This paper was presented to AARES 2009, the 53rd Annual Conference of the Australian Agricultural and Resource Economics Society, Cairns International Hotel, Northern Queensland, 10th- 13th February, 2009

Testing Different Types of Benefit Transfer in Valuation of Ecosystem Services: New Zealand Winegrowing Case Studies

Ramesh Baskaran, Ross Cullen, Sergio Colombo

Ramesh.Baskaran@lincoln.ac.nz, Commerce Division, Lincoln University, New Zealand

Ross.Cullen@lincoln.ac.nz, Commerce Division, Lincoln University, New Zealand

scolombo@ugr.es, Department of Agricultural Economics, IFAPA, Granada, Spain

Abstract

Most ecosystem services (ES) are neither priced nor marketed. Resource managers may fail to take into account degradation of unpriced services in their resource management decisions. Being able to estimate values for ES is fundamental to designing policies to induce resource users to provide (or improve) ES at levels that are acceptable to society. Conducting ecosystem valuation via non-market methods is costly and time consuming. Benefit Transfer (BT) using choice modeling (CM) is a potentially cost-effective method for valuing ES by transferring information from existing valuation studies (and study sites) to a target area of interest (policy sites). The prime objective of this paper is to examine the validity of BT and hence whether it is feasible to conduct the transfer process and assist policy making. The paper focuses on the environmental impact of winegrowing practices in two New Zealand winegrowing regions. The two sites, Hawke's Bay and Marlborough, have similar environmental issues and attributes but are geographically separated. The study estimates WTP and Compensating Surplus (CS) for ES applying CM and, subsequently, given the preferences of respondents across sites and populations, tests the transferability of unadjusted value transfer (WTP) and benefits function (CS) assessing four different types of BT.

Key Words: Benefit transfer, choice modeling, New Zealand winegrowing, ecosystem services

JEL Classification: Q1, Q2, Q5

1.0 Introduction

Non-market valuation methods have contributed an important set of new tools, in particular choice modeling, to estimate the value of Ecosystem Services (ES) where they are not priced via markets. Design of cost-effective public policy incentives for farmers to provide ES from agriculture requires estimates of how valuable improvements in the level of specific ES are to the public. Research is also required to determine which kinds of ES could provide the greatest overall welfare benefits to society (Swinton *et al*. 2007). Thus, ecosystem service valuation can potentially provide new ways to compare the costs and benefits of different agricultural strategies, using the dollar as the common metric of value.

However, non-market valuation studies are time consuming, labour intensive, and costly. Research funders are interested in finding ways to reduce costs of valuing ES and other non-market items. Value transfer uses value estimates from an existing study and transfers it to another site or alternative context that is of interest. This practice (benefit transfer) is attractive if it can provide acceptable estimates of value at lower cost than would new non-market value studies for each new site or context. Nevertheless, there are concerns about the accuracy of the values that are transferred and research is needed to determine in which circumstances benefit transfer will provide acceptable value estimates.

This paper has two objectives. The first is to estimate values for selected ES associated with winegrowing. The second objective involves checking if transfer of the estimated ES values across sites and populations is valid. An advanced choice modeling (CM) approach incorporating heterogeneity preferences, known as Random Parameter Logit (RPL) model is used to estimate the selected ES values. Surveys focused on the two largest New Zealand winegrowing regions, Marlborough and Hawke's Bay, are used as case studies. This research is conducted with a goal of applying Benefit Transfer (BT). The research treats each region as both a 'study' site (an original survey site from which to transfer values to other sites) and as a 'policy' site (whereby site values are transferred to the policy site from the original survey site). This study assesses the accuracy of such transfers. By comparing values, the study obtains an estimate of the 'transfer error' (i.e., the difference between the value obtained by surveying a given site and the value obtained by transfer from another site) and provides a series of recommendations about whether it is advisable, and in which conditions, to transfer the ES values resulting from various winegrowing management practices.

Several conditions necessary for performing effective and efficient benefit transfers have been considered in the study design, in particular the similarity of site characteristics (Desvousges et al. 1992). The two study regions exhibit some similarities in terms of the importance of winegrowing, recent changes in the quality of the environment; demographic profiles of the two populations; the extent and magnitude of the population that may be affected by resource use impacts; the type of value measurement (marginal value); and the period where the studies are carried out (temporality).

The study identifies four types of BT tests, and provides an extension of Morrison and Bergland (2006). A key difference in this study is the spatial dimension between the study site and the policy site where attitudes, tastes and perceptions of environmental issues may differ among the populations in the two regions. Thus, an important hypothesis can be tested: Do the geographically separated Hawke's Bay (North Island) and Marlborough (South Island) regions have the same Willingness To Pay (WTP) estimates for the winegrowing ES considered, and hence, is BT across sites and populations valid?

In this paper, the validity of four different types of BT is tested by determining whether the WTP estimates are statistically equivalent. Three tests are conducted for each type of BT transfer. In the first test, the equality of the model parameters is examined. Second, the equality of WTP and Compensating Surplus (CS) is tested. Lastly, the study applies a new statistical validity test proposed by Johnston and Duke (2008) incorporating the tolerance level of transfer error for policy purposes. Assessment of this error may allow analysts to judge if the transfer process is reliable and hence whether in the future it is valid to transfer values from study sites to policy sites without having to conduct new research or surveys.

2.0 Types of Benefit Transfer Tests

This research has identified at least four types of BT tests across sites and populations as illustrated in Figure 1. The four types of BT tests are as follows:

- Type 1 BT: Differences across populations only. This test examines the transferability of value estimates held by different populations towards one study site. It tests for population effects, as the site is held constant and the population varied. For example, Hawke's Bay and Marlborough populations valuing Marlborough site (B vs D) and vice versa for Marlborough population (A vs C).
- Type 2 BT: Differences across sites only. This test involves comparisons of the values held by a single population for multiple sites. The test asks whether the population has similar values for two sites (valuing their own as well as a different site). For example, Hawke's Bay population valuing their own and Marlborough site (A vs B) and vice versa (C vs D).
- Type 3 BT: Differences across sites and equivalent populations which are geographically separated. This involves comparing the values that respondents have for ES within their region to the values that another group of respondents who reside in a different region have for a similar ES in their region. For example, Hawke's Bay and Marlborough populations valuing ES in their respective region (A vs D). Since both regions are at different locations, it can be tested as A vs E and D vs F.
- Type 4 BT: Differences across sites and different populations. In general, this test compares values held by different populations towards sites other than their own. For example, Hawke's Bay population valuing Marlborough site (B) and Marlborough population valuing Hawke's Bay site (C).

Insert Figure 1 here

3.0 Method

Choice Modeling (CM)

In this study, a CM method, Random Parameter Logit (RPL) was implemented to elicit respondents' preferences for different hypothetical changes in the quality and quantity of environmental attributes linked to winegrowing practices.[1] The RPL specifications provide the analyst with valuable information incorporating unobserved heterogeneity in the data while estimating unbiased parameters estimates (Train 1998; Train 2003; Hensher et al. 2005; Hanley et al. 2006). In addition, in the context of BT analysis, RPL reduces the magnitude of the transfer error (Colombo et al. 2007). Therefore, a RPL modeling framework is applied to estimate the marginal WTP for improvements in the winegrowing ES and to determine the convergent validity of BT in valuing marginal changes in environmental quality which may differ across sites.

Data Collection

The choice modeling surveys contained multiple choice questions (choice situations) about alternative policies for improving four ES attributes associated with winegrowing. The questionnaire consisted of three parts. The first part contained questions regarding respondents' opinions and their awareness of current environmental impacts caused by winegrowing. These questions had the objective of introducing the respondent to the subject of winegrowing impacts on ES. The second part of the survey contained the choice situation questions. Respondents were first briefed about the selected ES attributes and associated cost to the household to achieve change in level of ES. The cost to the household (the payment vehicle) was defined as an additional annual payment to the regional council responsible for the management of the environment over the next five years. This payment vehicle did not raise protest responses during the pilot study and was judged to be suitable for the ES of interest.

In the choice questions, respondent were asked to select the option they favoured the most out of the three alternatives provided. Each option contained the four attributes and the cost to the household with various levels of attribute combinations. Attributes discussed were toxic chemical residues in wine, risk of toxic

[1] The random parameter logit (RPL) model is a generalisation of the standard conditional logit model that explicitly considers taste variation among individuals. Those who are interested in the theoretical underpinnings of RPL can refer to the papers of Train (1998), Chapter 6 of Train (2003) or Chapters 15 and 16 of Hensher et al. (2005).

chemicals reaching groundwater, greenhouse gas emissions per hectare per year, and the condition of native wildlife populations in vineyards. Each attribute was presented to respondents as several discrete levels. For example, the attribute of greenhouse gas emissions was presented as having three discrete levels: zero net emissions (highest improved level); 30% reduction; and 'no change' from current emission level. The attributes were coded employing effects coding instead of dummy coding because of the advantage of estimating the coefficient estimates of attribute variables in which the effects are uncorrelated with the intercept of the regression (Adamowicz et al. 1994; Louviere et al. 2000; Hensher et al. 2005; Bech and Gyrd-Hansen 2005). All of the attributes selected are factors that a policy maker can affect, directly or indirectly and they were selected based on expert advice, current debates in focus groups and information from wine industry literature. Table 1 provides a more complete description of all explanatory variables and their specified effects coding based on the levels. The final part of the survey contained questions regarding the respondents' socioeconomic status.

Insert Table 1 here

There are four attributes with three levels and the cost attribute with 6 levels (3^4 x 6^1) which were combined in fractional factorial main effects experimental design (Louviere et al. 2000), providing 18 profiles in order to form the choice sets. The choice sets were constructed following the procedure proposed by Street et al. (2005) obtaining choice sets with a 94.85% efficiency rate which were then blocked to 3 versions of 6 choice sets. Each choice question has three alternatives and the third alternative was always the status quo (current plan). In other words, each respondent in each choice set had to choose either an improved environmental management plan (Alternative 1 or 2) or the current plan (Alternative 3).

In the beginning of February 2008, pilot surveys were conducted on randomly selected residents in Canterbury, New Zealand. During April 2008 a pre-survey card, survey booklet, cover letter, and a reminder post-survey card were mailed to 4392 respondents selected from the New Zealand electoral roll using a random sampling design. The sample was divided into four strata: 1098 respondents were randomly selected from the Marlborough population to value ES in their own region (MARL); 1098 respondents from the Hawke's Bay population to value ES in their own region (HB); 1098 respondents from Hawke's Bay population to value ES in Marlborough region (HBPOP); and lastly, 1098 respondents from Marlborough population to value ES in Hawke's Bay region (MARLPOP). The study received a total of 330, 218, 192 and 262 completed questionnaire responses for the MARL, HB, HBPOP and MARLPOP samples. The effective response rates for these samples are 30%, 20%, 18% and 24% respectively.

4.0 Results and Discussion

Socioeconomic and Attitudinal Characterization of the Samples

Table 2 presents descriptive statistics and socio-demographic data for the four samples. The four samples do not differ much from each other but show greater differences in comparison to regional population. These differences between the samples and the total population may influence the model estimates. However, the CM method allows such biases to be corrected by using benefit function estimates rather than point estimates (mean WTP). The inclusion of socioeconomic variables in the benefit function hence allows the adjustment of the WTP (van Bueren and Bennett 2006).

Insert Table 2 here

HBPOP and HB Samples

There were 192 and 197 respondents who provided completed surveys from HBPOP and HB samples, respectively. Of the total number of respondents, 28 (15%) in HBPOP and 26 (13%) in HB samples expressed a protest answer regarding the proposed project. These protest bids were removed from the sample.[2] The majority of respondents who provided a protest response proclaim 'vineyards as business should pay for any damages they cause' as their main reason for not paying for the improvement plans. All respondents that displayed a genuine zero WTP by always choosing the current policy option (7% in HBPOP and 6% in HB), and those that chose either alternative 1 or 2 at least once were considered in the analysis, giving a total number of 974 and 962 observations for HBPOP and HB models estimation respectively.

[2] It is established in the literature that some respondents do not state their true value for the good in question. Respondents may state a zero WTP although their true WTP is higher than zero or they may state a very high amount which is much greater than their true WTP (Meyerhoff and Liebe 2008). If protest occurs, this will result in an incorrect economic value estimation of the good in question. In this study, those who expressed a protest answer were identified using a follow up question and deleted from analysis.

Table 3 reports the environmental attitudinal and beliefs on winegrowing practices of the four samples. The results show that more than three quarters of the sample in HBPOP declared ‘don't know' whether they are satisfied with the environmental quality in the Marlborough region. This is not surprising as about 56% of the respondents had not visited the area before and are not aware of the environmental conditions of the area. On the other hand, more than three quarters of the sample in HB are satisfied with the environmental quality in the Hawke’s Bay region and more than 75% of the sample live less than 5 kilometers away from a vineyard. In response to a statement ‘you enjoy views of vineyard landscapes that include native plant species’, about 85% of HBPOP and 53% of HB respondents answered in the affirmative. However, interviewees' preferences in HB are divided into two groups with 53% of the sample agreeing with the statement and 43% in disagreement. Respondents were also asked their opinions on whether winegrowing practices are harmful to groundwater quality, emit greenhouse gases, and affect health through chemical residues in wine. In general, respondents agree that winegrowing practices have the potential to damage the environment unless properly managed, but perceptions vary about these issues. It is surprising to find for both samples that a substantial percentage of respondents did not know the effect of winegrowing on these attributes. As might be expected, more than 75% of respondents in HBPOP would like to see wine bottles labelled so that consumers can be guaranteed that environmentally sustainable practices have been used in grapegrowing and winemaking. However, 79% of the HB respondents stated that they do not want to see wine bottles labeled so that consumers can be guaranteed that environmentally sustainable practices have been used in grapegrowing and winemaking. This suggests that the majority of Hawke’s Bay respondents, who are satisfied with their regional environmental quality, are more confident of their wine being produced in an environmentally sustainable manner and hence, labeling is unimportant. If wine is from other regions, they prefer to see the labels provide information on sustainability of production practices.

Insert Table 3 here

MARLPOP and MARL Samples

There are 262 and 301 completed surveys from respondents of MARLPOP and MARL samples respectively. Of the total number of respondents, 27% (MARLPOP) and 12% (MARL) expressed a protest answer regarding the proposed project and were removed from the sample. It is also observed that 54% (MARLPOP) and 57% (MARL) of the protest respondents want vineyards to accept the costs of changed production systems. All respondents that displayed a genuine zero WTP by always choosing the current policy option (6% in MARLPOP and 4% in MARL), and those that chose either alternative 1 or 2 at least once were considered in the analysis, giving a total number of 1134 (MARLPOP) and 1509 (MARL) observations for model estimations.

64% of the respondents stated they had never visited the HB region. Almost 83% of the sample in MARLPOP declared they lacked knowledge about the environmental quality in the Hawke's Bay region. In the MARL sample, nearly 88% were satisfied with the environmental quality in the region. The degree of satisfaction is the highest in this region relative to other samples. It is important to note that about 37% of the MARL respondents live less than 1 km from a vineyard compared to 31% in the HB sample. The MARLPOP respondents also have similar tastes for scenery to HBPOP respondents with approximately 81% of the sample concurring with the statement they enjoy views of vineyards landscape that include native plant species. In contrast, MARL respondents differ in the enjoyment they experience viewing vineyards landscape and approximately 90% of the sample disagreed with the statement. As described, respondents were also asked to affirm whether winegrowing practices are harmful for underground water quality, emit greenhouse gases and damage health via wine residue content. In the MARLPOP sample 59% of the respondents are concerned about groundwater quality issues. Likewise, 53% are aware that toxic chemicals in wine residues are dangerous for health. About 29% of the respondents agree that winegrowing produces greenhouse gas emissions, but 34% did not know whether winemaking contributes to greenhouse gases emissions. In contrast, MARL respondents generally disagree with the statements. In particular, 59% (24% did not know) of the sample disagreed that winegrowing can effect groundwater quality, 42% disagreed (34% did not know) that it contributes to greenhouse gases emissions and 57% disagreed (17% did not know) that pesticides in wine are dangerous for health. Finally, similar to Hawke's Bay residents, people in Marlborough would like to see wine bottles to be labelled especially if they come from other regions. In addition, they also judged that their local wine was produced via environmentally sound practices, and it is not necessary for labels to provide that information.

RPL Models

The choice data were analysed using NLOGIT 4.0 statistical software. Table 4 presents RPL models for the four samples in which the socioeconomic and attitudinal characteristics of respondents have been added. The models were estimated using 100 Halton draws and considered the random parameters to be independent.[3] In this study, all the attributes except COST which has a triangular distribution, are assumed to be random variables with normal distribution. The normal distribution for the non monetary attributes was used because respondents may be indifferent to increasing or diminishing quality or quantity of the attributes. For instance, people who completely trust the effectiveness of food safety regulations may not care even if the toxic chemical residue content in wine increases a bit. The cost attribute is assumed to follow a triangular distribution to ensure non negative WTP for winegrowing improvements over the entire range of the distribution which guarantees deriving behaviourally meaningful WTP measures while allowing taste heterogeneity for this attribute.[4]

Insert Table 4 here

In the analysis, all the attributes except for cost were effect coded, using the level of the current winegrowing management practices as reference point or base level for each attribute.[5] The effects coding system does not directly estimate the parameter of the base level; nonetheless it can be inferred from the estimation of the two effect-coded corresponding attribute parameters. In other words, the parameter of the current level of management practices for each attribute is equal to the negative sum of the estimated coefficients for that attribute. For example, the coefficient of the current (base level) water quality in HBPOP model is equal to $-(\hat{\beta}_3+\hat{\beta}_4)=-(1.0975+1.2212)=-2.3187$. The coefficients of all the attribute base levels

[3] All the random parameters models described in this report have been estimated using these settings.

[4] Following Hensher et al. (2005), a constraint triangular distribution was used in which the variance (spread) of the distribution is made equal to the mean, which is, Cost (t, 1). Such a constraint forces the same sign for the Cost estimate across the entire distribution. This is useful where a change of sign does not make sense.

[5] Refer Louviere et al. 2000 for detailed discussions of the different coding schemes.

have similar negative sum values. Thus, in Table 4, it is noticeable that respondents have an aversion to the current winegrowing management in all the four samples.

Overall the models are highly significant and show an excellent fit to the data.[6] All the significant attribute coefficients have the *a priori* expected sign for the models. In all the models, it is shown that the attribute RESORG is insignificant. This suggests that reducing the residue content in wine is a matter that significantly affects people's satisfactions only if the reduction is complete (zero level) rather than a marginal reduction. In addition, as noted in Table 3, this could also be due to the lack of knowledge and high levels of disagreement among respondents with the statement that chemical residues in wine are of concern. The effects that winegrowing has on underground water quality is deemed extremely important (highly significant and large coefficients) by the four samples of respondents, and a reduction in the risk of toxic chemicals reaching groundwater increases respondents' utility.

Reduction in the emissions of greenhouse gases is also of concern to all the sample respondents. Nevertheless, only a reduction of greenhouse gas emissions by 30% increases HB and MARLPOP respondents' utility, and they are indifferent about a reduction to zero emissions. All four sample's respondents prefer increasing the native wildlife population in vineyards by at least 10% or 30% relative to the current condition. However, they get more satisfaction from a 30% increase, except for MARL respondents who get more utility from a 10% increase. HB respondents favour increasing the native wildlife population in vineyards by at least 30% relative to the current condition. A smaller improvement is not of interest to them. As expected, cost is highly significant and has a negative sign for all the samples, showing that the higher the cost associated with a policy option, the less likely a given respondent is to choose that option.

By interacting individual socioeconomic and attitudinal variables with alternative specific constant (ASC), it is possible to enrich information about a particular sample and also to explain a part of respondent heterogeneity. It is surprising to note that most of the ASC in the models except MARL are negative with a large coefficient and are highly significant, showing that there are systematic reasons other than the attribute

[6]Simulations by Domencich and McFadden (1975) suggest values of ρ^2 between 0.2 - 0.4 are comparable to values between 0.7 - 0.9 for R^2 in the case of the ordinary linear regression.

values that drove respondents' when choosing the status quo option. MARL sample respondents who are satisfied with the environmental quality prefer to hold on to the current management practices instead of improving them. It is rather peculiar that the HB sample respondents who live closer to the vineyard are more likely to stick with the current winegrowing management.

Although the HB and HBPOP respondents are from the same region, the results only show some similarities in terms of socioeconomic interactions. For example, both sample respondents are in favour of winegrowing management practices that lead to more wildlife in the landscape, reduced wine residues, and more informative (environment friendly practice) labelling of wine bottles. The HBPOP respondents who are young and highly educated support better environmental management. By contrast, the HB residents, in particular females and older people are more likely to choose the improvement plans over the current environmental management. Highly educated residents are also not in favour of improving the present conditions. Household income is significant and has opposite signs for HB and HBPOP models indicating that higher income people are more likely to support the proposed winegrowing management practises in their own region but not outside their region.

Comparing MARL and MARLPOP samples, it is clear that males and younger people in MARL are more likely to choose the improvement plans, the opposite to what is observed in the MARLPOP model. For the MARL sample, neither household income nor education affects the choice of the improvement alternatives relative to the status quo. Respondent occupation significantly affects choice of the current situation relative to the various alternatives. In particular, people who work in the agriculture or resource based sectors are more likely to prefer the current winegrowing management over alternative management practices. This may be due to apprehension of incurring extra costs or losing income if there is a change in their management or a cultural reason to continue with the current management practises. It is also interesting to observe that in MARLPOP sample, higher income respondents indicate their support for reduced environmental degradation happening in Hawke's Bay region. Lastly, both MARL and MARLPOP respondents who found difficulty in understanding the environmental issues described in the questionnaire have less probability of choosing the two improvement alternatives relative to current winegrowing management.

All of the standard deviation terms are significant for all the models (except for RESORG in HB and NAT10 in MARLPOP and HB) indicating preference heterogeneity does indeed exist.[7] This may be expected given the differing opinions of respondents about the effects of winegrowing management on groundwater quality, wildlife, greenhouse gases emissions and health. As well, respondents' lack of knowledge about the issues may be a contributing factor that increases heterogeneity in respondents' choices.

In order to cope with site-specific features and populations who hold different preferences towards ES attributes, an alternative approach to valuation is estimation through a pooled dataset. The pooled model may integrate systematic differences in value estimates due to site-specific or sampling differences. In addition, it is possible to improve the statistical significance of the attributes where they are found to be insignificant in individual models by increasing the sample size. To accomplish this, a pooled model was estimated by combining HB, MARL, HBPOP and MARLPOP samples data. The results are presented in the last column of Table 4.

The estimated coefficients are found to be highly statistically significant (except RESORG) and of expected signs. The relative magnitudes of the parameter estimates indicate that increasing the attribute levels results in larger positive coefficients. Unlike other individual models, the pooled model also resulted in more significant variables, particularly when interacting with ASC (except GHGE and INCOME). This is an improvement in terms of model prediction as some of the attributes and interactions were found to be insignificant in the models estimated using individual datasets. For example, ASCSATIS is highly significant and negative, indicating that respondents who are satisfied with environmental quality in the region do not want to change the current winegrowing practices. Another variable, ASCJOB is significant and negative, indicating that on average, people who work in the agriculture or resource based sectors are more likely to prefer the current winegrowing management. Finally, all of the standard deviation terms are highly significant at the 1% level indicating preference heterogeneity does indeed exist. Given that the two regions comprise 72% of the

[7] Note that the parameter estimate for the standard deviation of the Cost is exactly same as that of the absolute value of its mean for all the samples which is due to the constraints imposed on the cost distribution.

national productive winegrowing areas, the results of the pooled model could be considered as New Zealand winegrowing regions residents' values of winegrowing environmental impacts.

Insert Table 5 here

Table 5 reports the estimates of mean WTP and Compensating Surplus (CS) derived from the models. There are two main approaches to BT in this study: marginal value transfer and function transfer. The mean WTP for all the attributes can be considered as marginal value transfer (unadjusted WTP), and this assumes that the welfare change experienced by the average person in the study site is the same as that experienced by the average person in the policy site. The function transfers encompass the transfer of a benefit function such as CS from a study site to a policy site that involves combination of multiple attributes using utility models of respondent choice behaviour. This will help to identify the extent to which benefit functions can be transferred between sites. It includes the calculation of the benefits that respondents receive from the environmental condition of winegrowing regions both before and after the change in alternative management that is being proposed. In this study, the function transfer was done without adapting the function to fit the specifics of the policy site such as socioeconomic characteristics. In other words, it is transferred with an absolute term (unadjusted CS) in the first place to test their equivalence between study site and policy site and, subsequently, can be used to forecast a benefit measure for the policy site by adjusting it with socioeconomic characteristics (adjusted CS).

The estimated values are marginal WTP annually for a period of five years for a change (improvement) in the ES attributes concerned, *ceteris paribus*. The mean WTP for all the attributes are positive, implying that on average, respondents have positive utilities (well being) for increases in the quality or quantity of each attribute regardless of whether these improvements occur in their region of residence or in a geographically distant region. The WTP for reduced chemical residues in organically produced wine (RESORG) is not statistically different from zero in all models. Notice that for all the samples, the risks of contamination of groundwater quality are highly valued and are viewed to be the most important attribute.

Values are much higher for the MARL sample even though the mean household income for this sample is not much different in comparison to other samples. A possible reason could be that Marlborough region has experienced very rapid change while growing to become the largest wine growing region in New Zealand. The rapid expansion of vineyards and their environmental impacts may be a nuisance and of concern to many residents. On the other hand, the Hawke's Bay region is the oldest winegrowing region in New Zealand, has experienced slower growth, vineyards are more dispersed in the region and impacts on ES are likely to be less intensive. This notion can be confirmed when comparing the WTP for the selected attributes between regions. It is interesting to observe that the mean WTP of the Hawke's Bay residents valuing their own region are lower than the mean WTP when valuing the Marlborough region. However, there is a completely diametrical point of view in the Marlborough region. The mean WTP for increasing the environmental quality in the vineyards of the region are higher than the mean WTP for increasing the similar environmental quality in the vineyards of the geographically distant Hawke's Bay region. Another reason could be due to the scale effect where the estimated values may be sensitive to the ways in which the issues are presented to the respondents (Bennett 2006). For example, valuing the groundwater quality attributes in Hawke's Bay (4,665 ha) and Marlborough (13,187 ha) regions may vary if the respondents were aware of the scale differences in terms of the size of the region and the intensity of the issues in that area. The pooled model WTP estimates more closely resemble the WTP estimates of the separate models, representing the "average" value of the estimates of the four samples. A closer look at the results revealed that due to larger sample size, the model increases the statistical efficiency of the estimated WTP (i.e., many highly significant coefficients and narrower confidence intervals) relative to individual samples. It can therefore be considered as a more general view of the overall samples and may be appropriate for policy evaluation as will be analyzed in subsequent sections.

As for the CS estimations, four options were created for policy analysis relative to a baseline of current conditions.[8] The first policy option calculated (CS1) includes lower levels of improvements (wine residue – organic, water quality – low risk, GHG reduction – 30%, Native species increase – 10%). The second policy option includes the best levels of improvements. As expected, the CS increases if there is improvement over the current (deteriorating) ES towards better environmental conditions in winegrowing. For a change from current conditions to improved conditions as in Policy 1, on average, respondents in HB are willing to pay NZ$147.35

[8] The CS estimation is based on unadjusted CS without adjusting with socioeconomic variables of the policy site.

each year over five years for the specified ES improvements. In contrast, greater improvements under Policy 2 increases the mean WTP to NZ$164.69. In addition, the results also indicate the importance of attribute tradeoffs when calculating CS for environmental improvements. For instance, Policy 1 and Policy 3 differ only in terms of native wildlife effects (with and without native wildlife improvement). The 'without native wildlife' effect reduces WTP by about 7.2% for Policy 3 compared to Policy 1. Comparing Policy 2 to Policy 4, trading off GHG reduction and native wildlife attributes reduces WTP by about 34%. Overall the respondents on average not only experience positive marginal utility for improvement in the selected ES attributes but also are willing to pay more for higher levels of environmental enhancement.

Insert Table 6 here

Based on these estimated values, it is possible to calculate the typical transfer errors of unadjusted mean WTP (simple absolute value difference) as presented in Table 6. In general, an *a priori* expectation is that study and policy site populations are similar so the errors associated with value (function) transfer should be relatively small. Conversely, where the preferences of study populations differ substantially from those at the policy site both function transfers and absolute value difference transfers may well produce relatively large errors. It is important to observe that attributes RESORG, GHGZERO and NAT10, which had insignificant coefficients in the models, have larger transfer errors, given that the 'true' value is close to zero.[9] Another feature of the results is that, the low average transfer errors in BT Type 4 for HBPOP and MARLPOP samples indicate that these samples are very close in terms of implied WTP for an improvement. This suggests either that people view these regions as close substitute for each other in terms of how they value their own regions, or that they are equally good indicators of value in other areas. Overall, policymakers would probably find the average levels of transfer error in Table 6 (within the range of 30 – 80%) quite acceptable for cost-benefit analysis purposes (Colombo et al. 2007). The following BT validity tests provide more detail study of these transfer errors.

[9] This is because the "true" value, which is assumed to be the value estimated at a particular site using an original study, enters in the formula for estimating the transfer error at the denominator (($|\text{predicted WTP}_{\text{site A}} - \text{observed WTP}_{\text{site A}}|$) / observed $\text{WTP}_{\text{site A}}$).

Benefit Transfer (BT) Tests

The validity of marginal value and benefit function (CS) transfer ought to be tested via statistical hypothesis concerning the equality of the benefit measure for policy and study sites. In other words, the underlying economic assumption is the equality of preferences across sites and populations. The study performs three types of statistical test in order to validate the BT analysis, given broad similarities in the sites and populations characteristics.

(1) Testing if the Model Parameters are Equivalent

A comparison of preference estimates between the two sites needs to allow for the fact that the estimated parameters are confounded with a scale parameter which is inversely proportional to the variance of the random term. The study thus performs a grid search technique as proposed by Swait and Louviere (1993).[10] The test is performed by estimating two models separately, and then a combined model. A likelihood ratio test is then carried out using the log-likelihoods from each model. The test results for each types of BT are shown in Table 7.

Insert Table 7 here

For example, in Type 3 BT, the estimated variance-scale ratio was found to be 1.0 after stacking both HB and MARL datasets, then rescaling the HB data relative to MARL data which implies that the MARL sample has on average the same response variability as the HB sample. The likelihood ratio test statistic for a comparison of the choice model parameters between the HB and MARL is 142. The critical chi-square value of 45 at the 5% significance level (31 degrees of freedom), is well below the calculated value. Therefore it can be concluded that a significant difference does exist between the two sites and we can reject the null hypothesis, even after taking scale differences into account. This means that using the models parameters for BT would be inaccurate or biased. As shown in Table 7, only POOL & MARL and POOL & MARLPOP samples seem to have no significant differences (null hypothesis not rejected) between the models parameters and thus, transferring values from these samples should, in principle, be valid.

[10] The procedure tests if the differences in the model parameters of the two datasets are due to scale parameter. It stacks the two datasets and then rescales one dataset relative to another dataset in order to cancel out the differences.

(2) Testing if the Mean WTP and CS are Equivalent

The complete combinatorial method proposed by Poe et al. (2005) has been carried out to identify where differences might be occurring, given that a significant difference exists in the model's parameters. Results are reported in Table 8.

Insert Table 8 here

The test suggests that there are no significant differences between the HB & HBPOP (Type 2 BT), POOL & HBPOP and POOL & MARLPOP WTP attributes and CS. This means that both marginal value and CS (benefit functions) are equivalent between these samples and suggests that BT would be appropriate. The results indicate that attributes RESZERO and GHG30 are shown to be not significantly different in all the models and hence, it can be implied that marginal value transfer is valid for these attributes for all types of BT across sites and populations. All the CS estimates in Types 1 & 3 BT and POOL & MARL indicate that there is no significant difference suggesting that benefit function transfer is not applicable for these models. However, Kristofersson and Navrud (2005) illustrated that the above results may provide a Type II Error null hypothesis, since welfare estimates with greater variances lead to a greater likelihood of finding transfers invalid (i.e., of failing to reject the null hypothesis that WTPs are the same when it is false). The authors further comment that it is also important for the analyst to choose a tolerance limit when testing for the transferability of the welfare measures. The following section illustrates this test.

(3) Testing if the Mean WTP and CS are Transferable

Johnston and Duke (2008) proposed an alternative equivalence test, denoted as the two one-sided convolutions (TOSC) test which is statistically valid regardless of the empirical distribution of welfare estimates.[11] This test incorporates the complete combinatorial convolutions approach of Poe et al. (2005) as well as the Kristofersson and Navrud (2005) equivalence test with a null hypothesis of WTP divergence (i.e., H_0: $WTP_{HB} - WTP_{MARL} \neq 0$). In order to implement the test, an analyst should choose the tolerance limit of difference between the welfare measures they are willing to accept and calculate the interval of tolerance.

[11] In this study, the empirical distributions of welfare estimates are non-normal. The standard "two one-sided t-test" (TOST) equivalence test may provide erroneous inference and thus, it is inappropriate (Johnston and Duke 2008).

Subsequently, the analyst must calculate the differences of the complete combinatorial of the two WTP distributions (for example, the distributions of WTP at the study and policy sites obtained by using the simulation procedure proposed by Hu et al. (2005)) and test if the resulting difference falls inside or outside the tolerance interval.

Insert Table 9 here

Table 9 shows the TOSC equivalence test results for unadjusted annual mean WTP and CS at $\alpha = 0.10$ using two different tolerance limits (TL) of 50% and 80% for the HB, MARL, HBPOP and MARLPOP samples and treating each as both a study site and as a policy site. For example, if the policy maker is willing to tolerate a 50% difference between the WTP estimated at HB and MARL (Type 3 BT), there is a failure to reject the null hypothesis of different WTP, implying that the two measures cannot be shown to be equivalent, and therefore, transfer is presumed invalid for both marginal value and benefit function transfers. On the contrary, for Type 4 BT, it is noticed that marginal values for groundwater quality can be transferred between MARLPOP and HBPOP samples if the policy maker willing is to tolerate a 50% error.

Insert Table 10 here

Table 10 shows the TOSC equivalence test results when the policy maker chooses an 80% TL. It is easily seen that as the tolerance limit increases to 80%, more WTP attributes seems to be equivalent; suggesting marginal values transfer is warranted across sites and populations. A further finding of the results indicates that only CS3 and CS4 distributions (Type 2 BT between HB and HBPOP samples) are revealed to be equivalent and suggestive of benefit function transfer.

Insert Table 11 here

Table 11 shows the TOSC equivalence test results when using the POOL data to transfer the estimated values. The results suggest that only groundwater quality attributes show no significant differences in transferring the values from POOL to HBPOP and POOL to MARLPOP samples respectively, at 50% TL. As expected, if the TL is increased to 80%, more WTP attributes and CS can be used as BT across sites and populations.

In summary, the empirical findings from this study (at least for the data at hand) suggest that the BT is not a reliable approach for the transfer of winegrowing ES benefits generated using CM. Even though the study controlled for a variety of factors that can affect the accuracy of BT, including the use of the same survey instrument, model specifications, valuing the same resource change, conducting the study in the same time period and similar demographic profile, it is still very difficult to select either one of them as the "study" site for transferring the resulting welfare measures. In addition, the statistical tests typically used in the literature for validating the BT show mixed results and thus, did not demonstrate strong plausibility of transferring the estimated values. For example, for Type 2 BT, the equivalent test for parameters strongly opposed using the model parameters especially for the benefit function transfer. On the contrary, the Poe et al. (2005) tests suggest that the CS are equivalent and in support of transferring the values.

5.0 Policy Implications

The successful application of benefit transfer methods remains a challenge. According to Navrud and Ready (2007), it is established in the literature that the statistical tests of equivalence (validity tests) may not be the most important criterion for deciding the usefulness of value transfer and most researchers are now focusing on the relative size of transfer error. For example, in this study, the TOSC test is applied to examine the tolerance levels of transfer error in a policy context. Although not reported in the study, a 30% TL did not indicate any transferability across sites and populations. As shown in Table 10, 50% and 80% TL may not able to convince policy makers about the merit of BT. Policy makers may seek much lower levels of TL before making policy decisions. The question is does an 80% TL give policy makers the liberty to make decisions on statistical grounds? The results in Table 6 indicate that for the majority of the attributes the average absolute transfer errors are in the range of 10 – 40% and by rights the TOSC tests should not be rejected for these attributes across sites and populations.

Are these transfer errors small enough for the purposes of the policy analyst? Transferring values from a study site to a policy site necessarily increases the errors in those values. This may be due to poorly conducted studies, incorrect model specifications, large measurement error and/or small sample size which results in increased variance of WTP and parameter estimates. This makes it more likely that the hypothesis of statistical equality of study and policy site WTP estimates will be rejected (Tests 1 & 2), even if the relative size of the transfer error may be small. In short, the results in this study deduced that the three tests performed to validate BT are not a good check on transfer reliability.

Analysts must judge how to provide policy advice in a timely manner, subject to the resource constraints they face. The levels of transfer error acceptable to the analyst will vary with their experience and professional judgment regarding the advisability of a policy or project. How expensive would it be to conduct a new study at the policy site? What level of error would the values from a new study have? How critical is preciseness of the attribute values to the analyst? The role of the benefit estimate in the policy process and the costs of a wrong decision are the two major issues that must be addressed when deciding whether to use a benefit transfer method instead of collecting primary data (Bergstrom and DeCivita 1999). Analysts should compare the cost of doing a new, original valuation study with the potential loss from making the wrong decision when based upon transferred estimates.

In light of this study, and the significance of the external costs of winegrowing, estimates of the benefits likely to be delivered by policy implementation should be directly valued and weighted against policy costs. The results in Table 6 may be particularly useful in policy contexts when results from a benefit cost analysis are key to decisions. Nevertheless, caution is advisable before using transferred values for major government policy decisions, in particular concerning public health and safety. Since the groundwater quality attribute is considered the most important and in need of urgent attention, implementing government policy such as Public Health Risk Management Plans in New Zealand (MOF 2005) for improving groundwater quality is crucial. Careful attention should be given to determining whether to conduct a new study based on primary data or using the benefit transfer approach. For example, a benefit cost analysis that uses BT Type 3 to estimate benefits may lead to overestimation (transferring values from MARL to HB) or underestimation (transferring values from HB to MARL) of the net present value of the policy. If the estimated groundwater quality values are transferred from Hawke's Bay (study site) to Marlborough (policy site), Hawke's Bay populations lower WTP estimates may result in underestimation of net benefits from policy change in the Marlborough region. This may lead to large welfare loss and misdirected resources because of the wrong policy decision. In addition, comparing the results from Table 6 and the validity tests, the policy analyst may find it difficult to make decisions due to ambiguous results. The validity tests for BT Type 3 indicate the groundwater attributes are inappropriate to be transferred across these sites and populations whereas typical transfer errors within acceptability range from Table 6 would suggest otherwise. In this case, the BT approach may not be preferred and it is advisable to conduct a new primary study. Indeed the costs of a new study are likely to be much smaller than the cost of a wrong decision.

6.0 Conclusion

The paper focuses on the environmental impact of winegrowing practices in two New Zealand winegrowing regions. The specific contribution of this paper is in employing data and results from a study that was designed with BT validity in mind and comparing different way of testing the validity of BT. The two sites, Hawke's Bay and Marlborough, have similar environmental issues and attributes but are geographically separated. Based on these sites and their populations, the study identified four types of BT tests. This study found a significant portion of the general public lack understanding about winemaking and there is heterogeneity in people's preferences regarding ES that are linked to winegrowing. Despite those limitations, the study found that respondents value programs which result in a significant total reduction in toxic chemical residue content in wine, a reduction of the risk of toxic substances reaching groundwater, a reduction of greenhouse gas emissions and an increase of natural environment and native wildlife populations in vineyards. The overall welfare estimation results show that respondents not only experience greater marginal satisfaction from improvements in these selected ES attributes but also are willing to pay more for higher levels of environmental enhancement.

The second purpose of this paper was to validate the transferability of unadjusted value transfer (WTP) and function transfer (CS) using different types of BT tests. Based on the results of this research, BT via CM is not a preferable approach to transfer values. Although the statistical tests showed conflicting results of whether to follow the transfer process, the findings of the study suggest that assessing how well benefit transfers can predict values at new sites and under which conditions they perform best are very subjective and depends on the professional judgement of the analyst. The statistical tests may be an important tool for examining the basic structure of the data regarding the degree of similarity between samples, but it does not necessarily robustly validate transfers of environmental values across sites and population. Further research in the development of the classical BT approach is necessary and should be investigated before they are recognized as a tool in environmental valuation. Therefore, from a practical point of view, the policy analyst may need to consider the trade-off in using the BT estimates between the risks of (i) under/over-estimated WTP values across sites and populations, (ii) saving in time and money resources of conducting a primary study, and finally, (iii) making costly policy decision mistakes.

References

Adamowicz, W., Louviere, J. and Williams M. 1994 "Combining revealed and stated preference methods for valuing environmental amenities." *Journal of Environmental Economics and Management*, 26(3): 271-292

Bech, M. and Gyrd-Hansen, D. 2005 "Effects coding in discrete choice experiments." *Health Economics* 14: 1079-1083

Bennett, J. 2006. Introduction. In Choice Modelling and the Transfer of Environmental Values, Rolfe, J and Bennett, J (Eds.), Edward Elgar, USA.

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

Colombo, S., Calatrava-Requena, 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., Naughton, M., and Parsons, G. 1992 "Benefit Transfer: Conceptual problems in estimating water quality benefits using existing studies. *Water Resources Research* 28(3): 675-683.

Domencich, T. and McFadden, D. L. 1975 Urban Travel Demand: A Behavioral Analysis. North-Holland Publishing Company.

Hanley, N., Colombo, S., Tinch, D., Black A. and Aftab A. 2006 "Estimating the benefits of water quality improvements under the Water Framework Directive: are benefits transferable?" *European Review of Agricultural Economics* 33(3): 391-413.

Hensher, D. A., Rose, J. M. and Greene, W. H. 2005 Applied Choice Analysis: A Primer. Cambridge University Press, U. K.

Hu, W., Veeman, M. and Adamowicz, W. 2005 "Labelling genetically modified food: Heterogeneous Consumer Preferences and the Value of Information." *Canadian Journal of Agricultural Economics* 53 (1): 79-99.

Johnston, R. J. and Duke, J. M. 2008 "Benefit Transfer Equivalence Tests with Non-normal Distributions". *Environmental and Resource Economics* 41:1–23.

Kristofferson, D. and Navrud, S 2005 "Validity tests of benefit transfer: are we performing the wrong tests?" *Environmental and Resource Economics* 30(3): 279-286.

Louviere, J. J., Hensher, D. A. and Swait, J. D., 2000. Stated Choice Methods: Analysis and Application, Cambridge University Press, Cambridge, UK.

Ministry of Health, New Zealand. 2005. Public Health Risk Management Plans. http://www.moh.govt.nz

Morrison, M. and Bergland, O. 2006. Prospects for the Use of Choice Modelling for Benefit Transfer. *Ecological Economics* 60: 420-28.

Morrison, M. and Bennett, J. 2006. Valuing New South Wales Rivers for Use in Benefit Transfer. In Choice Modelling and the Transfer of Environmental Values, Rolfe, J and Bennett, J (Eds.), Edward Elgar, USA.

Meyerhoff, J. and Liebe, U. 2008 "Do protest responses to a contingent valuation question and a choice experiment differ?" *Environmental and Resource Econ*omics 39: 433-446.

Navrud, S. and Ready, R. (2007). Lessons Learned for Environmental Value Transfer. In Environmental Value Transfer: Issues and Methods, Navrud and Ready (Eds.), Springer, The Netherlands.

Poe, G. L., Giraud, K. L. and Loomis, J. B. 2005 "Computational methods for measuring the difference of empirical distributions." *American Journal of Agricultural Economics* 89(2):353-65.

Street, D. J., Burgess, L. and Louviere, J. J. 2005 "Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments." *International Journal of Research in Marketing* 22: 459-470.

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: 305-14.

Swinton, S. M., Lupi, F., Robertson, G. P. and Hamilton, S. K. 2007 "Ecosystem services and agriculture: Cultivating agricultural ecosystems for diverse benefits." *Ecological Economics* 64: 245-252.

Train, K. E. 1998 "Recreation demand models with taste differences over people." *Land Economics* 74: 230 – 239

Train, K. 2003 Discrete Choice Methods with Simulation. Cambridge University Press, U.K.

van Bueren, M. and Bennett, J. 2006. Generalising Environmental Values: The Case of the National Land and Water Resources Audit in Australia. In Choice Modelling and the Transfer of Environmental Values, Rolfe, J and Bennett, J (Eds.), Edward Elgar, USA.

Table 1 Definition and coding of variables

Variable	Description
Attribute variable	
RESORG	Organic wine with fewer residue levels Effect Coding: 1 if organic wine; 0 if zero residue; -1 if current level
RESZERO	Wine with no detectable residue levels Effect Coding: 1 if zero residue; 0 if organic wine; -1 if current level
WATLOW	Low risk of toxic chemical reaching groundwater Effect Coding: 1 if low risk; 0 if no risk; -1 if high risk
WATNO	No risk of toxic chemical reaching groundwater Effect Coding: 1 if no risk; 0 if low risk; -1 if high risk
GHG30	30% reduction on greenhouse gas emissions per hectare per year Effect Coding: 1 if 30% reduction; 0 if zero reduction; -1 if current level
GHGZERO	Zero greenhouse gas emissions per hectare per year Effect Coding: 1 if zero reduction; 0 if 30% reduction; -1 if current level
NAT10	10% increase of natural environment and native wildlife populations Effect Coding: 1 if 10% increase; 0 if 30% increase; -1 if current level
NAT30	30% increase of natural environment and native wildlife populations Effect Coding: 1 if 30% increase; 0 if 10% increase; -1 if current level
COST	Cost to household per year for the next 5 years - NZ$0, 15, 30, 45, 60, 75, 90
Non-attribute variable	
ASC	Alternative-specific constant on value of 1 for Alternative 1 and 2, and 0 for the current level
SATIS	How satisfied is respondent with environmental quality (1=not; 3=highly)
VISIT	How many times visited the region in the last five years (1= none; 4=>5 times)
CLOSE	How close is respondent from the nearest vineyard (1=>20Km; 5=<200m)
VINLAND	Respondents enjoy vineyards with native plant species (1= strongly disagree; 4=strongly agree)
WQ	Respondents think that winegrowing damages groundwater (1= strongly disagree; 4=strongly agree)
GHGE	Respondents think that winegrowing increase greenhouse gases (1= strongly disagree; 4=strongly agree)
HEALTH	Respondents think that winegrowing leaves dangerous residues in wine (1= strongly disagree; 4=strongly agree)
WINELABEL	Respondents would like wine bottles to be labelled to show environmental friendly practises in winegrowing (1= strongly disagree; 4=strongly agree)
MALE	Respondent sex (1=male; 0=female)
AGE	Respondent age
EDU	Respondent education (1=primary school; 4=degree/professional)
JOB	Respondent occupation (1= based on agriculture sector; 0 = otherwise)
INCOME	Respondent income (1= ≤ $20,000; 6= > $100,000)
UNDER	Respondents think the survey was easy to follow (1= strongly disagree; 4=strongly agree)

Table 2: Principal socio-economic characteristics of survey samples

	HBPOP	HB	Population Census	MARLPOP	MARL	Population Census
Total number of respondents	192	197	147,783	262	301	42,558
Genders (%)						
Males	47	43.2	48.6	49	56.1	49.9
Females	53	56.8	51.4	51	43.9	50.1
Age (mean)	55.6	55.1	37.5#	54.1	53.4	41.7#
Education (%)						
Primary School	1.1	1.2	27.5*	1.7	1.0	25.9*
High School	37.4	41.3	43.5	29.6	36.0	45.3
Trade/technical	22.9	23.2	8.5	30.1	31.0	8.9
Degree/professional	38.6	34.3	9.1	38.5	32.0	8.3
Occupation (%)						
Agricultural/resource	28.0	14.6	19.1	37.8	32.1	20
Manufacturing and transportation	12.3	13.8	18.2	10.9	8.6	20.7
Banking/financial	6.5	4.9	1.4	1.2	1.8	1.2
Education	13.3	10.6	7.8	6.7	8.3	4.7
Health services	12.6	16.9	9.5	9.1	11.8	8.1
Accommodation, retail, and leisure	11.9	13.8	17.4	11.9	12.2	19.5
Government and defence services	6.2	12.1	2.5	6.0	8.3	4.4
Others	9.2	13.4	23.9	16.4	17.0	21.3
Income (%)						
Less than $20000	6.2	14.9	45.4	6.3	11.0	43.7
$20001 to $40000	28.9	26.6	32.0	19.8	24.2	33.5
$40001 to $60000	17.7	18.1	8.7	19.3	23.3	8.9
$60001 to $80000	20.0	15.7	8.6	16.4	15.3	8.5
$80001 to $100000	12.9	10.4	3.0	15.9	12.1	2.9
More than $100000	14.2	14.3	2.3	22.3	14.2	2.5

Note: * - No qualification; # - Median

Source: Populations censuses were obtained from www.stats.govt.nz; Hawke's Bay Region Quarterly Review December 2007 (SNZ); and Marlborough Region Quarterly Review December 2007 (SNZ).

Table 3: General environmental attitudes and beliefs on winegrowing managing

	HBPOP	HB	MARLPOP	MARL
Total number of respondents	192	197	262	301
How satisfied are you with environmental quality in the region (%)?				
Highly satisfied	1.6	14.8	1.3	46.2
Satisfied	22.1	62.0	11.0	41.3
Not satisfied	2.2	14.8	5.2	4.7
Don't know	74.2	7.9	82.6	7.1
How many times visited the region in the last five years (%)?				
More than 5 times	2.1	NA	2.5	NA
2 – 5 times	18.1	NA	13.7	NA
Once	23.2	NA	20.1	NA
None	56.5	NA	63.8	NA
How close is the nearest vineyard to your home (%)?				
Less than 200m	NA	6.4	NA	29.1
Less than 1 Km	NA	24.7	NA	18.0
1-5 Km	NA	43.4	NA	10.4
5-20 Km	NA	19.3	NA	23.1
More than 20 Km	NA	6.3	NA	19.4
I enjoy views of vineyard landscapes that include native plant species (%)				
Strongly agree	27.5	4.3	26.6	4.9
Agree	57.7	47.7	54.2	11.5
Disagree	5.6	43.2	6.4	52.9
Strongly disagree	2.2	4.3	6.6	26.2
Don't know	7.0	4.8	6.2	4.4
Grape growing and winemaking practices are damaging the quality of groundwater (%)				
Strongly agree	2.2	6.8	22.0	3.1
Agree	13.7	37.8	37.0	14.5
Disagree	28.3	12.1	19.2	28.8
Strongly disagree	6.7	4.3	6.2	29.8
Don't know	49.0	39.0	15.6	23.7
Grape growing and winemaking practices are adding to greenhouse gas emissions levels (%)				
Strongly agree	1.7	5.8	7.4	5.9
Agree	15.2	36.0	28.8	18.9
Disagree	30.7	19.2	20.1	29.9
Strongly disagree	12.2	4.3	9.9	11.7
Don't know	40.2	34.7	33.7	33.6

Note: Not applicable (NA) – Did not ask in the survey questionnaire

Weed killers, insecticides and fungicides in grape growing are dangerous to my health in terms of wine residue content (%)				
Strongly agree	5.7	4.9	22.9	5.7
Agree	27.4	30.2	30.0	20.4
Disagree	33.3	32.3	19.6	33.6
Strongly disagree	6.1	6.9	7.7	23.1
Don't know	27.5	25.6	19.7	17.3
I would like wine bottles to be labelled so that I am guaranteed that environmentally sustainable practices have been used (%)				
Strongly agree	22.8	3.8	32.2	2.3
Agree	54.2	15.1	47.5	11.6
Disagree	12.9	51.5	13.4	38.9
Strongly disagree	3.2	26.9	1.2	41.0
Don't know	7.0	2.7	5.7	6.2

Table 4: RPL model results for HB, HBPOP, MARL and MARLPOP

Variable	HB	HBPOP	MARL	MARLPOP	POOL
Random Parameters					
RESORG	-0.1476	-0.0235	-0.0841	-0.0809	-0.0753
RESZERO	0.3162**	0.3372**	0.2647***	0.2202**	0.2476***
WATLOW	0.9633***	1.0975***	0.9059***	0.8639***	0.9054***
WATNO	1.0528***	1.2212***	1.1871***	1.0765***	0.9461***
GHG30	0.5649***	0.3574**	0.2224**	0.4544***	0.2842***
GHGZERO	0.1709	0.5503***	0.4408***	-0.0418	0.2962***
NAT10	-0.0486	0.2582*	0.4712***	0.1843*	0.2524***
NAT30	0.5824***	0.3887***	0.2980***	0.3888***	0.3152***
COST	-0.0385***	-0.0362***	-0.0195***	-0.0263***	-0.0235***
Non-random Parameters					
ASC	-13.8075***	-16.0329***	0.7256	-14.4547***	-4.3076***
ASCSATIS	-0.9589	0.2903	-1.4961***	-0.4947	-0.6557***
ASCVISIT		0.0204		0.4055	
ASCCLOSE	-1.5122***		0.1247		
ASCVINLAND	1.7121***	0.9195**	0.4064	1.2863***	0.5890***
ASCWQ	0.6836	0.9356*	0.0145	1.2695***	0.5161***
ASCGHGE	-0.1013	-0.7729*	0.2620	1.6587***	-0.1516
ASCHEALTH	0.2937	1.7350***	0.5261	0.9586**	0.4823***
ASCWINELABEL	2.2091***	1.9368***	0.5036	-0.5136	0.7922***
ASCMALE	-1.7411***	0.7043	0.8138**	-1.2178**	-0.4133**
ASCAGE	0.0336*	-0.0477***	-0.0283**	0.0143	-0.0169***
ASCEDU	-0.5607*	0.5770*	0.4316	-0.0099	0.2669**
ASCJOB	-1.2128	-0.1570	-1.0253***	0.1571	-0.0796**
ASCINCOME	0.7009***	-0.4174**	-0.1218	0.8135***	-0.0518
ASCUNDER	1.0474*	0.2544	-1.0895***	-1.3794***	-0.5679***
Standard Deviation					
NsRESORG	0.1306	0.8921***	0.6465***	0.7116***	0.5897***
NsRESZERO	0.6587***	0.3364**	0.6850***	0.4539**	0.5313***
NsWATLOW	1.1437***	1.5257***	0.8677***	0.7495***	0.9534***
NsWATNO	0.8849***	1.2116***	1.5643***	1.1668***	1.3039***
NsGHG30	0.7047***	0.6057***	0.7147***	0.4183*	0.7476***
NsGHGZERO	0.8218***	1.1937***	0.5930***	0.5699***	0.8009***
NsNAT10	0.0724	0.9174***	0.8336***	0.3138	0.3575***
NsNAT30	0.8222***	0.8469***	0.9011***	0.8054***	0.7800***
TsCOST	0.0385***	0.0362***	0.0195***	0.0263***	0.0235***
Model Statistics					
N (Observations)	962	974	1509	1134	4551
Log Likelihood	-584.71	-602.64	-962.15	-679.10	-2958.65
McFadden Pseudo R^2 (%)	44.7	43.7	41.9	45.5	40.8
χ^2 (degrees of freedom)	944.30***(31)	934.81***(31)	1391.31***(31)	1133.45***(31)	4082.27***(30)

Notes: Standard errors in parentheses; single (*), double (**) and triple (***) asterisks denote significance at the 10%, 5% and 1% levels respectively.

Table 5: Mean annual WTP per household for the HBPOP, HB, MARLPOP, MARL and POOLED attributes.

Attribute	HBPOP	HB	MARLPOP	MARL	POOLED
RESORG	7.26#	0.75#	2.33#	4.40#	3.49#
	(-9, 24)	(-9, 11)	(-16, 20)	(-19, 26)	(-9, 15)
RESZERO	15.47	10.69	11.91	19.56	14.81
	(1, 31)	(-1, 23)	(-4, 29)	(-3, 47)	(3, 27)
WATLOW	80.12	64.98	91.04	132.44	98.85
	(51, 118)	(45, 88)	(62, 134)	(84, 213)	(75, 127)
WATNO	83.17	67.11	97.82	145.29	100.41
	(54, 120)	(48, 90)	(66, 145)	(87, 237)	(75, 130)
GHG30	29.99	28.40	28.36	39.37	31.05
	(12, 53)	(15, 43)	(13, 50)	(14, 75)	(17, 47)
GHGZERO	34.34	19.68#	12#	48.59	31.54
	(14, 62)	(6, 35)	(-5, 31)	(24, 89)	(17, 48)
NAT10	21.30	10.54#	24.59	55.13	29.49
	(4, 42)	(-0.5, 22)	(9, 45)	(27, 101)	(18, 42)
NAT30	24.64	24.53	31.75	47.81	31.84
	(8, 44)	(11, 39)	(11, 56)	(19, 88)	(17, 49)
CS1	156.17	147.35	227.90	287.54	191.76
	(103, 228)	(111, 193)	(156, 336)	(183, 466)	(154, 237)
CS2	175.13	164.69	235.04	317.44	207.48
	(118, 254)	(125, 216)	(162, 341)	(203, 518)	(163, 258)
CS3	134.87	136.81	203.30	232.41	162.27
	(86, 198)	(103, 183)	(139, 300)	(150, 373)	(128, 201)
CS4	104.88	108.41	174.94	193.03	131.21
	(68, 155)	(82, 146)	(121, 261)	(126, 304)	(102, 165)

Notes: Confidence intervals (CIs) in parentheses at 95% level; the unconditional mean WTPs and CIs are calculated following the simulation procedure proposed by Hu et al. (2005); single (*), double (**) and triple (***) asterisks denote significance at the 10%, 5% and 1% levels respectively; # - non significant coefficients

Table 6: Transfer error (%) for Unadjusted Value Transfer WTP (Absolute Value Difference)

Models	RESORG	RESZERO	WATLOW	WATNO	GHG30	GHGZERO	NAT10	NAT30	AVERAGE
Type 1 BT									
(HBPOP – MARL)/MARL	**65**	21	40	43	24	29	61	48	38
(HBPOP – MARL)/HBPOP	**39**	26	65	75	31	41	159	94	70
(HB – MARLPOP)/MARLPOP	**68**	10	29	31	0.1	**64**	**57**	23	19
(HB – MARLPOP)/HB	**201**	11	40	46	0.1	**39**	**133**	29	25
Type 2 BT									
(HB – HBPOP)/HBPOP	**90**	31	19	19	5	**43**	**51**	0.4	15
(HB – HBPOP)/HB	**868**	45	23	24	6	**75**	**102**	0.4	20
(MARLPOP – MARL)/MARL	**47**	39	31	33	28	**75**	55	34	37
(MARLPOP – MARL)/MARLPOP	**89**	64	46	49	39	**305**	124	51	62
Type 3 BT									
(HB – MARL)/MARL	**83**	45	51	54	28	**60**	**81**	49	46
(HB – MARL)/HB	**487**	83	104	116	39	**147**	**423**	95	87
Type 4 BT									
(HBPOP – MARLPOP)/MARLPOP	**212**	30	12	15	6	**186**	13	22	16
(HBPOP – MARLPOP)/HBPOP	**68**	23	14	18	5	**65**	15	29	17
POOL Data									
(Pool – HB)/HB	**365**	39	52	50	9	**60**	**180**	30	36
(Pool – MARL)/MARL	**21**	24	25	31	21	35	47	33	31
(Pool – HBPOP)/HBPOP	**52**	4	23	21	4	8	39	29	18
(Pool – MARLPOP)/MARLPOP	**50**	24	9	3	10	**163**	20	0.3	11
AVERAGE	175	32	36	39	16	28	59	35	

Note: Bolded transfer errors have insignificant coefficients and are not included in the average.

Table 7: Swait and Louviere (1993) procedure results for testing equivalent of model parameters

Models	Variance-scale ratio	Log Likelihood Ratio $\chi^2 = -2\left(LL_{Model1+Model2} - \left(LL_{Model1} + LL_{Model2}\right)\right)$	Critical Value $(\chi^2_{(\alpha, K_1+K_2-K_{Pool})})$	Reject H_0
Type 1 BT				
HBPOP & MARL	$\mu = 1.2$	$\chi^2 = -2\left(-1601.28 - (-604.64 - 962.15)\right) = 69$	$\chi^2_{(0.05,32)} = 46$	Yes
HB & MARLPOP	$\mu = 1.2$	$\chi^2 = -2\left(-1324.63 - (-584.71 - 679.10)\right) = 122$	$\chi^2_{(0.05,32)} = 46$	Yes
Type 2 BT				
HB & HBPOP	$\mu = 1.2$	$\chi^2 = -2\left(-1230.38 - (-584.71 - 604.64)\right) = 82$	$\chi^2_{(0.05,32)} = 46$	Yes
MARLPOP & MARL	$\mu = 1.2$	$\chi^2 = -2\left(-1684.05 - (-679.10 - 962.15)\right) = 86$	$\chi^2_{(0.05,32)} = 46$	Yes
Type 3 BT				
HB & MARL	$\mu = 1.0$	$\chi^2 = -2\left(-1617.68 - (-584.71 - 962.15)\right) = 142$	$\chi^2_{(0.05,31)} = 45$	Yes
Type 4 BT				
HBPOP & MARLPOP	$\mu = 1.3$	$\chi^2 = -2\left(-1320.06 - (-604.64 - 679.10)\right) = 73$	$\chi^2_{(0.05,31)} = 45$	Yes
POOL Data				
Pool & HB	$\mu = 0.8$	$\chi^2 = -2\left(-3575.93 - (-2958.65 - 584.71)\right) = 65$	$\chi^2_{(0.05,32)} = 46$	Yes
Pool & MARL	$\mu = 1.1$	$\chi^2 = -2\left(-3938.43 - (-2958.65 - 962.15)\right) = 35$	$\chi^2_{(0.05,32)} = 46$	No
Pool & HBPOP	$\mu = 1.3$	$\chi^2 = -2\left(-3603.92 - (-2958.65 - 604.64)\right) = 81$	$\chi^2_{(0.05,32)} = 46$	Yes
Pool & MARLPOP	$\mu = 1.1$	$\chi^2 = -2\left(-3657.58 - (-2958.65 - 679.10)\right) = 40$	$\chi^2_{(0.05,32)} = 46$	No

Table 8: Poe et al. (2005) test results

Models	RESORG	RESZERO	WATLOW	WATNO	GHG30	GHGZERO	NAT10	NAT30	CS1	CS2	CS3	CS4
Type 1 BT												
HBPOP & MARL	0.5836	0.3902	**0.0558**	**0.0429**	0.3129	0.2349	**0.0295**	**0.1068**	**0.0225**	**0.0237**	**0.0377**	**0.0204**
HB & MARLPOP	0.4384	0.4589	**0.1049**	**0.0728**	0.5214	0.7579	**0.0855**	0.3057	**0.0352**	**0.0751**	**0.0536**	**0.0245**
Type 2 BT												
HB & HBPOP	0.2442	0.3069	0.2313	0.2111	0.4631	0.1502	0.1516	0.5055	0.4253	0.4168	0.5413	0.5708
MARLPOP & MARL	0.4370	0.2959	0.1170	0.1127	0.2713	**0.0131**	**0.0467**	0.2196	0.2264	0.1658	0.3357	0.3743
Type 3 BT												
HB & MARL	0.3730	0.2519	**0.0063**	**0.0055**	0.2679	**0.0311**	**0.0012**	**0.0915**	**0.0059**	**0.0066**	**0.0249**	**0.0145**
Type 4 BT												
HBPOP & MARLPOP	0.6628	0.6352	0.3348	0.2843	0.5501	**0.9368**	0.3997	0.3176	**0.0885**	0.1457	**0.0728**	**0.0329**
POOL Data												
Pool & HB	0.6428	0.6858	**0.9768**	**0.9705**	0.6011	0.8617	**0.9889**	0.7547	**0.9258**	0.8958	0.8265	0.8496
Pool & MARL	0.4692	0.3676	0.1513	**0.1073**	0.3322	0.1639	**0.0624**	0.2003	**0.0602**	**0.0517**	**0.0887**	**0.0645**
Pool & HBPOP	0.3533	0.4755	0.8184	0.7941	0.5485	0.4399	0.7829	0.7318	0.8269	0.7881	0.8028	0.8471
Pool & MARLPOP	0.5481	0.6229	0.665	0.5693	0.6096	0.9468	0.6922	0.5167	0.2501	0.3243	0.1821	0.1123

Note: Bolded denote as significance level at *p*-values lower than 0.10 or greater than 0.90 (i.e., Reject the null hypothesis that WTPs or CSs are equivalent)

Table 9: TOSC equivalence test results for unadjusted annual mean WTP and CS for TL of 50%

Models	RESORG	RESZERO	WATLOW	WATNO	GHG30	GHGZERO	NAT10	NAT30	CS1	CS2	CS3	CS4
Type 1 BT												
HBPOP vs MARL - P_L	0.3146	0.3831	0.6050	0.6731	0.3547	0.4196	0.8915	0.6983	0.9741	0.9689	0.8346	0.8416
P_U	0.4749	0.1942	0.0029	0.0019	0.0837	0.0506	0.0059	0.0233	0.0196	0.0181	0.0069	0.0022
MARL vs HBPOP - P_L	0.3534	0.3273	0.3099	0.3477	0.2663	0.2839	0.5942	0.4501	0.9603	0.9572	0.9413	0.9596
P_U	0.5179	0.1552	0.0005	0.0001	0.0489	0.0238	0.0002	0.0034	0.0126	0.0127	0.0238	0.0100
HB vs MARLPOP - P_L	0.5463	0.3317	0.3505	0.4161	**0.1081**	0.0648	0.7971	0.3382	0.9641	0.9051	0.7633	0.8033
P_U	0.4232	0.2587	0.0013	0.0006	**0.1062**	0.4302	0.0281	0.0693	0.0344	0.0585	0.0065	0.0012
MARLPOP vs HB - P_L	0.5137	0.3102	0.1760	0.1959	**0.1083**	0.1135	0.5494	0.2504	0.9251	0.8862	0.9284	0.9401
P_U	0.3914	0.2393	0.0000	0.0000	**0.1065**	0.5676	0.0040	0.0381	0.0145	0.0468	0.0394	0.0088
Type 2 BT												
HB vs HBPOP - P_L	0.7432	0.4729	0.1886	0.1819	0.1524	0.6193	0.6945	0.1431	0.5707	0.5303	0.1512	0.0843
P_U	0.2320	0.1416	0.0061	0.0043	0.0993	0.0357	0.0619	0.1373	0.4213	0.3653	0.1839	0.1243
HBPOP vs HB - P_L	0.6209	0.3726	0.1118	**0.1019**	0.1387	0.4109	0.4927	0.1419	0.4233	0.4866	0.3971	0.2615
P_U	0.1418	0.0933	0.0018	**0.0011**	0.0881	0.0092	0.0209	0.1361	0.2802	0.3235	0.4777	0.3799
MARLPOP vs MARL - P_L	0.5289	0.5409	0.4168	0.4434	0.4022	0.9683	0.8259	0.4804	0.7687	0.8141	0.3789	0.2679
P_U	0.4038	0.1679	0.0070	0.0062	0.0678	0.0053	0.0083	0.0533	0.2216	0.1472	0.1237	0.1049
MARL vs MARLPOP - P_L	0.4986	0.4316	0.2249	0.2414	0.2882	0.7474	0.5246	0.3304	0.7093	0.8139	0.5881	0.5019
P_U	0.3744	0.1097	0.0025	0.0017	0.0346	0.0008	0.0008	0.0214	0.1716	0.1470	0.2662	0.2632
Type 3 BT												
HB vs MARL - P_L	0.6148	0.5956	0.8664	0.9009	0.3865	0.8736	0.9955	0.7075	0.9939	0.9912	0.8788	0.8825
P_U	0.3606	0.1376	0.0000	0.0000	0.0529	0.0049	0.0002	0.0148	0.0058	0.0049	0.0033	0.0008
MARL vs HB - P_L	0.5517	0.4574	0.4566	0.4993	0.2731	0.5751	0.8247	0.4441	0.9864	0.9889	0.9664	0.9653
P_U	0.3027	0.0761	0.0000	0.0000	0.0223	0.0003	0.0000	0.0011	0.0025	0.0039	0.0183	0.0054
Type 4 BT												
HBPOP vs MARLPOP - P_L	0.2343	0.1472	0.1224	0.1441	0.1089	0.0051	0.2742	0.3501	0.8993	0.8187	0.7077	0.7510
P_U	0.5451	0.3509	0.0222	0.0157	0.1613	0.6276	0.1342	0.0899	0.0777	0.1150	0.0118	0.0033
MARLPOP vs HBPOP - P_L	0.3019	0.1866	**0.0875**	**0.0937**	0.1201	0.0269	0.2333	0.2663	0.8533	0.7678	0.8882	0.9339
P_U	0.6234	0.4150	**0.0139**	**0.0086**	0.1760	0.8657	0.1084	0.0563	0.0504	0.0843	0.0457	0.0157

p-values lower than 0.10 indicate no differences in the two distributions (bolded)

Table 10: TOSC equivalence test results for unadjusted annual mean WTP and CS for TL of 80%

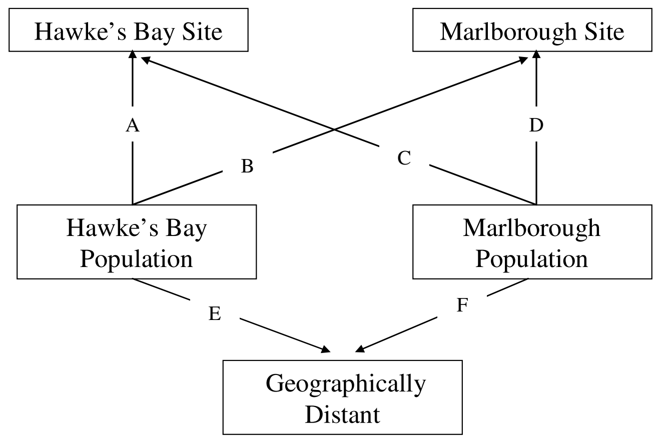
Models	RESORG	RESZERO	WATLOW	WATNO	GHG30	GHGZERO	NAT10	NAT30	CS1	CS2	CS3	CS4
Type 1 BT												
HBPOP vs MARL - P_L	0.2601	0.2635	0.3302	0.4054	0.1982	0.2328	0.8013	0.5406	0.9719	0.9636	0.6906	0.6466
P_U	0.4100	0.1139	0.0006	0.0003	0.0290	0.0170	0.0020	0.0073	0.0179	0.0153	0.0029	0.0008
MARL vs HBPOP - P_L	0.3174	0.1948	**0.0799**	**0.0997**	0.1111	**0.1044**	0.2687	0.2065	0.9453	0.9402	0.9248	0.9406
P_U	0.4781	0.0749	**0.0000**	**0.0000**	0.0114	**0.0054**	0.0000	0.0001	0.0089	0.0087	0.0179	0.0066
HB vs MARLPOP - P_L	0.5371	0.2272	0.1192	0.1485	**0.0332**	0.0254	0.6953	0.1757	0.9636	0.8919	0.5868	0.5912
P_U	0.4142	0.1656	0.0000	0.0000	**0.0228**	0.2393	0.0126	0.0186	0.0340	0.0499	0.0019	0.0001
MARLPOP vs HB - P_L	0.4849	0.1998	**0.0274**	**0.0286**	**0.0333**	0.0671	0.2789	**0.0936**	0.8890	0.8571	0.9155	0.9048
P_U	0.3636	0.1427	**0.0000**	**0.0000**	**0.0229**	0.4388	0.0003	**0.0049**	0.0083	0.0344	0.0326	0.0046
Type 2 BT												
HB vs HBPOP - P_L	0.7355	0.3401	**0.0434**	**0.0381**	**0.0505**	0.4511	0.5785	**0.0495**	0.5684	0.4991	**0.0631**	**0.0237**
P_U	0.2249	0.0802	**0.0002**	**0.0000**	**0.0234**	0.0122	0.0335	**0.0398**	0.4189	0.3352	**0.0664**	**0.0275**
HBPOP vs HB - P_L	0.5299	0.2072	**0.0147**	**0.0128**	**0.0424**	0.1742	0.2648	**0.0488**	0.3388	0.4301	0.3617	0.1832
P_U	0.0974	0.0368	**0.0000**	**0.0000**	**0.0183**	0.0011	0.0045	**0.0390**	0.2059	0.2713	0.4392	0.2741
MARLPOP vs MARL - P_L	0.5085	0.4387	0.1801	0.2062	0.2359	0.9471	0.6930	0.3056	0.7657	0.8012	0.2380	0.1341
P_U	0.3838	0.1130	0.0019	0.0013	0.0236	0.0033	0.0028	0.0178	0.2187	0.1368	0.0596	0.0411
MARL vs MARLPOP - P_L	0.4600	0.2788	**0.0558**	**0.0656**	0.1204	0.4142	0.2205	0.1425	0.6672	0.8009	0.5407	0.4279
P_U	0.3385	0.0527	**0.0006**	**0.0003**	0.0082	0.0002	0.0000	0.0040	0.1435	0.1365	0.2279	0.2068
Type 3 BT												
HB vs MARL - P_L	0.6072	0.4954	0.6431	0.7213	0.2208	0.7706	0.9912	0.5407	0.9938	0.9896	0.7658	0.7333
P_U	0.3534	0.0903	0.0000	0.0000	0.0133	0.0015	0.0000	0.0033	0.0057	0.0041	0.0009	0.0001
MARL vs HB - P_L	0.5055	0.2904	0.1243	0.1524	**0.1087**	0.2473	0.4586	0.1988	0.9786	0.9850	0.9599	0.9447
P_U	0.2638	0.0312	0.0000	0.0000	**0.0027**	0.0000	0.0000	0.0000	0.0014	0.0028	0.0151	0.0028
Type 4 BT												
HBPOP vs MARLPOP - P_L	0.1828	0.0746	**0.0245**	**0.0292**	**0.0336**	0.0016	0.1383	0.1899	0.8915	0.7946	0.5050	0.4975
P_U	0.4711	0.2070	**0.0035**	**0.0023**	**0.0526**	0.3518	0.0558	0.0320	0.0716	0.0990	0.0047	0.0013
MARLPOP vs HBPOP - P_L	0.2816	0.1144	**0.0132**	**0.0129**	**0.0401**	0.0157	**0.1002**	**0.1065**	0.8075	0.7043	0.8584	0.9029
P_U	0.6039	0.2908	**0.0023**	**0.0009**	**0.0629**	0.8032	**0.0375**	**0.0127**	0.0351	0.0587	0.0343	0.0101

p-values lower than 0.10 indicate no differences in the two distributions (bolded)

Table 11: TOSC equivalence test results for unadjusted annual mean WTP and CS for TL of 50% and 80% using POOL data

Models	RESORG	RESZERO	WATLOW	WATNO	GHG30	GHGZERO	NAT10	NAT30	CS1	CS2	CS3	CS4
TL 50%												
Pool vs HB - P_L	0.2768	0.0908	0.0000	0.0000	0.0434	0.0074	0.0000	0.0134	0.0672	0.0708	0.0072	**0.0022**
P_U	0.5549	0.3508	0.1730	0.1662	0.1126	0.3603	0.6997	0.2062	0.9178	0.8518	0.1917	**0.1060**
Pool vs MARL - P_L	0.4726	0.4043	0.2779	0.3880	0.3041	0.4929	0.7004	0.4627	0.9354	0.9313	0.5953	0.5462
P_U	0.4110	0.1763	0.0009	0.0012	0.0679	0.0179	0.0042	0.0319	0.0559	0.0383	0.0039	0.0009
Pool vs HBPOP - P_L	0.5797	0.2337	**0.0042**	**0.0046**	0.0984	0.1785	0.0246	0.0307	0.1623	0.1677	0.0201	0.0074
P_U	0.2905	0.1970	**0.0656**	**0.0603**	0.1286	0.0983	0.2734	0.2324	0.8162	0.7362	0.2644	0.1801
Pool vs MARLPOP - P_L	0.3866	0.1509	**0.0143**	**0.0238**	0.0675	0.0049	0.0434	0.1263	0.7366	0.6179	0.3766	0.3807
P_U	0.4808	0.3313	**0.0219**	**0.0173**	0.1408	0.6319	0.1793	0.1267	0.2369	0.2694	0.0085	0.0012
TL 80%												
Pool vs HB - P_L	0.2333	0.0344	**0.0000**	**0.0000**	**0.0076**	0.0011	0.0000	**0.0008**	0.0629	0.0553	**0.0012**	**0.0000**
P_U	0.5000	0.1846	**0.0038**	**0.0046**	**0.0183**	0.1133	0.2779	**0.0467**	0.9128	0.8203	**0.0243**	**0.0055**
Pool vs MARL - P_L	0.4382	0.2820	**0.0994**	0.1689	0.1556	0.2849	0.4959	0.2807	0.9326	0.9197	0.3801	0.2939
P_U	0.3769	0.1003	**0.0000**	0.0000	0.0153	0.0022	0.0005	0.0061	0.0535	0.0315	0.0003	0.0000
Pool vs HBPOP - P_L	0.5377	0.1176	**0.0008**	**0.0004**	**0.0286**	**0.0676**	**0.0058**	**0.0053**	0.1567	0.1442	**0.0062**	**0.0024**
P_U	0.2559	0.0928	**0.0011**	**0.0012**	**0.0280**	**0.0220**	**0.0752**	**0.0631**	0.8097	0.7019	**0.0555**	**0.0159**
Pool vs MARLPOP - P_L	0.3490	0.0759	**0.0031**	**0.0032**	**0.0190**	0.0018	**0.0118**	**0.0395**	0.7286	0.5827	0.1795	0.1569
P_U	0.4404	0.1880	**0.0001**	**0.0001**	**0.0283**	0.3197	**0.0371**	**0.0308**	0.2293	0.2382	0.0005	0.0000

p-values lower than 0.10 indicate no differences in the two distributions (bolded)



Type 1: Differences across populations only (B vs D and A vs C)

Type 2: Differences across sites only (A vs B and C vs D)

Type 3: Differences across sites and equivalent populations but geographic scale (A vs D, D vs F, and A vs E)

Type 4: Differences across sites and different populations (B vs C)

Figure 1: Types of benefit transfer tests