5%), *** (1%)

The sign of the interaction effect indicates how the population views the importance of the relevant attribute. For example, the North Shore interaction (High Income x D5) is highly significant and positive, indicating that North Shore High Income households place a higher value than other households on natural channel form in degraded streams. Table 3 presents part worth estimates and their 95% confidence intervals for the models in Table 2.

Table 3: Part Worths ($/household)

		North Shore Mean	95% confidence interval	South Auckland Mean	95% confidence interval
Natural Stream	Water clarity	$66	$43~$110	$67	$42~$114
	Native fish species	$11	$6~$20	$5	$0~$12
	Fish habitat	-$1	-$12~$9	-$3	-$15~$8
	Moderate vegetation	$28	-$1~$68	$16	-$10~$49
	Plentiful vegetation	$21	$2~$50	$41	$17~$75
Degraded Stream	Water clarity	$48	$28~$84	$73	$47~$123
	Native fish species	$4	-$6~$17	$0	-$13~$14
	Fish habitat	$13	$5~$27	$5	-$6~$18
	Moderate vegetation	$21	-$5~$53	$36	$8~$76
	Plentiful vegetation	$20	$0~$48	$55	$28~$97
	Channel	$58	$38~$97	$42	$21~$73

Part worths are different for different people. For example, Channel is valued at $94 by very high income North Shore residents, but is valued at $33 by other North Shore residents. In Table 3 personal characteristics have been set to their population means. Consequently, the North Shore Channel part worth ($58) reflects the proportion of the North Shore population in the very high income category.

Marginal rates of substitution between any two attributes can be identified from the coefficients in Table 2. For example, on the North Shore it is necessary to increase native fish habitat by about 0.8 km on a degraded stream to offset the loss of one native fish species on a natural stream [β_j/β_i = N2/D3 = 0.1082 ÷ 0.1315 = 0.823]. Marginal rates of substitution are relevant guides for policy where mitigation occurs through manipulation of the natural environment. Of course, there is an infinite combination of attributes that yield the same level of utility, allowing design of alternative mitigation scenarios.

Benefit Transfer

Two separate populations within the same metropolitan area have been used. The populations differ in several respects. People living on the North Shore are generally more affluent and better educated than South Auckland residents. While age and sex distributions and home ownership rates are very similar, North Shore households are more likely to consist of only 1 or 2 people. Large households are common in South Auckland. The ethnic mixes of the two communities are also different. These two diverse communities were chosen to test for potential differences in values, and to provide a test of value transfer between communities. Each population was asked to value streams in their own area[7] (one natural stream, one degraded stream, each stream type defined to have the same characteristics in each location). The same streams may differ in value by location simply because of availability of complements and substitutes. Consequently, differences in values between populations reflect differences in people, and in study site values. For simplicity the aggregates of these effects are referred to as differences between populations.

[7] People living in North Shore valued North Shore streams and people living in South Auckland valued South Auckland streams.

Benefit transfer is based upon the underlying assumption that people with the same characteristics in different locations possess similar values for the same items in the same context. Tests of the underlying assumption are frequently undertaken by assessing convergent validity – testing whether benefits measured at one site are the same as those predicted at another. As with non-market valuation method convergent validity tests, it is important to control for as many factors as possible in order to remove explainable reasons for differences. Some of the sources of difference include (Boyle & Bergstrom, 1992; Brouwer, 2000; Brouwer & Spaninks, 1999; Desvousges *et al.*, 1992; Loomis, 1992; Oglethorpe *et al.*, 2000; Shrestha & Loomis, 2001):

1. The nature of the valued sites themselves
2. The changes valued at each of the sites
3. Valuation methods (hedonic, contingent valuation, choice model, …)
4. Time of study (season or year)
5. Availability of substitutes and complements for each of the sites
6. Differences in the people valuing the sites (demographic, social, economic, cultural, …)

The choice model study of Auckland streams valued identical changes to streams with identical characteristics at each site. Furthermore, identical methodology was employed concurrently at each site to avoid elicitation method and temporal impacts that could have affected estimated values. Population differences arising from the influences of age, sex, ethnicity, household size, home ownership and education can be statistically investigated during data analysis. Substitutes, complements, and other contextual differences cannot be controlled using the Auckland study design. Because it removes sources of differences 1-4 and provides partial control over 6, the Auckland study provides an excellent opportunity to measure the convergent validity of benefit measures across sites and populations.

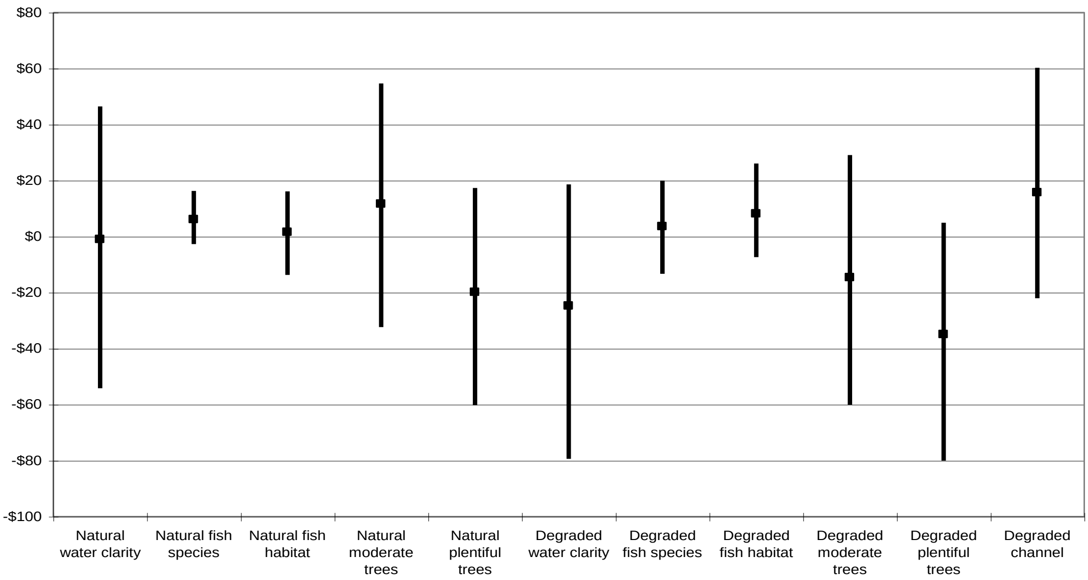
The three principal methods of transferring benefits from a study site or sites to a policy site are direct transfer, benefit function transfer, and meta-analysis. In direct transfer mean values estimated at the study site, or several study sites, are used directly at the policy site, without adjustment to reflect policy site characteristics. For benefit function transfer a valuation function derived at the study site is applied to the policy site using policy site parameters. Benefit function transfer provides control over site and/or population differences, and is generally thought to be more accurate than direct benefit transfer (Rosenberger & Loomis, 2003; VandenBerg *et al.*, 2001). Meta-analysis is another form of valuation function benefit transfer. It uses results from valuation studies completed at many sites to identify statistically the influences of site and personal attributes. Direct transfer and benefit function transfer are both possible using the Auckland Stream study results, but there are insufficient data to apply meta-analysis.

Transferring the Auckland Benefit Estimates

The simplest convergent validity test of benefit transfer accuracy entails comparison of benefit estimate confidence intervals for the two populations. This is a test of direct benefit transfer. It is a weak test because it fails to account for any of the potential reasons that benefits could differ between sites. However, non-overlapping confidence intervals can indicate potential problems with benefit transfer. There are no cases where North Shore and South Auckland part worth confidence intervals do not overlap substantially.

The overlapping confidence intervals test is relatively weak. The possibility of drawing two results in the opposite tails of the distributions is much less than the significance level of the individual confidence intervals (in this case 5%) (Poe et al., 1994). Consequently, it is possible for confidence intervals to overlap even if differences in part worths are significantly different from zero. Figure 3 depicts part worth difference confidence intervals. Rather than reporting two separate distributions for part worths, each developed independently using a Monte Carlo procedure, a single distribution of part worth differences is developed by subtracting the vector of Monte Carlo part worths for one site from the vector of Monte Carlo part worths for the other site (Poe et al., 2001). None of the distributions of part worth differences is significantly different from zero at the 5% level. However, non-significant differences do not imply that benefit estimates at one site are good predictors of benefits at the other site.

Figure 3: Part Worth Difference (North – South) 95% Confidence Intervals



Point Benefit Direct Transfers

An alternative measure of the merits of direct benefit (part worth) transfer validity is the percentage error in using one population point estimate to predict another population point estimate[8] (Rosenberger & Loomis, 2003; Vandenberg *et al.*, 2001). Errors arising from using point estimates from one population to predict point estimates in another population using the direct transfer approach (Table 4) show wide variability, with errors ranging from 2% to 704%. These error magnitudes are similar to those found in other studies (Rosenberger & Loomis, 2003).

Care should be exercised in interpreting these results. Several part worths are not significantly different from zero. Consequently, a small change in one part worth can result in large percentage differences. Further, even changes in sign may not be

[8] Percentage error is defined as 100 x (*Estimate* – *Actual*)/*Actual*. Where *Estimate* is the point measure of benefits at the study site and *Actual* is the point measure of benefits at the policy site.

significant. When consideration is given only to cases in which both part worth point estimates are significantly different from zero the errors are somewhat smaller. Benefit transfer errors in these cases range from 2% to 114%.

Benefit Function Point Estimate Transfers

Benefit function transfers of part worth point estimates[9] produce similar outcomes to direct transfers of point benefit estimates[10]. Consequently, function transfer part worth difference confidence intervals are not reported here[11]. Again, there is overlap on all measures. However, when either benefit function is used to produce part worths for the other location, the differences in part worths for *Degraded Stream Plentiful Vegetation* are significant at the 8% level. South Auckland residents appear to place higher value than North Shore residents on *Degraded Stream Plentiful Vegetation*.

Table 4: Direct and Valuation Function Point Benefit Transfer Errors

				Direct Transfer		Function Transfer	
		North Shore Part Worth (NSPW)	South Auckland Part Worth (SAPW)	Error in predicting SAPW from NSPW	Error in predicting NSPW from SAPW	Error in predicting SAPW from NSPW	Error in predicting NSPW from SAPW
Natural	Water clarity	$66.23	$67.26	-2%	2%	-2%	2%
	Fish species	$11.01	$4.89	125%	-56%	125%	-56%
	Fish habitat	-$1.32	-$2.89	-54%	119%	-12%	314%
	Moderate veg	$28.08	$16.42	71%	-42%	71%	-42%
	Plentiful veg	$21.42	$41.31	-48%	93%	-48%	114%
Degraded	Water clarity	$48.38	$73.12	-34%	51%	-38%	59%
	Fish species	$4.10	$0.51	704%	-88%	704%	-169%
	Fish habitat	$13.38	$5.25	155%	-61%	155%	-67%
	Moderate veg	$21.47	$36.11	-41%	68%	-41%	68%
	Plentiful veg	$20.12	$55.09	-63%	174%	-63%	174%
	Channel	$57.65	$41.94	37%	-27%	39%	-21%

Population means of independent variables are used throughout
Not significantly different from zero at 5% level

Whether the, apparently, large errors in Table 4 are an indictment of benefit transfer is debateable. On the one hand, it is apparent that very large percentage errors can occur from use of transferred point estimates. However, it should be acknowledged that the confidence intervals for individual study sites are large - meaning that use of point estimates <u>at study sites</u> is risky. Comparison of two uncertain values introduces the opportunity of compounding that error. Just as very low errors from point estimate transfers can arise by chance and consequently do not guarantee that

[9] Using study population parameters with the models in Table 2.
[10] Valuation function benefit transfers result in some predictions better than and some predictions worse than direct benefit transfers.
[11] The benefit function transfer analogue of the direct transfer estimates in Figure 3.

benefits transfer is valid, large percentage errors in transferring point estimates do not necessarily indicate that benefits transfer is invalid.

Whereas the simple overlapping confidence interval test offers an unjustified, overly enthusiastic endorsement of benefit transfer, errors associated with transfer of point benefit estimates are likely to provide an overly pessimistic view of the reliability of benefit transfer because they do not account for the confidence intervals surrounding both sets of benefit estimates. Benefit difference confidence intervals provide an approach that is intermediate to these extremes by recognising point differences and their confidence intervals concurrently. Consequently, benefit difference confidence intervals provide better indicators of the reliability of benefits transfer. Using the benefit difference approach with valuation functions, benefit transfer indicates significant differences (albeit at a low level of confidence) between populations for the value of *Degraded Stream Plentiful Vegetation*.

Pooled Models

Further tests of benefit transfer are provided by pooled models, which allow detection of population differences. Pooled models allow for site-specific differences in attribute coefficients and in the role of socio-economic characteristics for each population. The different interaction variables occurring in each model in Table 2 indicate that location differences are likely to occur, with only (Degree x N3) being significant in Table 2 for both groups. Here, the potential use of pooled models for benefit transfer purposes is explored with the aid of two tests.

Test 1

The hypothesis that one utility function applies to both populations is tested by fitting the same model to each group, as well as to the pooled responses from both groups (Table 5). The interactions specified in these models include all significant interactions identified in the individual population models fitted in Table 2.

A method proposed by Swait & Louviere (1993) was used to identify the optimal relative scale of error terms for the two data sets in the pooled model. The relative scale parameter is very close to one. The Swait-Louviere test result indicates that allowing non-uniform errors did not significantly improve model fit relative to the naïve pooled model that assumes identical errors. A likelihood ratio test measures the significance of improvement in fit from use of separate models. The test statistic is distributed χ^2 with degrees of freedom equal to the number of estimated parameters. The result is highly significant.

$$\chi^2 = -2*(LL_{Pooled} - (LL_{North\ Shore} + LL_{South\ Auckland})) = 67.616$$

Together, these results indicate that different utility functions apply to the two populations and that the differences occur in the estimated coefficients, not in the scale factor (Swait & Louviere, 1993). Different utility functions imply that transferring valuation functions between populations may lead to estimation errors.

Test 2

A Pooled model is developed that includes location dummy variables interacted with site attributes and personal characteristics. This type of model has the advantage that it concurrently produces population-specific coefficients within a single model. Coefficient differences between populations are automatically identified without the need for comparison of separate model coefficients, part-worths or their confidence intervals. The location-related interactions take two forms. Two-way interactions (e.g. South x N2) show the direct impact of location on the value of the attribute. Three way interactions (e.g. South x Degree x N3) show differences by location in the way personal characteristics influence the values of specific attributes. Results are reported in Table 6. As with Test 1, the Swait-Louviere procedure was used to identify the relative scale parameter, which at 0.990 is not significantly different from one. A likelihood ratio test [$\chi^2 = -2*(LL_R - LL_{UR})$][12] indicates that location variables are highly significant as a group.

Five personal characteristics (High Income, Age, Degree, Homeowner, Household Size) affect attribute values independent of location. Three attribute part worths differ between locations, independent of personal characteristics. The value of *Natural Stream Fish Species* abundance is greater for North Shore residents, while South Auckland residents place higher values on *Plentiful Vegetation* at both types of stream. The significant two-way interactions between attributes and location (South x N2; South x N4B; South x D4B) indicate that, despite overlapping 95% confidence intervals, part worths for *Natural Stream Fish Species* and *Plentiful Vegetation* on both stream types are significantly different at the 95% confidence level.

There are seven three-way interactions that differentiate the impact of personal characteristics by location. Of particular note is the diverse influence on *Degraded Stream Channel* form because of income. High Income causes increased willingness to pay for a more natural channel form on the North Shore (β=0.5536), but has no significant effect in South Auckland (β=0.5536-0.5868=-0.0332). However, South Auckland displays a strong impact from Very High Income that does not occur in North Shore.

[12] LL_{UR} is the log likelihood of the full (31 parameter) model that includes location variables. LL_R is the log likelihood of the same excluding the 10 location variables.

Table 5: Pooled and Independent Models

	Attribute	North Shore	South Auckland	Pooled
Natural Stream Attributes	Water Clarity (N1)	0.6035***	0.6940***	0.6412***
	Fish Species (N2)	0.09836***	0.0517*	0.07787***
	Fish Habitat (N3)	-0.3447***	-0.1621	-0.2664***
	Moderate Vegetation (N4A)	0.2268*	0.1998	0.1980**
	Plentiful Vegetation (N4B)	0.04974	0.6627***	0.3288***
Degraded Stream Attributes	Water Clarity (D1)	0.6473***	0.8107***	0.6606***
	Fish Species (D2)	0.2298*	0.1145	0.1939**
	Fish Habitat (D3)	0.1683**	0.2052**	0.1945***
	Moderate Vegetation (D4A)	0.1735	0.3750**	0.2519***
	Plentiful Vegetation (D4B)	0.1629*	0.5854***	0.3318***
	Channel (D5)	0.2843***	0.3999***	0.3414***
	Money	-0.009232***	-0.01039***	-0.009229***
Personal Attributes	Age x D2	-0.004082*	-0.000378	-0.002812
	Age x N3	0.006911***	0.003714	0.005761***
	Degree x N3	0.1358	-0.4023**	0.01696
	Degree x D1	0.3582**	0.1393	0.2182*
	Degree x D5	-0.4202***	-0.2229	-0.3579***
	People x N4B	0.03691	-0.1128**	-0.04293
	People x D1	-0.08636*	-0.04657	-0.04768*
	Homeowner x D3	-0.07346	-0.2286**	-0.1447**
	High Income x D5	0.5055***	-0.09889	0.2724***
	Very High Income x N4B	0.1828	1.0363*	0.3877**
	Very High Income x D1	-0.005306	0.8662**	0.1344
	Very High Income x D2	-0.07834	-0.7153**	-0.2025**
	Very High Income x D5	0.2913	0.9144*	0.5272***
ASCs	Status Quo	0.4417**	0.4706	0.4393**
	Second option	0.1154	-0.2026	0.03208
HEV Scale Parameters	Status Quo	1.4645	1.0943	1.2845
	Second option	1.1302	0.7605	1.0041
Relative scale parameter				0.934
N		1331	1256	2587
LL_R		-1433.811	-1361.700	-2797.702
LL_{UR}		-1302.836	-1242.487	-2579.131
Rho^2		0.091	0.088	0.078
LR test of pooled versus separate models		$\chi^2 = 67.616$	$P(\chi^2,30) = 1.01 \times 10^{-5}$	
Swait-Louviere test of relative scale parameter		$\chi^2 = 0.410$	$P(\chi^2,1) = 0.522$	

Significance levels * (10%), ** (5%), *** (1%)

Table 6: Pooled Model with Location Variables

	Attribute	Coefficient	
Natural Stream Attributes	Water Clarity (N1)	0.6514***	
	Fish Species (N2)	0.1169***	
	Fish Habitat (N3)	-0.2718***	
	Moderate Vegetation (N4A)	0.2038**	
	Plentiful Vegetation (N4B)	0.2098**	
Degraded Stream Attributes	Water Clarity (D1)	0.7429***	
	Fish Species (D2)	0.06438*	
	Fish Habitat (D3)	0.2041***	
	Moderate Vegetation (D4A)	0.2662***	
	Plentiful Vegetation (D4B)	0.2228**	
	Channel (D5)	0.3318***	
	Money	-0.009727***	
Personal Attributes	High Income x D5	0.5536***	
	Age x N3	0.006338***	
	Degree x D5	-0.3021***	
	Homeowner x D3	-0.1593***	
	People x D1	-0.05490**	
Location Variables	South x N2	-0.07203**	
	South x N4B	0.3665**	
	South x D4B	0.2687**	
	South x Degree x N3	-0.4040***	
	South x People x N4B	-0.09817***	
	South x High Income x D5	-0.5868***	
	South x Very High Income x N4B	0.9624**	
	South x Very High Income x D1	0.8234**	
	South x Very High Income x D2	-0.6119***	
	South x Very High Income x D5	0.7734**	
Alternative-specific constants	Status Quo	0.4357**	
	Second option	-0.03615	
HEV Scale Parameters	Status Quo	1.2471	
	Second option	0.9299	
Relative scale parameter		0.990	
N		2587	
$LL_{\text{Constant only}}$		-2797.702	
$LL_{\text{No location variables}}$		-2591.157	
$LL_{\text{Full model}}$		-2557.716	
Rho^2		0.086	
LR test of location variables		$\chi^2 = 66.882$	$P(\chi^2,10) = 1.77 \times 10^{-10}$
Swait-Louviere test of relative scale parameter		$\chi^2 = 0.010$	$P(\chi^2,1) = 0.9203$
Significance levels * (10%), ** (5%), *** (1%)			

Table 7 reports site-specific part worth estimates and 95% confidence intervals for each location from the pooled model. In each case results are modelled for a 45 year old respondent with a university degree from a high-income, home owning household of three people.

Table 7: Part Worths – Pooled Model ($/household)

45 year old homeowner with a degree. Household income > $50,000 p.a. 3 people in household.		North Shore	95% confidence interval	South Auckland	95% confidence interval
Natural stream	Water clarity	$67	$49~ $94	$67	$49~ $94
	Fish species	$12	$7 ~ $18	$5	$0 ~ $10
	Fish habitat	$1	-$6 ~ $9	-$40	-$68 ~ -$16
	Moderate vegetation	$21	-$1 ~ $45	$21	-$1 ~ $45
	Plentiful vegetation	$22	$3 ~ $43	$29	$10 ~ $54
Degraded stream	Water clarity	$59	$43 ~ $86	$59	$43 ~ $86
	Fish species	$7	$0 ~ $15	$7	$0 ~ $15
	Fish habitat	$5	-$4 ~ $14	$5	-$4 ~ $14
	Moderate vegetation	$27	$8 ~ $50	$27	$8 ~ $50
	Plentiful vegetation	$23	$4 ~ $44	$51	$28 ~ $77
	Channel	$60	$34 ~ $92	$0	-$33 ~ $32

The shaded cells in Table 7 highlight attributes for which part worths are invariant between populations irrespective of personal characteristics. *Degraded Stream Water Clarity* and *Degraded Stream Fish Species* do not differ in the case reported in Table 7 because their differential effects only occur for very high-income households. The simple non-overlapping confidence interval test indicates highly significant differences between populations for *Degraded Stream Channel Form* and *Natural Stream Fish Species* part worths. The other three part worths that are affected by personal characteristics exhibit confidence interval overlaps.

Values in Table 8 have been derived from Monte Carlo simulation of the differences in part worths for the five attributes in Table 7 that differ by location. In each case the estimated South Auckland part worth has been subtracted from the estimated North Shore part worth to yield a simulated distribution of part worth differences. In only one case (*Natural Stream Plentiful Vegetation*) does the 95% confidence interval include zero. These results indicate that, even after controlling for personal characteristics, North Shore residents in this demographic profile place significantly higher values on abundance of *Natural Stream Fish Species*, availability of *Natural Stream Fish Habitat*, and *Degraded Stream Channel Form*. South Aucklanders value *Degraded Stream Plentiful Vegetation* more highly than do North Shore residents.

Table 8: Pooled Model Part Worth Differences

45 year old homeowner with a degree. Household income > $50,000 p.a. 3 people in household.	Part Worth Differences (North minus South)	95% confidence interval
Natural Stream Fish Species	$7	$1 ~ $15
Natural Stream Fish Habitat	$42	$17 ~ $70
Natural Stream Plentiful Vegetation	-$7	-$30 ~ $14
Degraded Stream Plentiful Vegetation	-$28	-$58 ~ -$3
Degraded Stream Channel	$60	$29 ~ $100

Differences arise irrespective of personal characteristics because of the significant two-way interaction variables (South*N2, South*N4B and South*D4B) in Table 6. While, part worth differences occur regardless of personal characteristics, differences vary by demographic profile. Consequently, the non-significance of *Natural Stream Plentiful Vegetation* in Table 8 is not inconsistent with the model in Table 6. For example, changing household income to more than $100,000 per annum (while leaving all other characteristics unchanged) produces significant part worth differences for this attribute, as well as for *Degraded Stream Water Clarity* and *Degraded Stream Fish Species* abundance.

Conclusions

The study provides important insights into benefits transfer. Overlapping part worth confidence intervals indicate similar values between the two populations, but provide an overly optimistic view of benefits transfer when compared to confidence intervals of attribute part worth differences. Point estimate transfers, whether direct or benefit function transfers, resulted in some very large errors. However, point transfers do not account for uncertainty in the estimates at either site and so percentage errors of point transfers provide poor tests of benefit transfer. Tests of part worth differences and two pooled model tests were used to overcome deficiencies in overlapping confidence interval and point estimate tests. Part worth difference tests identified significant (albeit at low levels) differences in one part worth using both direct and benefit transfer approaches.

Two different pooled model tests have been used to show that the same utility function does not apply to both populations. Because it has a larger sample size and the ability to control for other factors, tests based on the pooled model with location variables have more power to identify differences than do tests based on independently estimated models for each site. In addition, pooled models identify the sources of part worth differences. Part worth difference distributions from the pooled model that includes location effects are significantly different from zero, consistent with the significance of location variables in the model. The pooled models indicate that errors will arise from transfer of benefits between locations. Those errors were not identified by overlapping confidence interval or part worth difference tests based on independent models. Significant part worth differences remaining in the pooled model, after controlling for socio-economic effects, suggest that there are unaccounted-for differences between values in the two population groups. Such differences may arise because of contextual differences, unaccounted for socio-economic differences, or simply because people living at the two locations value stream attributes differently. These results caution against benefit transfer.

Studies of the type conducted here have the luxury of estimated values for both the original and target sites (or populations). When benefit transfer is undertaken for policy purposes it is not known what the true value at the policy site is, or even the range of values that include the true value. If that information were available there would be no need for benefit transfer. In that situation it is not possible to compare value distributions or point estimates of value, or to fit pooled models. The analyst has three options – direct transfer of benefits, transfer of valuation functions, or don't transfer benefits at all. What would happen if valuation functions or point estimates were transferred in these cases? It is not possible to provide an unambiguous answer to that question, as it depends on the policy proposal being evaluated. When off-site mitigation is undertaken, several attributes may change at each stream, which means that errors may compound - or they may cancel each other out. While the potential to be wrong is moderated in this situation, the implications of being wrong may be very serious. It is apparent that use of point estimates has the potential to produce highly biased results. The implications when confidence intervals are developed for welfare changes are less likely to be problematic, but, because errors may compound across several attributes, there is still the potential to obtain extremely misleading indicators of welfare change. Overall, the evidence presented here adds weight to the growing literature that has identified large potential errors from benefit transfer, even under close to ideal conditions.

Acknowledgements

This study was commissioned and funded by Auckland Regional Council. The conclusions reached belong to the authors and there is no implication that this paper represents the views of ARC. Part of the research was undertaken while Geoff Kerr was Visiting Fellow at the Asia Pacific School of Economics and Government at the Australian National University, Canberra. We are extremely appreciative of the encouragement and support offered by APSEG. We gratefully acknowledge the assistance of Jeff Bennett, John Rolfe, Chris Hatton and Graeme Ridley in designing and reviewing this research. Two anonymous referees offered suggestions that improved presentation. Remaining errors are ours alone.

References

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

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

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

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.

Horowitz, J.K. and McConnell, K.E. (2002). A review of WTA/WTP studies. *Journal of Environmental Economics and Management* 44(3): 426-447.

Kerr, G.N. and Sharp, B.M.H. (2002). *Community perceptions of Stream Attributes: Focus Group Results*. Report to Auckland Regional Council. Agribusiness and Economics Research Unit, Lincoln University, Canterbury.

Kerr, G.N. and Sharp, B.M.H. (2003). *Community mitigation preferences: A choice modelling study of Auckland streams*. Agribusiness and Economics Research Unit Research Report No. 256. Lincoln University.

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

Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000). *Stated choice Methods: Analysis and Application.* Cambridge University Press.

Oglethorpe, D., Hanley, N., Hussain, S. and Sanderson, R. (2000). Modelling the transfer of the socio-economic benefits of environmental management. *Environmental Modelling and Software* 15: 343-356.

Poe, G.L., Giraud, K.L. and Loomis, J.B. (2001). Simple computational methods for measuring the differences of empirical distributions: Application to internal and external scope tests in contingent valuation. *Cornell University Applied Economics and Management Staff Paper 2001-05.*

Poe, G.L., Severance-Lossin, E.K. and Welsh, M.P. (1994). Measuring the difference (X-Y) of simulated distributions: A convolutions approach. *American Journal of Agricultural Economics* 76: 904-915.

Rosenberger, R. and Loomis, J.B. (2003). Benefit Transfer. Chapter 12 in Champ, P.A.; Boyle, K.J. and Brown, T.C. (eds) *A Primer on Nonmarket Valuation.* Dordrecht: Kluwer Academic Publishers.

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

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: 305-314.

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: Bergstrom, J.C., Boyle, K.J. and Poe, G.L. (Eds) *The economic value of water quality.* Cheltenham: Edward Elgar.

Whitten, S.M. and Bennett, J.W. (2001). Non-market values of wetlands: A choice modelling study of wetlands in the Upper South East of South Australia and the Murrumbidgee River floodplain in New South Wales. *Private and Social Values of Wetlands Research Report No.8*. School of Economics and Management, University College, University of New South Wales, Canberra.