Calibration of values in benefit transfer to account for variations in geographic scale and scope: Comparing two choice modelling experiments

John Rolfe[1], Jill Windle[1], Jeff Bennett[2] and Kasia Mazur[2]

[1] *Centre for Environmental Management*
CQUniversity, Rockhampton.
[2] *Crawford School of Economics and Government,*
Australian National University, Canberra.

Contributed paper presented at the 57th Australian Agricultural and Resource Economics (AARES) 2013 Annual conference, 5 – 8 February 2013, Sydney, Australia.

Abstract

Two choice modelling studies in Australia were designed to test for the effects of variations in geographic scale and scope on WTP values. One case study assessed values for improved natural resource management in a river catchment, and the other assessed values for improved protection of the Great Barrier Reef. The results show that increases in the amount of an amenity offered are valued positively and display diminishing marginal utility. Unit value estimates vary inversely with increases in the geographic scope over which an amenity improvement was offered. In the case studies, marginal values for the same unit of environmental improvement could be several thousand times higher when only very small areas were considered compared to when the whole amenity was framed. These results confirm that calibration factors are needed in benefit transfer applications between different geographic scopes. A close inverse relationship was identified between the ratio of quantities involved and the ratio of the WTP amounts. A log-log form of this relationship is recommended as a simple and efficient way of performing this calibration.

1. Introduction

Time and cost savings make it attractive for analysts and policy-makers to transfer value estimates from non-market valuation studies into other situations, a process known as benefit transfer (Brouwer 2000; Rolfe and Bennett 2006; Johnston and Rosenberger 2010). This is particularly relevant to stated preference experiments, where substantial time and effort is required to conduct primary studies. While there is general agreement that benefit transfer should be appropriate under idealistic conditions where the source and target case studies and relevant populations are identical (Boyle and Bergstrom 1992), this cannot be achieved in real world applications. Substantial effort has been invested in identifying how the accuracy of benefit transfer may be sensitive to variation between source and target case studies and populations, and to identify adjustments to improve the benefit transfer process (Brouwer 2000; Rolfe and Bennett 2006; Johnston and Rosenberger 2010).

Methods of benefit transfer include the simple transfer of unit values for an amenity of interest and the transfer of a benefit function, with the latter approach often preferred because these allow for adjustments to be made for a variety of other influences including site and population differences (Loomis 1992; Rolfe 2006; Johnston and Rosenberger 2010). A number of reviews of benefit transfer applications involving stated preference techniques (for example, Brouwer 2000; Bergland *et al.* 2002; Rolfe and Bennett 2006) have noted that tests of convergent validity are difficult to satisfy even when there are only modest differences between source and target sites. Transferring benefit functions also allow calibration factors to be employed where preferences are adjusted to take account of variations between source and target studies (Smith *et al.* 2002, 2006; Johnston and Rosenberger 2010).

A particular area of interest for calibration in benefit transfer studies is the potential for unit values to vary according to the amount of the amenity being valued and the extent of the context in which the amenity is being offered (Johnston and Rosenberger 2010). Where the size of the trade-off and its context are different between source and target studies, then sensitivity to unit value differences makes the benefit transfer process problematic without calibration. There is evidence that unit values are much higher for amenities when they are offered in a limited context compared to wider scoped amenities, and that there is potential for adjustment factors to be used in any subsequent benefit transfer process (e.g. van Bueren and Bennett 2004). However the reasons why values may be sensitive to such 'scope variation' and the identification of more general rules for preference calibration remain areas for further study.

The focus of this paper is to compare the results of two choice modelling studies in Australia that assessed values for different scopes of environmental protection. The first study involved assessing community values for improved natural resource management in the Namoi River catchment in New South Wales where improvements in 10 per cent, 50 per cent and 100 per cent of the catchment were considered (Mazur and Bennett 2009; Mazur 2011). The second study explored issues relating to the environmental health of the iconic Great Barrier Reef (GBR) in Queensland where values for improved protection in local, regional and the whole GBR case studies were assessed (Rolfe and Windle 2010a,b). The comparison of these studies helps to identify how unit values are sensitive to the scope of the context in which they are valued, and whether there is potential for the systematic use of related adjustment factors in benefit transfer applications.

The results show that values vary inversely with increases in the geographic scope of an amenity. In the case studies, marginal values for the same unit of environmental improvement could be several thousand times higher when the context involved only very small areas (e.g. a small percentage of the catchment or a local reef) were considered compared to when the amenity on offer was framed within a larger geographic context (the whole catchment or reef). These results confirm that calibration factors are needed in benefit transfer applications between different geographic scopes. An analysis of the relationship between quantities involved and subsequent WTP estimates in the different scope tests has allowed one potential form of calibration to be identified.

The paper is outlined as follows. In the next section, conceptual issues are discussed and research hypotheses outlined. Background information about the two case studies is presented in the third section. The results of the comparative scale and scope tests are presented in the fourth section followed by tests of calibration for scope effects. Discussion and conclusions follow in the final section.

2. Conceptualising the issues

Early development of the contingent valuation technique identified that values were sometimes insensitive to changing dimensions of and the context in which an amenity was being offered, an effect termed as part-whole bias (Mitchell and Carson 1989) or scope insensitivity (Arrow *et al.* 1993). This effect generated heated and divergent views about the usefulness of the contingent valuation method (Schulze *et al.* 1998), with the NOAA panel review recommending the conduct of a so-called 'scope test' to ensure that respondents understood the scenarios accurately (Portney 1994). A number of subsequent contingent valuation studies have passed the test; failed the test or had mixed results (Bateman *et al.* 2004; Heberlein *et al.* 2005; Czajkowski and Hanley 2009).

The broad definition of scope insensitivity used in the contingent valuation literature has been the source of some confusion. To clarify the situation Mazur and Bennett (2009) and Rolfe and Wang (2011) define two different effects relating to the dimensions of the amenity on offer and its context. Scale refers to the quantity of the amenity being considered. Sensitivity to scale means that more of the amenity is valued more highly. The possibility of diminishing marginal utility is also part of scale sensitivity. Scope refers to the context in which values are being estimated. Sensitivity to scope involves value estimates being different when the context or frame of the valuation exercise varies. The context can involve space, time or policy dimensions.

Scale and scope effects are often intertwined, as increasing amounts of an amenity (scale) can also involve changes in the extent of the policy, time or space in which the good is being offered (scope) and vice versa. An example of the latter is where increasing the geographic region over which an amenity is to be improved (an increase in scope) necessarily involves more of the amenity being provided (an increase in scale).

Insensitivity to changes in the scale of a good being valued have been the focus of part-whole bias and 'scope' tests with the contingent valuation technique. Much of the early debate was about whether 'scope' insensitivity (i.e. the estimate of value for a smaller scale of provision being equal to the value estimated for a larger scale of provision) invalidated the use of stated preference approaches. Both theoretical reasons (e.g. Hanemann 1994; Randall and Hoehn

1996) and methodological considerations (e.g. Heberlein *et al.* 2005: Czajkowski and Hanley 2009) were advanced to explain why scale insensitivity may occur.

The development of the choice modelling technique has helped to address scale issues because the repeated presentation of attributes at varying levels makes it clear that scale varies and a significant coefficient indicates that an internal scale test has been satisfied (Hanley *et al.* 1998; Rolfe *et al*. 2002). Typically only a single (average) value estimate is generated for each attribute in a choice experiment, although some reduction of values with increased amounts could be expected as a consequence of diminishing marginal utility. Larger scale effects between studies, where 'average' marginal values for small changes are typically higher than 'average' marginal values for large changes, are expected as a consequence of diminishing marginal utility (Hoehn and Randall 1989; Hoehn 1991). This can be termed an external test of scale effects, where 'average' marginal value estimates are lower when larger amounts of the asset are involved across choice modelling experiments (Rolfe and Wang 2011). The tests can be visualised as involving different sections around a utility frontier, where the average slope changes for each section or range of change that is considered.

To reflect these scale issues, we identify two specific scale hypotheses to test in applications of choice modelling experiments. First, an internal scale test is applied to test that greater quantities of attributes representing an improvement are preferred:

$$H_1: \ \beta_1 ATT_1 > 0$$

where ATT_1 refers to the attribute being valued and β_1 is the attribute coefficient calculated in the choice model. Second is an external scale test where it is expected that 'average' (across the range of the amenity presented to respondents) marginal benefit estimates diminish as the overall size (scale) of the asset involved increases.

$$H_2: \text{Marginal WTP}_{ATTsmall} > \text{Marginal WTP}_{ATTmedium} > \text{Marginal WTP}_{ATTlarge}$$

where Marginal WTP_{ATT} refers to the implicit price for an attribute within a choice experiment.

In terms of scope effects, the default assumption in the benefit transfer of stated preference values is that scope effects have little impact on value estimates. This allows analysts to transfer unit values estimated for instance at one level of geographic scope (i.e. a local river catchment) to target sites at different scope levels (i.e. a regional river catchment). If this default assumption does not hold, then benefit transfers across scopes should also involve some application of adjustment factors to take account of the impacts on unit value estimates (van Bueren and Bennett 2004; Johnston and Rosenberger 2010).

Scope effects are likely to be generated when there are changes in the frame or context of the amenity of interest and hence the pool of substitute and complement goods that may be considered. Respondents may consider different substitutes and complements as the set of resource possibilities (scope) expands (Pate and Loomis 1997; Rolfe *et al.* 2002; Hanley *et al.* 2003; Bateman *et al.* 2006), particularly if there are ordering effects (Randall and Hoehn 1996), or sub-additivity effects (Hanemann 1994). Values may also be higher with smaller scoped tradeoffs because respondents are more familiar or comfortable with such contexts, identify different substitutes, or because of different perceptions of responsibility and effectiveness (Pate and Loomis 1997; Rolfe *et al.* 2002; Hanley *et al.* 2003; Bateman *et al.*

2006; Rolfe and Wang 2011). Varying geographic scopes may have implications for loyalty, proximity, and perceptions of responsibility and institutional arrangements (Rolfe and Wang 2011). These factors mean that increased scope is likely to be associated with lower marginal values.

However, the evidence about geographic scope effects is mixed. A number of benefit transfer reviewers (e.g. Brouwer 2000; Bergstrom *et al.* 2002; Rolfe and Bennett 2006) have noted that tests of convergent validity for value estimates are often difficult to satisfy even though the amounts involved (scale) are similar. This suggests that scope and other context differences may have a significant influence on values. Many convergent validity tests have confirmed that values are equivalent between similar scoped amenities, such as two similar size rivers (Hanley *et al.* 2006), river catchments (Rolfe *et al.* 2006), or regional and sub-regional protection of vegetation, soils and waterways (Rolfe and Windle 2008). Other studies appear to show significant scope effects, such as for land and water protection between regional and national levels in Australia (van Bueren and Bennett 2004).

To reflect these scope issues, we identify a specific hypothesis to test whether marginal value estimates are sensitive to changes in scope:

H_3: Marginal WTP $_{SCOPEsmall}$ > Marginal WTP$_{SCOPElarge}$

where Marginal WTP $_{SCOPE}$ refers to the implicit price of attribute changes at a particular level of scope.

Significant scale and scope effects mean that the marginal values of attributes measured in a smaller geographic context will be larger than values measured in a larger context. The use of adjustment or calibration factors has been suggested to help transfer values between geographic or other contexts (Smith *et al.* 2002; van Bueren and Bennett 2004; Smith *et al.* 2006, Johnston and Rosenberger 2010), with two main approaches suggested. The first has been the use of meta-analysis to identify how scale and scope factors may have systematic influences on values (e.g. Loomis 2006; Rolfe and Brouwer 2011). The second approach has been to calculate specific adjustment factors so that a valuation at one level of scope can be adjusted for benefit transfer to another level of scope. Under this second approach adjustment factors of up to 26 times have been identified to convert from a national to a regional frame (van Bueren and Bennett 2004.

The theoretical justifications for calibrating benefit transfer values to account for 'scope' differences (including both scope and scale effects) are that diminishing marginal utility can be expected as the quantity (scale) of the amenity increases (Randall and Hoehn 1996), and that changes in the context of the amenity on offer (scope) can lead to different substitutes and tradeoffs being considered (Mazur and Bennett 2009; Rolfe and Wang 2011). However the estimation of calibration factors for 'scope' changes is complex as the ratio of marginal values approach suggested by van Bueren and Bennett (2004) may be confounded with case study influences on value estimates, and to date there has been no systematic testing of the convergent validity of such calibrating factors. To begin this process it is possible to explore whether adjustments based on values are related in some systematic way to the geographic scope parameters (i.e. the geographic area involved). This test is summarised in a fourth hypothesis:

H_4: WTP $_{ATTsmall}$ / WTP $_{ATTlarge}$ = f [Quantity $_{ATTlarge}$ / Quantity $_{ATTsmall}$]

3. Case study details

The hypotheses of interest in this study are tested with data from two case studies. The first study explored issues related to natural resource management in the Namoi River catchment within the Murray-Darling Basin in New South Wales. It covers an area of 42,000 sq km (Figure 1). The second study explored issues relating to the environmental health of the iconic GBR in Queensland which covers an area of 346,000 sq km (Figure 1).

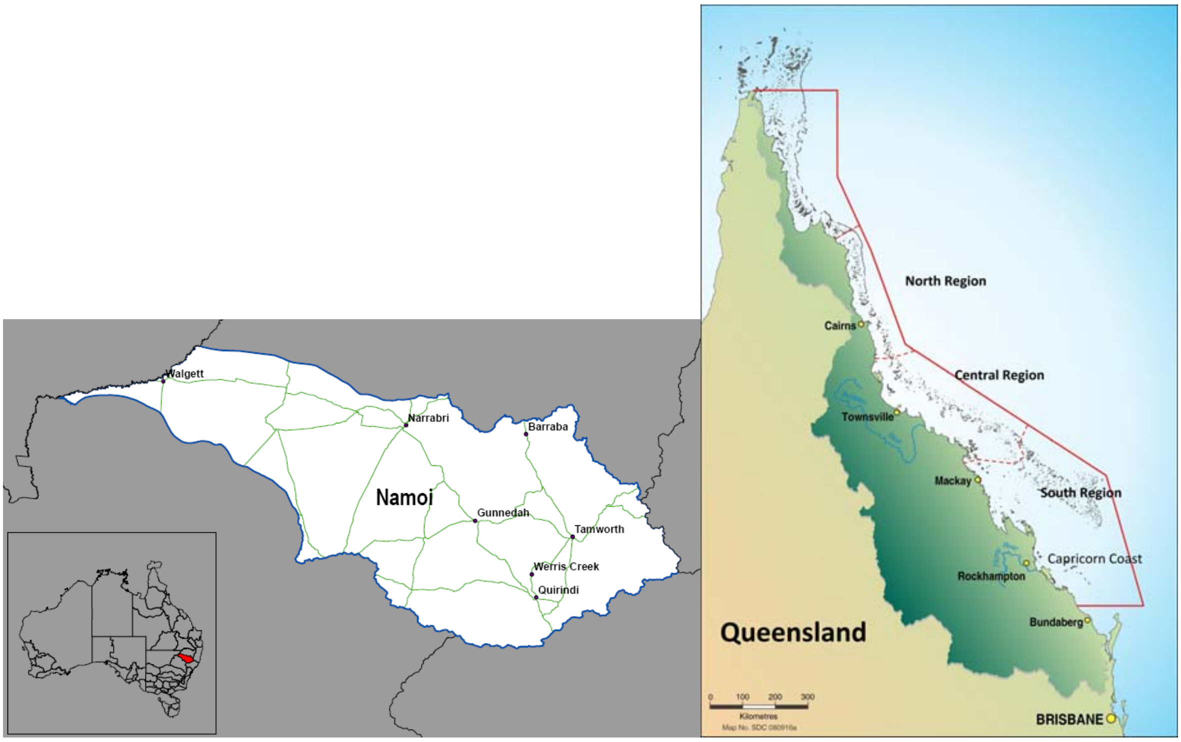


Figure 1 The Namoi catchment and Great Barrier Reef study areas in Australia

Both studies applied choice modelling experiments to elicit community values for environmental improvements and used split sample surveys to examine how WTP estimates varied across three levels of geographic scope. In each study, the same set of attributes was used in each split sample survey, but the level of attributes (scale) differed according to the geographical scope involved. For most attributes, there were larger amounts (scale) involved as the geographic scope increased. The main features of the two studies are outlined in Table 1 below. In the interests of brevity only relevant details of each study are reported here, with further detail of the catchment study available in Mazur and Bennett (2009) and Mazur (2011), and of the GBR study in Rolfe and Windle (2010a,b). More specific details on how the attribute levels (scale) varied with the different scopes of the experiments are provided in Appendix 1.

There was a difference between the two studies in terms of the way scope was defined. The Namoi catchment study focused on three levels of scope – 100%, 50% and 10% of the catchment area. There was little difference in the choice sets apart from some small changes in the wording of the choice set question (Figure 2a). The wording for the 100% scope referred to "in the Namoi catchment" while the 50% context refers to "in parts of the Namoi catchment" and the 10% context to "on selected farms in the Namoi catchment". The GBR study also covered three levels of scope – the whole GBR (100%), a regional section of the GBR (25%) and a local section (3% or less). The geographical area of the GBR that was relevant to each experiment was specified in each survey (Figure 2b).

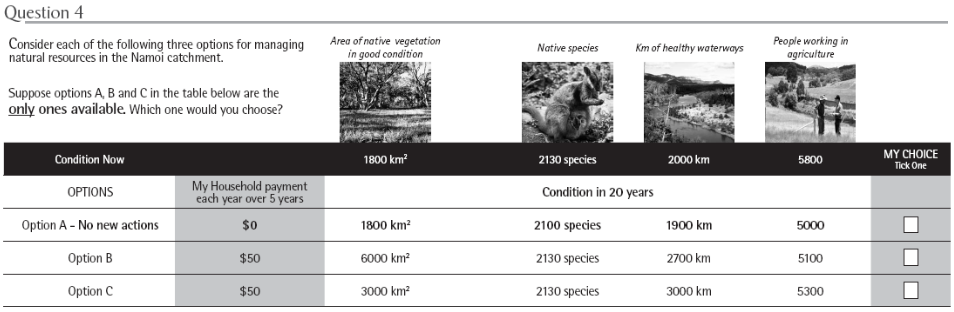
Table 1 Study area design features

The Namoi River catchment study	The Great Barrier Reef study:

Choice design:
• Framed with a future base (20 years)
• Status Quo plus 2 improvement options
• Five attributes:
  o Cost (annual payment for five years)
  o Area of Native vegetation in good condition
  o Native species
  o Healthy waterways
  o People working in agriculture
• Attribute levels described in absolute values only
• 5 choice sets
	Choice design:
	• Framed with a future base (25 years)
	• Status Quo plus 3 improvement options
	• Four attributes:
	  o Area of coral reef in good health
	  o No of fish species in good health
	  o Area of seagrass in good health
	  o Cost (annual payment for five years)
	• Attribute levels described in both absolute and relative (%) terms
	• 6 choice sets

Scope dimensions:
• Three levels of geographic scope:
  o Natural resource management in the Namoi catchment (100%)
  o Natural resource management in parts of the Namoi catchment (50%)
  o Natural resource management on selected farms in the Namoi catchment (10%)
• The non-cost attribute levels varied according to the proportions of the scope involved
• The cost attribute levels remained constant across all surveys and were not adjusted for scope
	Scope dimensions:
	• Three levels of geographic scope:
	  o Whole GBR (100%)
	  o Regional section of the GBR (25% each)
	    ▪ northern, central and southern sections
	  o Local case studies: Cairns (North) (0.4%); Townsville (Central) (3%); and Capricorn Coast (South) (0.7%)
	    ▪ Current and future conditions levels varied
	    ▪ Attribute levels varied but overall improvement remained constant across attributes and cases studies (total improvement of 25%)
	• In the regional studies the non-cost attribute levels were 25% of the whole GBR study
	• In the local studies the non-cost attribute levels reflected actual situations, and varied slightly from the share of the GBR involved
	• The cost attribute levels remained constant across all surveys and were not adjusted for scope

Population samples:
• Two population samples:
  o Local catchment population (Tamworth, Gunnedah)
    ▪ Valuation scope: 100% and 10%
  o Distant urban population (Sydney)
    ▪ Valuation scope: 100%; 50% and 10%
	Population samples:
	• Two population samples:
	  o Local catchment population (Cairns – North; Townsville – Central; and Capricorn Coast – South
	    ▪ Valuation scope: Townsville: whole, regional and local; Cairns & Cap. Coast: local only
	  o Distant urban population (Brisbane)
	    ▪ Valuation scope: Whole, regional, local



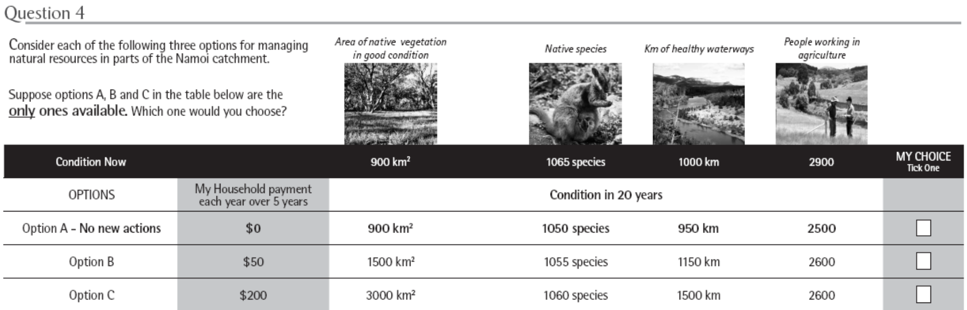
Question 4

Consider each of the following three options for managing natural resources in the Namoi catchment.

Suppose options A, B and C in the table below are the only ones available. Which one would you choose?

Condition Now		1800 km²	2130 species	2000 km	5800	MY CHOICE Tick One
		Area of native vegetation in good condition	Native species	Km of healthy waterways	People working in agriculture	
OPTIONS	My Household payment each year over 5 years	Condition in 20 years				
Option A - No new actions	$0	1800 km²	2100 species	1900 km	5000	☐
Option B	$50	6000 km²	2130 species	2700 km	5100	☐
Option C	$50	3000 km²	2130 species	3000 km	5300	☐

100%



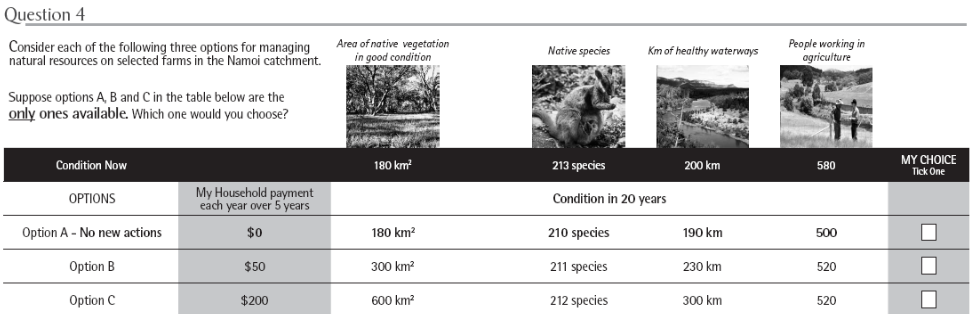
Question 4

Consider each of the following three options for managing natural resources in parts of the Namoi catchment.

Suppose options A, B and C in the table below are the only ones available. Which one would you choose?

Condition Now		900 km²	1065 species	1000 km	2900	MY CHOICE Tick One
		Area of native vegetation in good condition	Native species	Km of healthy waterways	People working in agriculture	
OPTIONS	My Household payment each year over 5 years	Condition in 20 years				
Option A - No new actions	$0	900 km²	1050 species	950 km	2500	☐
Option B	$50	1500 km²	1055 species	1150 km	2600	☐
Option C	$200	3000 km²	1060 species	1500 km	2600	☐

50%



Question 4

Consider each of the following three options for managing natural resources on selected farms in the Namoi catchment.

Suppose options A, B and C in the table below are the only ones available. Which one would you choose?

Condition Now		180 km²	213 species	200 km	580	MY CHOICE Tick One
		Area of native vegetation in good condition	Native species	Km of healthy waterways	People working in agriculture	
OPTIONS	My Household payment each year over 5 years	Condition in 20 years				
Option A - No new actions	$0	180 km²	210 species	190 km	500	☐
Option B	$50	300 km²	211 species	230 km	520	☐
Option C	$200	600 km²	212 species	300 km	520	☐

10%

Figure 2a Example choice sets in the catchment study

In both studies, surveys were collected from a local population sample as well as from respondents in the more distant state capital city. Details of survey sampling are presented in Table 2.

Table 2. Survey sample sizes

	Whole	**Medium**	**Small**
Catchment survey scope	**100%**	**50%**	**10%**
Local (Namoi)	268		272
Distant (Sydney)	255	258	249
GBR survey scope	**100%**	**25%**	**<1% to3%**
Local (Cairns) – north		73	72
Local (Townsville) – central	587	89	144
Local (Capricorn Coast) – south		72	73
Distant – Brisbane	250		
North		160	160
Central		244	159
South		159	160

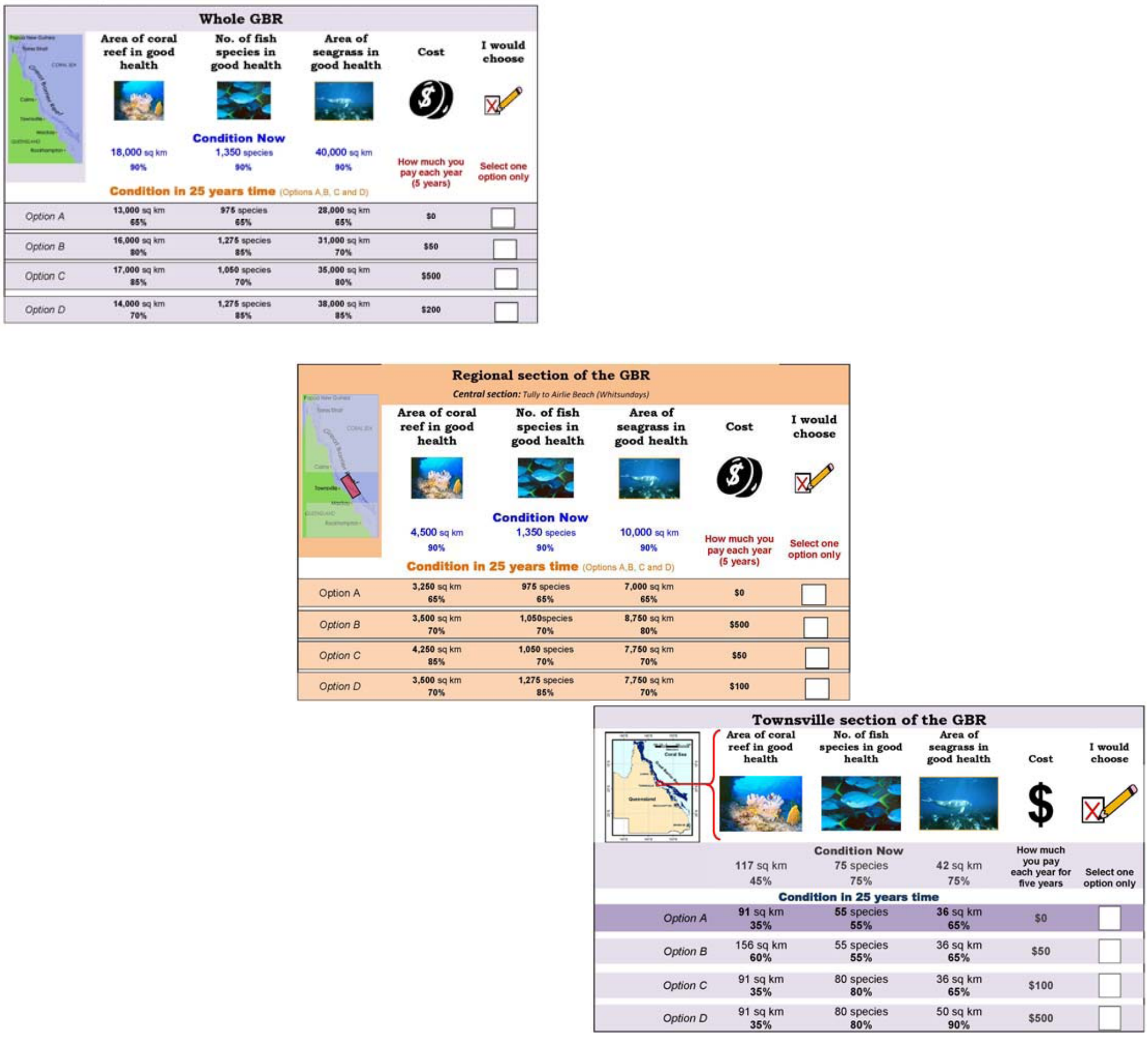


Figure 2b Example choice sets in the GBR study

4. Results

In the case studies, error component and random parameter logit models were used to identify the influences of attributes and labels on choices, and to capture respondent heterogeneity. The summaries of WTP estimates by attributes and split-samples are shown for the two case studies in Tables 3 and 4. The results in the Namoi case study indicate that the first hypothesis of significant internal scale effects is accepted for nearly all attributes (the cost attribute was significant in every experiment). The ‘healthy waterways’ attribute was not significant in one split-sample, and the ‘people in agriculture’ attribute was not significant in three split-samples.

Table 3 WTP estimates and confidence intervals (CI) for the Namoi catchment study

Sample	Native vegetation	Native species	Healthy waterways	People in agriculture
100%	*1800 sq km*	*2130 species*	*2000 km*	*5800 persons*
Namoi	**$0.02**** ($0.00 - $0.03)	**$2.90***** ($0.80 - $4.90)	**$0.12***** ($0.06 - $0.17)	**$0.16** ($0.07 - $0.38)
Sydney	**$0.02***** ($0.01 - $0.03)	**$3.01***** ($1.26 - $4.76)	**$0.02** (-$0.04 - $0.07)	**$0.24**** ($0.01 - $0.46)
50%	*900 sq km*	*1065 species*	*1000 km*	*2900 persons*
Sydney	**$0.06***** ($0.04 - $0.09)	**$6.02**** ($1.15 - $11.02)	**$0.28***** ($0.15 - $0.42)	**$0.11** (-$0.40 - $0.62)
10%	*180 sq km*	*213 species*	*200 km*	*580 persons*
Namoi	**$0.15***** ($0.06 - $0.25)	**$15.53*** (-$2.76 - $32.79)	**$1.02***** ($0.55 - $1.47)	**$2.73***** ($0.93 - $4.57)
Sydney	**$0.24**** ($0.00 - $0.47)	**$32.95*** ($0.13 - $66.52)	**$0.84*** (-$0.11 - $1.82)	**$0.84** (-$2.64 - $4.31)

*** = significant at the 1% level, ** = significant at the 5% level; * = significant at the 10% level

The results in the GBR case study indicate that the first hypothesis is accepted for nearly all attributes (the cost attribute was significant in every experiment), apart from 'Seagrass' in two of the local population split-samples.

Table 4 WTP estimates and confidence intervals (CI) for the GBR study

Sample	REEF	FISH	SEAGRASS
WHOLE: 346,000 sq km	*20,000 sq km*	*1,500 species*	*44,000 sq km*
Brisbane (BNE)	**$0.06***** ($0.05-$0.07)	**$0.53***** ($0.37-$0.70)	**$0.01***** ($0.01-$0.02)
Townsville (TSV)	**$0.08***** ($0.05-$0.12)	**$0.90***** ($0.79-$1.01)	**$0.02**** ($0.00-$0.03)
REGION: 86,500 sq km	*5,000 sq km*	*1,500 species*	*11,000 sq km*
Brisbane sample			
BNE: north region	**$0.24***** ($0.17-$0.33)	**$0.95***** ($0.75-$1.14)	**$0.07***** ($0.05-$0.10)
BNE: central region	**$0.19***** ($0.15-$0.23v	**$0.60***** ($0.47-$0.75)	**$0.04***** ($0.02-$0.06)
BNE: south region	**$0.16***** ($0.11-$0.23)	**$0.66***** ($0.49-$0.84)	**$0.05***** ($0.02-$0.08)
Local sample			
Cairns (CNS): north region	**$0.21***** ($0.10-$0.33)	**$0.50***** ($0.16-$0.84)	**$0.01** ($0.05$-0.07)
TSV: central region	**$0.13**** ($0.01-$0.24)	**$0.64*****($0.20-$1.15)	**$0.05** ($0.00-$0.10)
Capricorn coast (CAP): south region	**$0.45***** ($0.26-$0.75)	**$1.39***** ($0.84-$2.16)	**$0.11***** ($0.04-$0.21)
LOCAL case studies			
Cairns: 1,515sq km	*376 sq km*	*100 species*	*25 sq km*
Townsville: 9,700sq km	*260 sq km*	*100 species*	*56 sq km*
Capricorn coast: 2,425 sq km	*27 sq km*	*100 species*	*7 sq km*
Brisbane sample			
BNE: Cairns area	**$1.51***** ($1.0-$2.0)	**$8.22***** ($5-$12)	**$28.66***** ($20-$39)
BNE: Townsville area	**$4.27***** ($1.0-$7.0)	**$15.79***** ($10-$22)	**$9.44***** ($3-$16)
BNE: Capricorn Coast area	**$51.20***** ($30-$72)	**$16.18***** ($10-$23)	**$74.78***** ($51-$101)
Local sample			
CNS: Cairns area	**$3.34***** ($1.0-$6.0)	**$16.12***** ($7-$26)	**$37.41***** ($18-$64)
TSV: Townsville area	**$4.63***** ($3.0-$7.0)	**$12.15***** ($7-$17)	**$12.29***** ($8-$18)
CAP: Capricorn Coast area	**$26.59*** ($0-$54)	**$12.86***** ($5-$21)	**$69.61***** ($40-$104)

*** = significant at the 1% level, ** = significant at the 5% level; * = significant at the 10% level

The scale and scope effects relating to the second and third hypotheses are largely confounded in the tests as both scale and scope increase together for most attributes in the two case studies (Appendix 1). In all cases in the Namoi catchment study, the WTP for a marginal improvement decreases as the amounts of the amenity increases (Table 3), supporting both scale and scope hypotheses. The Poe *et al.* (2005) procedure is used to estimate whether the proportion of differences greater than zero are significant, with comparisons only made where attribute significance permits (Table 5). The results indicate there is only one case where there is no significant difference between the two estimates (native species comparison for Sydney between the 10% and 50% levels), supporting both Hypothesis 2 and Hypothesis 3.

Table 5 Catchment study tests of difference between WTP estimates

Sample	**Comparison**	**Vegetation** p-value	**Native species** p-value	**Waterways** p-value	**People in ag** p-value
Namoi	WTP_{10}– WTP_{100}	0.003***	0.077*	0.000***	0.004***
	WTP_{10}– WTP_{100}	0.037**	0.040**	0.046**	-
Sydney	WTP_{50}– WTP_{100}	0.072*	0.060*	-	-
	WTP_{10}– WTP_{50}	0.002***	0.125	0.000***	-

*** = significant at the 1% level, ** = significant at the 5% level; * = significant at the 10% level

In the GBR study, marginal WTP values held by the Brisbane sample declined for all attributes as the amount of the GBR under study increased (Table 4). The same holds true for the local population sample apart from the fish attribute, where values for the north and central regions are lower than at the whole level of scope. The Poe *et al.* (2005) tests (Table 6) verify the general significance of declines in value with scope increases, supporting both Hypothesis 2 and Hypothesis 3.

Table 6 GBR study tests of difference between WTP estimates

	Brisbane sample			**Local sample**		
Comparison	**REEF**	**FISH**	**SEAGRASS**	**REEF**	**FISH**	**SEAGRASS**
whole - north region	1.000***	0.999***	1.000***	0.991***	0.077*	-
whole - central region	1.000***	0.759	0.999***	0.776	0.211	-
whole - south region	1.000***	0.861	0.996***	1.000***	0.899	0.994***
whole - north local	0.999***	1.000***	1.000***	0.997***	0.999***	1.000***
whole - central local	1.000***	1.000***	0.996***	1.000***	1.000***	1.000***
whole - south local	1.000***	1.000***	1.000***	0.970**	0.999***	1.000***
north region - north local	0.999***	1.000***	1.000***	0.995***	0.999***	-
central region- central local	1.000***	1.000***	0.996***	1.000***	1.000***	-
south region- south local	1.000***	1.000***	1.000***	0.969**	0.998***	1.000***

The 'fish' attribute in the GBR study provides a test of differences between scale and scope effects, as the levels were the same (1500 species) for both the whole and regional scopes. In four out of the six comparative examples across the two population samples, there was no significant difference in the WTP estimates for the whole and regional studies, and there was only weak evidence of a difference (at the 10% level) in a fifth case. These results suggest that there is only limited evidence for hypothesis 3 between regional and whole effects, although scope effects may be more significant between whole and local case studies.

5. Calibration factors for 'scope' effects

The results of the 'scope' experiments reported in the section above indicate that some form of calibration would be needed to transfer values between studies with different geographic scopes. The extent of the calibration that might be required, following the ratio of WTP values approach of van Bueren and Bennett (2004), is shown in Tables 7 and 8. The results of both case studies demonstrate that if the marginal WTP values at the small/local level were simply extrapolated to the largest case study without adjustment, total values would be overestimated (by between 10 and 7,500 times). In most cases the local populations have lower adjustment factors (and hence smaller 'scope' differences) than more distant populations.

The variation in the ratio of WTP values between case studies indicates that it will not be appropriate to use generic categorical calibration factors, such as from local to regional or from regional to national frames. Instead, calibration factors for benefit transfer will need to be tailored in some way to the characteristics of the different source and target studies. The results of the two case studies provide some guide to a suitable approach. A comparison of the ratios of WTP values to the ratios of the geographic scope quantities involved in each experiment is also available from Tables 7 and 8. This demonstrates that differences in WTP ratios across case studies are broadly related to the proportional change in the quantities involved.

Table 7 Ratio of WTP estimates relative to the 100% level in catchment study

	Ratio of quantities ***(100% scope / smaller)***	**Native vegetation** (WTP smaller/WTP 100% catchment)	**Native species** (WTP smaller/WTP 100% catchment)	**Healthy waterways** (WTP smaller/WTP 100% catchment)	**People in agriculture** (WTP smaller/WTP 100% catchment)
Distant Sydney sample					
10%	*10*	12.0	10.9	n.s.	n.s.
50%	*2*	3.0	2.0	n.s.	n.s.
Local Namoi sample					
10%	*10*	7.5	5.4	8.5	n.s.

Note: n.s. means at least one source WTP is not significant

The correlation coefficient between the ratios of the quantities involved and the WTP estimates for each 'scope' test across the two studies is estimated at 0.695. The relationship is stronger when both the quantity and the WTP ratios are expressed in (natural) log form, with the correlation coefficient rising to 0.986. The extent of correlation between the two ratios for each scope test is shown below in Figure 3, with the data pooled across the two studies and in ascending order by the ratios of quantity changes.

Tests of possible relationships between the ratio of quantities involved and the ratio of WTP estimates for the different scope tests identify that improved model fits are obtained with log-log models rather than with other forms. Regression models (Table 9) show that the relationship between quantity and value ratios is close to unity, that regional populations tend to have lower adjustments for geographic scope than urban populations, and that no difference in results could be distinguished between the Namoi catchment studies and GBR studies.

Table 8 Ratio of WTP estimates relative to the whole level in GBR study

	Reef		Fish		Seagrass	
	Ratio of quantities ***(whole GBR / smaller)***	Ratio of WTP (WTP smaller / WTP whole)	***Ratio of quantities*** ***(whole GBR / smaller)***	Ratio of WTP (WTP smaller / WTP whole)	***Ratio of quantities*** ***(whole GBR / smaller)***	Ratio of WTP (WTP smaller / WTP whole)
Distant Brisbane sample						
Region						
North region	*4*	4.0	*1*	1.8	*4*	7.0
Central region	*4*	3.2	*1*	1.1	*4*	4.0
South region	*4*	2.7	*1*	1.2	*4*	5.0
Local						
Cairns local	*56*	25.2	*16*	15.5	*1,652*	2,866.0
Townsville local	*109*	71.2	*16*	29.8	*760*	944.0
Cap Coast local	*739*	853.3	*16*	30.5	*6,333*	7,478.0
Local sample						
Region						
North region	*4*	2.6	*1*	0.6	*4*	n.s.
Central region	*4*	1.6	*1*	0.7	*4*	n.s.
South region	*4*	5.6	*1*	1.5	*4*	5.5
Local						
Cairns local	*56*	41.8	*16*	17.9	*1,652*	1,870.5
Townsville local	*109*	57.9	*16*	13.5	*760*	614.5
Cap Coast local	*739*	332.4	*16*	14.3	*6,333*	3,480.5

Note: n.s. means at least one source WTP is not significant

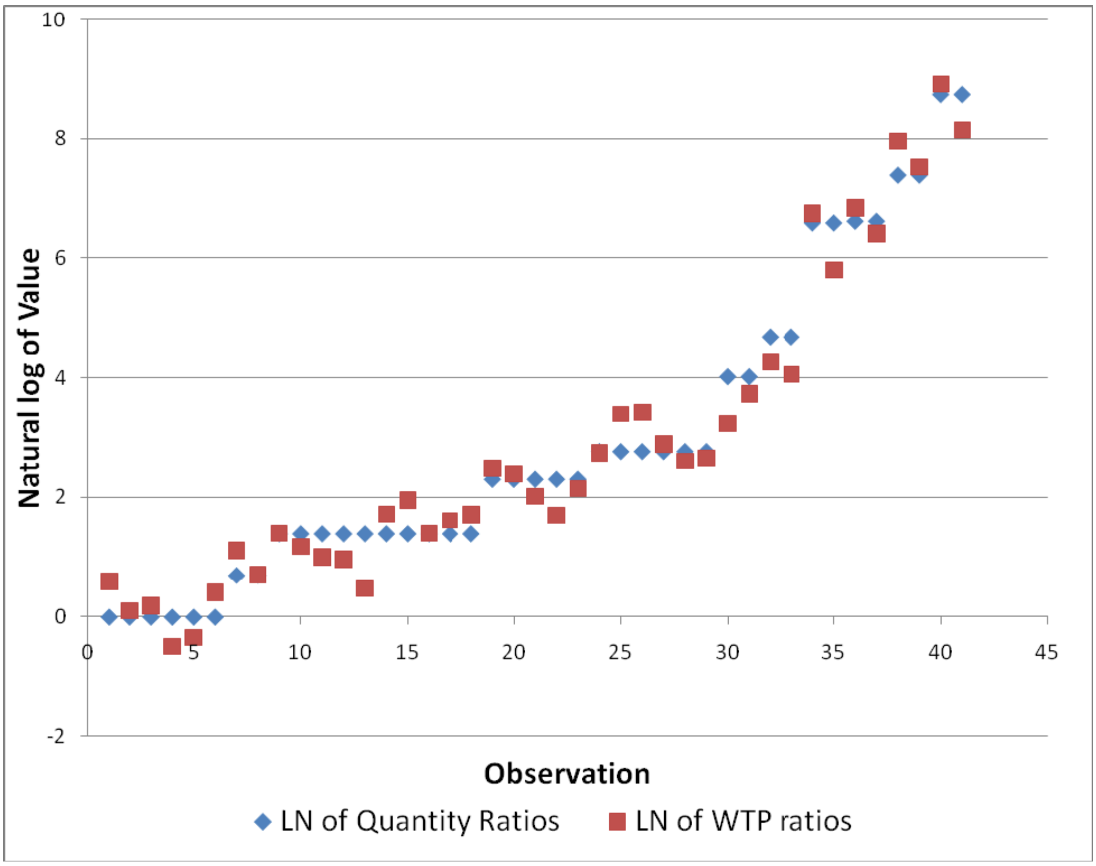


Figure 3 Comparison of quantity and value ratios (in natural logs) across geographic scopes from both case studies

Table 9 Regression model predicting LN of ratio of WTP across scope tests

	Pooled tests across Namoi catchment and GBR		**Pooled tests with additional comparisons from van Bueren & Bennett**	
	Coefficient	*St. Error*	*Coefficient*	*St. Error*
Constant	0.1894	0.1129	0.0345	0.8072
Natural Log of ratio of quantities in scope test	0.9807***	0.0243	1.0191***	0.0000
Dummy for regional population	-0.3744***	0.1199	-0.3014***	0.0192
Dummy for Namoi catchment	-0.0494	0.1616	0.0035	0.9871
Dummy for van Bueren & Bennett 2000 data			0.0927	0.6733
Model statistics				
No of Observations	41		67	
Log L	-16.60		-47.54	
Restricted Log L	-95.53		-151.81	
Finite sample: AIC	-1.83		-1.27	
R-sqrd	0.979		0.956	

*** significant at the 1% level;

To explore these relationships further, retrospective tests have also been conducted on the geographic scope studies for Australia and two regional areas reported by van Bueren and Bennett (2000, 2004). The attribute levels and WTP values were sourced from van Buren and Bennett (2000), and the ratios of quantity and values were calculated for both the available national-regional comparisons, and for the region-region comparison. Several of the comparisons were replicated across outside-urban and within-region populations, so that 26 valid observations were available after observations with missing values were dropped. The correlation coefficient between the ratios of the quantities involved and the WTP estimates is 0.894, increasing to 0.940 when the ratios are expressed in (natural) log form. The latter relationship is shown in Figure 4.

A regression analysis of these additional observations pooled with the Namoi catchment/GBR data set is shown in Table 9. The insignificant dummy variable for the added data indicates that results are consistent across the studies, while the strength of the regression confirms a systematic relationship between the geographic scope of experiments and the attribute WTP that is generated.

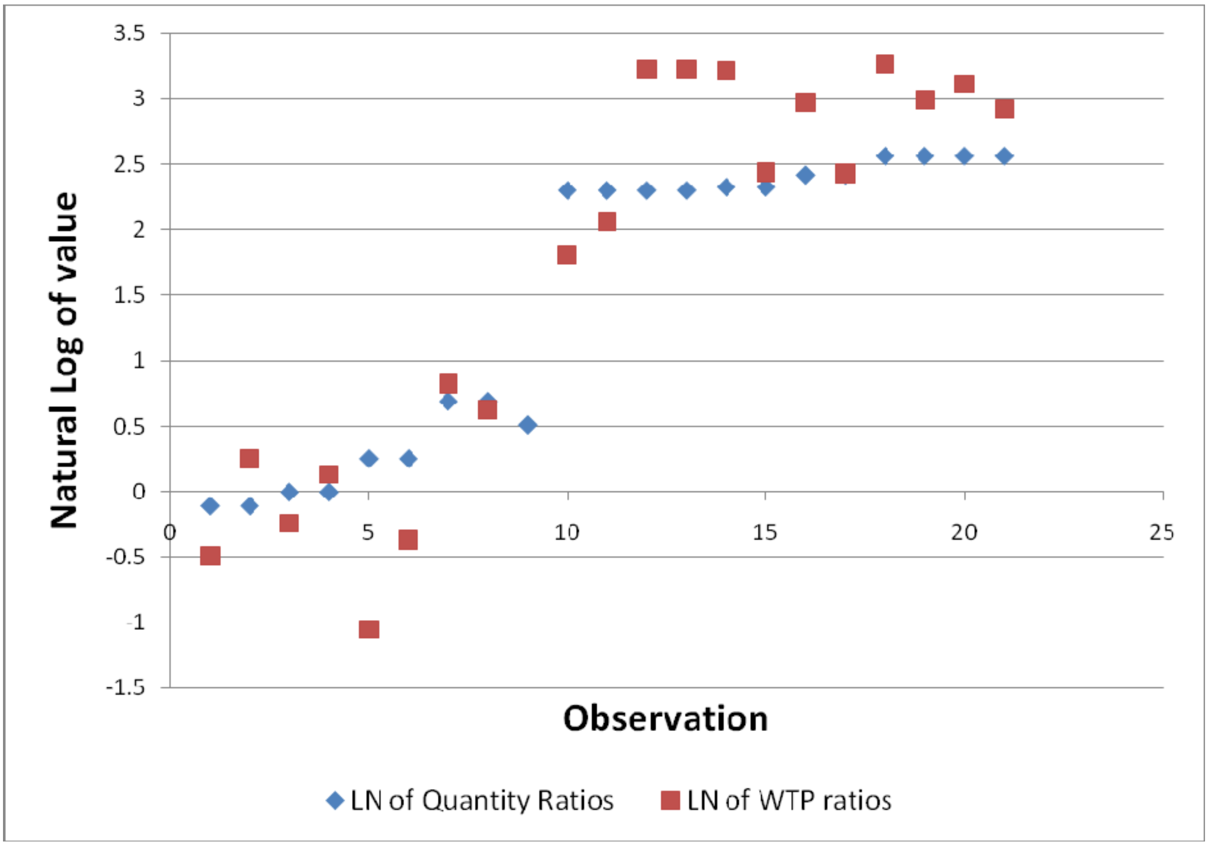


Figure 4 Comparison of quantity and value ratios (in natural logs) across geographic scopes from van Bueren and Bennett (2000) data

6. Discussion and conclusions

The review of two natural resource management focused choice modelling studies in this paper provides key insights into the impacts of scale and scope factors on WTP estimates, and the potential use of adjustment or calibration factors in benefit transfer applications when the geographic amount of an amenity varies between source and target sites. A key distinction underpinning the analysis is that scale effects are defined in terms of the amount of the amenity that is considered, and scope effects are defined as the context in which an amenity is being valued. Where geographic extent is being considered, then scale and scope effects are usually intertwined, as increases in the context of a good's provision (scope) usually increases the amount involved (scale) as well.

The results of the studies confirm that attribute significance in choice experiments is a strong test of internal scale effects, and also show that 'average' marginal WTP values decline significantly as the geographic scope increases between studies. These declines in average WTP values may be both a consequence of diminishing marginal utility as increasing amounts of the amenity are involved (scale), as well as other framing and substitute effects (scope). The results demonstrate that average WTP values can be up to 7,500 times higher in small local case studies compared to large broadly scoped studies.

The significant variations in marginal WTP across the case studies confirm that calibration factors are likely to be required in benefit transfer where geographic 'scope' changes are involved (Smith *et al.* 2002; van Bueren and Bennett 2004; Smith *et al.* 2006, Johnson and Rosenberger 2010). Where target and source studies have different geographic dimensions, then simple transfer and extrapolation of unit values may be very misleading. The large variation in WTP values between differently scoped case studies demonstrate that simple adjustment factors by scope categories will not be appropriate, and adjustment factors will need to be calibrated for each case study.

Across the 41 different scope tests that have been examined in the two case studies of interest, and the 26 additional scope tests sourced from van Bueren and Bennett (2000), there are very high levels of correlation between the ratios of the quantities involved in each scope test and the ratios of the corresponding WTP estimates. Expressing these ratios in log-log form gives almost perfect predictive power. We recommend that where benefit transfer applications involve differences in geographic scope, and primary studies are not available, that value transfers should be calibrated by the following relationship:

$$LN(WTP_{ATTsmall} / WTP_{ATTlarge}) = LN(Quantity_{ATTlarge} / Quantity_{ATTsmall})$$

where WTP refers to the average implicit price for different case studies, and the Quantity refers to the amount of the attribute involved across different levels of geographic scope.

We note that one implication of using this calibration factor is that total WTP will be insensitive to increases in geographic scope. The calibration factor implies that respondents identify an amount they are willing to pay to address a specific environmental issue and then distribute that amount according to the different scopes of the good presented. Hence, for smaller scopes, the total willingness to pay is split into larger unit values whereas for larger scopes, the same total is split into smaller unit values. This feature may be exacerbated by the use of the same range for the cost attribute in each of the differing scope split samples for the two case study applications. These are important issues that remain to be tested in future research.

Acknowledgement

This research has been supported through the Environmental Economics Research Hub, funded by the Australian Department of Environment and Water Heritage and the Arts under the Commonwealth Environment Research Facility

References

Arrow, K., Solow, R., Portney, P., Leamer, R., Radner, R. and Schuman, H. (1993). Report of the NOAA Panel on contingent valuation, *Federal Register* 58 (10), 4602–4614.

Bateman, I., Cole, M., Cooper, P., Georgiou, S., Hadley, D. and Poe, G. (2004). On visible choice sets and scope sensitivity, *Journal of Environmental Economics and Management* 47, 71–93.

Bateman, I., Day, B., Georgiou, S. and Lake, I. (2006). The aggregation of environmental benefit values: Welfare measures, distance decay and total WTP, *Ecological Economics* 60, 450-460.

Bergland, O., Magnussen, K. and Navrud, S. (2002). Benefit transfer: testing for accuracy and reliability, in Florax, R., Nijkamp, P. and Willis, K. (eds) *Comparative Environmental Economic Assessment*, Edward Elgar, Cheltenham, U.K.

Boyle, K. J. and J. C. Bergstrom, (1992). Benefit transfer studies: myths, pragmatism and idealism, *Water Resources Research* 28 (3), 657-63.
Brouwer, R. (2000). Environmental value transfer: state of the art and future prospects, *Ecological Economics* 32, 137-152.
Czajkowski, M. and Hanley, N. (2009). Using labels to investigate scope effects in stated preference methods, *Environmental and Resource Economics* 44(4), 521-535.
Hanemann, W.M. (1994). Valuing the environment through contingent valuation, *Journal of Economic Perspectives* 8, 19–43.
Hanley N., Wright, R. and Adamowicz, W. (1998). Using choice experiments to value the environment: Design issues, current experience and future prospects, *Environmental Resource Economics* 3-4, 413-428.
Hanley, N., Wright, R. and Alvarez-Farizo, B. (2006). Testing for benefit transfer over water quality benefits, in Rolfe, J. and Bennett, J. (eds) *Choice Modelling and the Transfer of Environmental Values*, Edward Elgar, Cheltenham, UK.
Hanley, N., Schlapfer, F. and Spurgeon, J. (2003). Aggregating the benefits of environmental improvements: distance decay functions for use and non-use values, *Journal of Environmental Management* 68, 297-304.
Heberlein, T.A., Wilson, M.A., Bishop, R.C. and Schaeffer, N.C. (2005). Rethinking the scope test as a criterion for validity in contingent valuation, *Journal of Environmental Economics and Management* 50, 1–22.
Hoehn, J. P. (1991). Valuing the multidimensional impacts of environmental policy: Theory and methods, *American Journal of Agricultural Economics* 73, 289-299.
Hoehn, J. P. and Randall, A. (1989). Too many proposals pass the benefit cost test, *American Economic Review* 79, 544-551.
Johnston, R.J. and Rosenberger, R. (2010). Methods, trends and controversy in contemporary benefit transfer, *Journal of Economic Surveys* 24(3), 479-510.
Loomis, J. B. (1992). The evolution of a more rigorous approach to benefit transfer: benefit function transfer, *Water Resources Research* 28 (3), 701-5.
Loomis, J. (2006). Stated preference benefit transfer approaches for estimating passive use value of wild salmon, in Rolfe, J. and Bennett, J. (eds) *Choice Modelling and the Transfer of Environmental Values*, Edward Elgar, Cheltenham, UK.
Mazur, K. (2011). *Choice Modelling in the Development of Natural Resource Management Strategies in NSW Catchments*, Unpublished PhD thesis, Australian National University, Canberra.
Mazur, K. and Bennett, J. (2009). *Scale and Scope Effects on Communities' Values for Environmental Improvements in the Namoi Catchment: A choice modelling approach*, Environmental Economics Research Hub Research Report No.42. http://www.crawford.anu.edu.au/research_units/eerh/publications/.
Mitchell, R.C. and Carson, R.T. (1989). Using *Surveys to Value Public Goods: the Contingent Valuation Method*, Resources for the Future, Washington, DC.
Pate, J. and Loomis, J. (1997). The effect of distance on willingness to pay values: a case study of wetlands and salmon in California, *Ecological Economics* 20, 199-207.
Poe, G.L., Giraud, K.L. and Loomis, J.B. (2005). Computational methods for measuring the differences of empirical distributions, *American Journal of Agricultural Economics* 87(2), 353-365.
Portney, P.R. (1994). The contingent valuation debate: why economists should care, *Journal of Economic Perspectives* 8 (4), 3–17.
Randall, A. and Hoehn, J.P. (1996). Embedding in market demand systems, *Journal of Environmental Economics and Management* 30, 369-80.
Rolfe, J. (2006). Theoretical issues in using choice modelling data for benefit transfer, in

Rolfe, J. and Bennett, J. (eds) *Choice Modelling and the Transfer of Environmental Values*, Edward Elgar, Cheltenham, UK.
Rolfe, J. and Bennett, J. (eds) (2006). *Choice Modelling and the Transfer of Environmental Values*. Edward Elgar, Cheltenham, UK.
Rolfe, J. and Brouwer, R. (2011). *Testing for Value Stability with a Meta-analysis of Choice Experiments: River Health in Australia*, Environmental Economics Research Hub Research Report No.95. http://www.crawford.anu.edu.au/research_units/eerh/publications/.
Rolfe, J. and Wang, X. (2011). Dealing with scale and scope issues in stated preference experiments, in Bennett, J. (ed) *The International Handbook on Non-market Environmental Valuation*, Edward Elgar, Cheltenham, UK (in press).
Rolfe, J. and Windle, J. (2008). Testing for differences in benefit transfer values between state and regional frameworks, *Australian Journal of Agricultural and Resource* Economics 52(2), 149-168.
Rolfe, J. and Windle, J. (2010a). *Valuing Environmental Improvements in the Great Barrier Reef: Ecological and Preference Heterogeneity in Local Area Case Studies*, Environmental Economics Research Hub Research Report No.66. http://www.crawford.anu.edu.au/research_units/eerh/publications/
Rolfe, J. and Windle, J. (2010b). *Testing for Geographic Scope and Scale Effects with Choice Modelling: Application to the Great Barrier Reef*, Environmental Economics Research Hub Research Report No.69. http://www.crawford.anu.edu.au/research_units/eerh/publications/
Rolfe, J., Bennett, J. and Louviere, J. (2002). Stated values and reminders of substitute goods: Testing for framing effects with choice modelling, *Australian Journal of Agricultural and Resource Economics* 46(1), 1-20.
Rolfe, J., Loch, A. and Bennett, J. (2006). Testing benefit transfer with water resources in Central Queensland, Australia, in Rolfe, J. and Bennett, J. (eds) *Choice Modelling and the Transfer of Environmental Values*. Edward Elgar, Cheltenham, UK.
Schulze, W. D., McClelland, G., Lazo, J. and. Rowe, R. D. (1998). Embedding and calibration in measuring non-use values, *Resource and Energy Economics* 20, 163–178.
Smith, V.K., van Houtven, G. and Pattanayak, S.K. (2002). Benefit transfer via preference calibration: 'prudential algebra' for policy, *Land Economics* 78, 132–152.
Smith, V.K., Pattanayak, S.K. and van Houtven, G. (2006). Structural benefit transfer: an example using VSL estimates, *Ecological Economics* 60, 361–371.
van Bueren, M. and Bennett, J. (2000). *Estimating Community Values for Land and Water Degradation Impacts*, project report for the National Land and Water Resources Audit, Canberra.
van Bueren, M. and Bennett, J. (2004). Towards the development of a transferable set of value estimates for environmental attributes, *Australian Journal of Agricultural and Resource Economics* 48 (1), 1-32.

Appendix 1 Attribute levels by experiment scope

Table A1 Attribute levels for the Namoi catchment study

Sample	**Cost** ($)	**Native vegetation** (sq km)	**Native species** (species)	**Healthy waterways** (km)	**People in agriculture** (persons)
100%					
Current level		1800	2130	2000	5800
Status Quo	0	1800	2100	1900	5000
Alternatives	50, 200, 300	3000, 5,000, 6000	2110, 2120, 2130	2300, 2700, 3000	5100, 5200, 5300
50%					
Current level		900	1065	1000	2900
Status Quo	0	900	1050	950	2500
Alternatives	50, 200, 300	1500, 2500, 3000,	1050, 1060, 1065	1150, 1350, 1500	2550, 2600, 2650
10%					
Current level		180	213	200	580
Status Quo	0	180	210	190	500
Alternatives	50, 200, 300	300, 500, 600	211, 212, 213	230, 270, 300	510, 520, 530

Table A2 Attribute levels for the GBR study

Sample	**Reef** (sq km)	**Fish** (species)	**Seagrass** (sq km)	**Cost** ($)
WHOLE				
Total Area: 346,000 sq km	20,000	1,500	44,000	
Current level	18,000 (90%)	1,350 (90%)	40,000 (90%)	
Status Quo	13,000 (65%)	975 (65%)	28,000 (65%)	0
Alternatives	14,000, 16,000, 17,000 (70%, 80%, 85%)	1,050, 1,200, 1,275 (70%, 80%, 85%)	31,000, 35,000, 38,000 (70%, 80%, 85%)	50, 100. 200, 500
REGION				
Total Area: 86,500 sq km	5,000	1,500	11,000	
Current level	4,500 (90%)	1350 (90%)	10,000 (90%)	
Status Quo	3,250	975 (65%)	7,000	0
Alternatives	3,500, 4,000, 4,250 (70%, 80%, 85%)	1,050, 1,200, 1,275 (70%, 80%, 85%)	7,750, 8,750, 9,500 (70%, 80%, 85%)	50, 100. 200, 500
LOCAL case studies				
Cairns				
Total Area: 1,515sq km	376	100	25	0
Current level	282 (75%)	75 (75%)	23 (90%)	
Status Quo	207 (55%)	55 (55%)	16 (65%)	0
Alternatives	244, 282, 301 (65%, 75% ,80%)	65, 75, 80 (65%, 75%, 80%)	19, 21, 23 (75%, 85%,90%)	50, 200, 500
Townsville				
Total Area: 9,700sq km	260	100	56	
Current level	117 (45%)	75 (75%)	42 (75%)	
Status Quo	91 (35%)	55 (55%)	36 (65%)	0
Alternatives	117, 143, 156 (45%, 55%, 60%)	65, 75, 80 (65%, 75%, 80%)	42, 48, 50 (75%, 85%, 90%)	50, 200, 500
Capricorn coast				
Total Area: 2,425 sq km	27	100	7	
Current level	23 (85%)	75 (75%)	5 (70%)	
Status Quo	16 (60%)	55 (55%)	3 (50%)	0
Alternatives	19, 22, 23 (70%, 80%, 85%)	65, 75, 80 (65%, 75%, 80%)	4, 5, 6 (60%.70%,75%)	50, 200, 500