The 84th Annual Conference of the Agricultural Economics Society

Edinburgh

29th to 31st March 2010

Reducing Bias from Choice Experiments Estimates in the Demand for Recreation

Alberto Longo*
Emma Rowan
W. George Hutchinson

Gibson Institute for Food, Land and the Environment
School of Biological Sciences
Queen's University Belfast
Belfast, UK

Abstract: In valuing the demand for recreation, the literature has grown from using revealed preference methods to applying stated preference methods, namely contingent valuation and choice modelling. Recent attempts have merged revealed and stated preference data to exploit the strengths of both sources of data. We use contingent behaviour and choice experiments data to show that, with choice experiments exercises, when respondents are asked to choose which improvement programme they prefer for a site with recreational opportunities, failing to consider the information explaining the number of visits that respondents intend to take to a recreational site under each hypothetical programme leads to biased coefficients estimates in the models for the choice experiments data.

Keywords: travel cost, contingent behaviour, choice experiments, revealed preferences, stated preferences

JEL: Q51, Q26

1. Introduction and previous studies

In valuing the demand for recreation, the literature has grown from using revealed preference (RP) methods (Clawson & Knetsch, 1966; Bhat, 2003; Bhat & Gossen, 2004) that are only able to assess use values of a recreational site to applying stated preference (SP) methods, namely contingent valuation (Bateman et al, 1994) and choice modelling (Louviere and Timmermans, 1990) that, in addition to use values, are able to capture non-use values. Both approaches have been criticized for their respective limitations: RP methods can only assess use values; while SP methods are based on hypothetical scenarios. To overcome the shortcomings of each method, recent attempts have merged RP and SP data to exploit the strengths of both sources of data (Adamowicz et al, 1994, 1997, 1998; Alberini and Longo 2006; Alberini et al 2007; Cameron, 1992; Cameron et al, 1996; Christie et al 2007; Englin and Cameron 1996; Eom and Larson, 2006; Hanley et al 2003; Huang et al 1997; Whitehead et al 2000, 2005, 2007, 2008 and Bhat 2003). Among recent applications, Huang et al (1997) combine RP and SP to estimate WTP for quality improvements at recreational sites in North Carolina by using travel cost, contingent behaviour (CB) and contingent valuation questions. Bhat (2003) combines travel cost and CB to value the non-market recreational benefits of reef quality improvements. Eom and Larson (2006) apply the information obtained from a study incorporating elements of both RP and SP techniques by combining data from a travel cost model and a contingent valuation survey. Hanley et al (2003) value the benefits of coastal water quality improvement by applying travel cost and CB data to predict the welfare change arising from trip numbers as water quality improves. Christie et al (2007) study forest recreation in Great Britain using choice experiments (CE) and CB. They use both CE and CB to compare the benefit estimates from the data. Data were collected through in person interviews between May and September 2005 in seven forests throughout Great Britain chosen for their ability to cover the range of activities being investigated. Respondents were presented with information on how the forest they were in could be improved before being shown a single CB scenario. From this, they were asked to outline whether their intended trips to this forest would change within the next 12 months should these improvements be implemented. They were then given a series of four CE choice tasks, where they were asked to state their preferred scenario and the number of expected trips under the chosen scenario.

In this paper, we carry forward the approach by Christie et al (2007), as we collect information on travel cost, CB and CE related to a recreational site. Compared to Christie et al (2007), we collect the information on the expected number of trips under each hypothetical scenario used in the CE questions, and exploit this information in the CE data estimation. To avoid problems of endogeneity in our econometric models, we use an approach similar to the one employed by Guevara and Ben-Akiva (2006) and Petrin and Train (2003), who use a 'control function' to help explain unobservable variables in their CE models with a two step estimation approach.

2. The economic model

In this paper we expand the strength of merging different sources of data by using data from CB and CE questions to improve the valuation of the demand for recreation to the Mournes Mountains in Northern Ireland, UK. We claim that exploiting the information researchers can obtain from the demand for recreation may help improving the estimates from CE data.

We start from a single site travel cost survey, where respondents provide us with welfare estimates for the access value to the recreational area. This gives us the use value that respondents experience from the recreational area. We then ask respondents a set of CB questions on how many trips they intend to take, in a twelve month period, under the current situation and under four other hypothetical programmes of improvement to the recreational site. The four hypothetical programmes are described by different levels of four attributes: access to the site, infrastructure at the site, management of the area, and planning controls. After the CB questions, respondents answer two CE questions, where each hypothetical programme is described by the same attributes and levels used in the CB questions, in addition to an annual tax attribute. The scenarios used in the first (second) CE question are the same ones presented to respondents in the first (third) and second (fourth) CB question. In this way, we have all the information to control for the use value part of utility that respondents receive under each scenario, including the current situation.

We first run a Random Effects Poisson Model from the CB data to explain the number of trips that respondents take under each scenario:

(1) $r_{ij} = \beta_0 + \beta_1 TC + \beta_2 A_j + \beta_3 I_j + \beta_4 M_j + \beta_5 P_j + X\beta_6 + \eta$

Where r is the number of trips taken under scenario j by respondent i, TC is the travel cost to the recreational area, A is the level of access, I is the level of infrastructure, M is the level of management, P is the level of planning controls, X is a vector of respondent's characteristics, βs are parameters to be estimated and η is the error term. There are five scenarios j for each respondent, with one scenario being the current situation and the other four present hypothetical changes to the current situation. Different respondents received different hypothetical changes to the current situation. We then calculate the residuals $\hat{r}_{ij} - r_{ij}$, that capture the part of the utility not explained by the variables included in equation (1), that we will later use to help better explain the CE data.

Next, the CE data are analysed within the Random Utility Model (RUM) framework that assumes that a respondent chooses the option that gives him/her the highest level of utility, given the attributes and the levels of each option within a choice set:

(2) $V_{ij} = \alpha_0 + \alpha_1 A_j + \alpha_2 I_j + \alpha_3 M_j + \alpha_4 P_j + \alpha_5 C_j + \varepsilon_j$

Where *V* is the indirect utility from scenario *j* for respondent *i*, and *C* is the tax for implementing scenario *j*.

In this paper we want to investigate whether model (2) suffers from an omitted variable bias: when respondents answer the CE question, to what extent the information on the number of trips that they intend to take to the site under each hypothetical scenario, including the current situation, affects their choices? The variables used in model (2) already explain both use and non-use value components of the utility that respondents receive under each hypothetical scenario. The attributes of the CE questions also already include the information on the number of trips that respondents would take under each hypothetical scenario, albeit in an implicit way. If we augmented model (2) with an additional variable reporting the number of trips obtained from the CB data that respondents would do under each hypothetical scenario we would cause a problem of endogeneity in model (2), as the number of trips are determined by the attributes in model (1). Following Guevara and Ben-Akiva (2006) and Petrin and Train (2003), we the use a 'control function' approach by using the residuals calculated from (1) and add them as an additional variable to model (2):

(3) $V_{ij} = \alpha_0 + \alpha_1 A_j + \alpha_2 I_j + \alpha_3 M_j + \alpha_4 P_j + \alpha_5 C_j + \alpha_6 residuals_j + \varepsilon_j$

The residuals bring the information from the use value model not explained by the attributes nor variables used in (1) or (2). We then use Mixed Logit Models to estimate (2) and (3). Specifically, we use random parameters models, to relax the IIA property of the multinomial logit model, and assume that the coefficients of the variables in models (2) and (3) are not assumed to be fixed across individuals, but are composed by a fixed part, α_j common to all individuals, and by a part that varies across individuals, u_{ij}, where u is a random component, that we assume to be distributed following a triangular distribution with mean and spread constrained to be equal. Using a t-test on α_6, we can investigate the importance of the use value part of utility not explained by the variables in (1) and (2), hence helping to explain the choices made by individuals in the CE exercises. A not significant coefficient estimate for α_6 would recommend that model (2) is the correct one, and that no bias is present in the coefficients estimated in equation (2). A significant coefficient estimate for α_6 would suggest that equation (2) estimates are biased, that respondents do consider the number of trips they expect to take when choosing their preferred option in the CE questions, and that model (2) suffers from an omitted variable bias. In addition, respondents' utility would be explained by use value components not well captured by the variables used in (2). A positive and significant estimate for α_6 would indicate that unobservable use value aspects not captured by the attributes nor the variables used in (1) are important determinants in the choice of hypothetical programs. A negative and significant coefficient estimate for α_6 would stress that unobservable use value aspects actually decrease the utility of respondents in the choice experiments questions.

3. The case study and the survey instrument

Northern Ireland is the only administrative division within the United Kingdom that does not have a National Park. The idea for a National Park in Northern Ireland was first raised by the Planning Advisory Board in their 1946 report "The Ulster Countryside" (Northern Ireland Planning Advisory Board 1947). It identified the Mourne Mountains in particular and requested its immediate designation. The Mournes area is one of the most striking mountain districts in Ireland. It comprises twelve peaks each rising above 600m (1968.5 feet). Much of the area is included within the Mourne Area of Outstanding Natural Beauty (AONB) in recognition of the quality of its landscape. The area boasts the first National Nature Reserve to be designated in Ireland and has an abundance of pure water reserves within its 9,000-acre catchment area demarcated by the 22-mile long Mourne Wall, which supplies the local Mournes area and much of Belfast (Kirk 2002).

In September 2002, the Minister for the Environment expressed a commitment to progress towards a Mourne National Park provided there was sufficient public support for such a designation. A study in 2002 identified the Mournes area as being the place most suited to a National Park designation and becoming Northern Ireland's first National Park (Europarc 2002). The Mourne National Park Working Party (MNPWP) was established in 2004 by the government as an independent body whose role was to commission research on a National Park boundary and to investigate the prospect of National Park designation for the Mournes area (EHS NI 2004). The MNPWP undertook an extensive public consultation exercise within the Mournes area which ran from August 2006 through until January 2007 (Inform Communications 2007) with their final report highlighting that residents of Northern Ireland deem important the following attributes for a Mournes National Park: (i) access to the area, (ii) infrastructures at the area, (iii) planning restrictions a the area, (iv) the type of management of the site. Following the results from the extensive public consultation carried out by the MNPWP, for this survey we use these four key attributes as attributes for the CE and for describing the hypothetical scenarios in the CB questions. We set these attributes at two different levels – high and low (being the status quo low too). The attributes therefore are *infrastructure* (toilets, parking facilities, rest spots availability, visitor centre and information provision), *access* (onto public and private lands), *planning* restrictions (controls for design of buildings and materials used), the type of *management* for the area as well as a *cost* attribute. These attributes are represented by symbols and have been clearly set out in the survey instrument. For the CE part of the survey, a *cost* attribute is presented to each respondent in each of the options given in the choice sets. The payment vehicle devised is an annual environmental tax to be collected from all wages, salaries, pensions and social security payments. The maximum and minimum amounts people were willing to pay for the alternatives presented to them were determined through the focus groups, and set at 5 levels (£2, £6, £10, £15 and £25). We create the experimental design following Johnson et al (2007) using SAS 9.1. Figures 1 and 2 present the scenarios used for the first CB question and for a CE question in one of the questionnaires. The reader can notice that the first alternative in the CE question is described with the same levels used for the CB question.

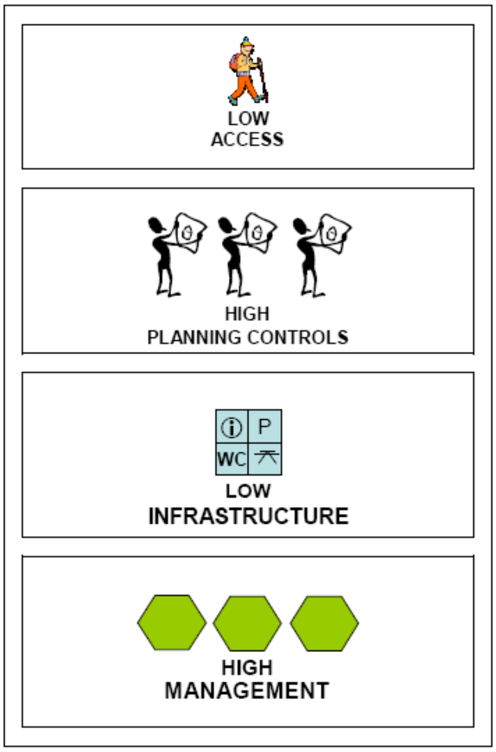
The survey instrument was administered by mail after the summer of 2008 to a sample of the population of Northern Ireland.

Figure 1. Example of CB question

PART 3: Consider your next trips to the Mournes area

The following questions ask you to choose how many times you would visit the Mournes area within the **next 12 months** period if these particular programmes were implemented.

1. Would you visit the Mournes area if the following were to be implemented?

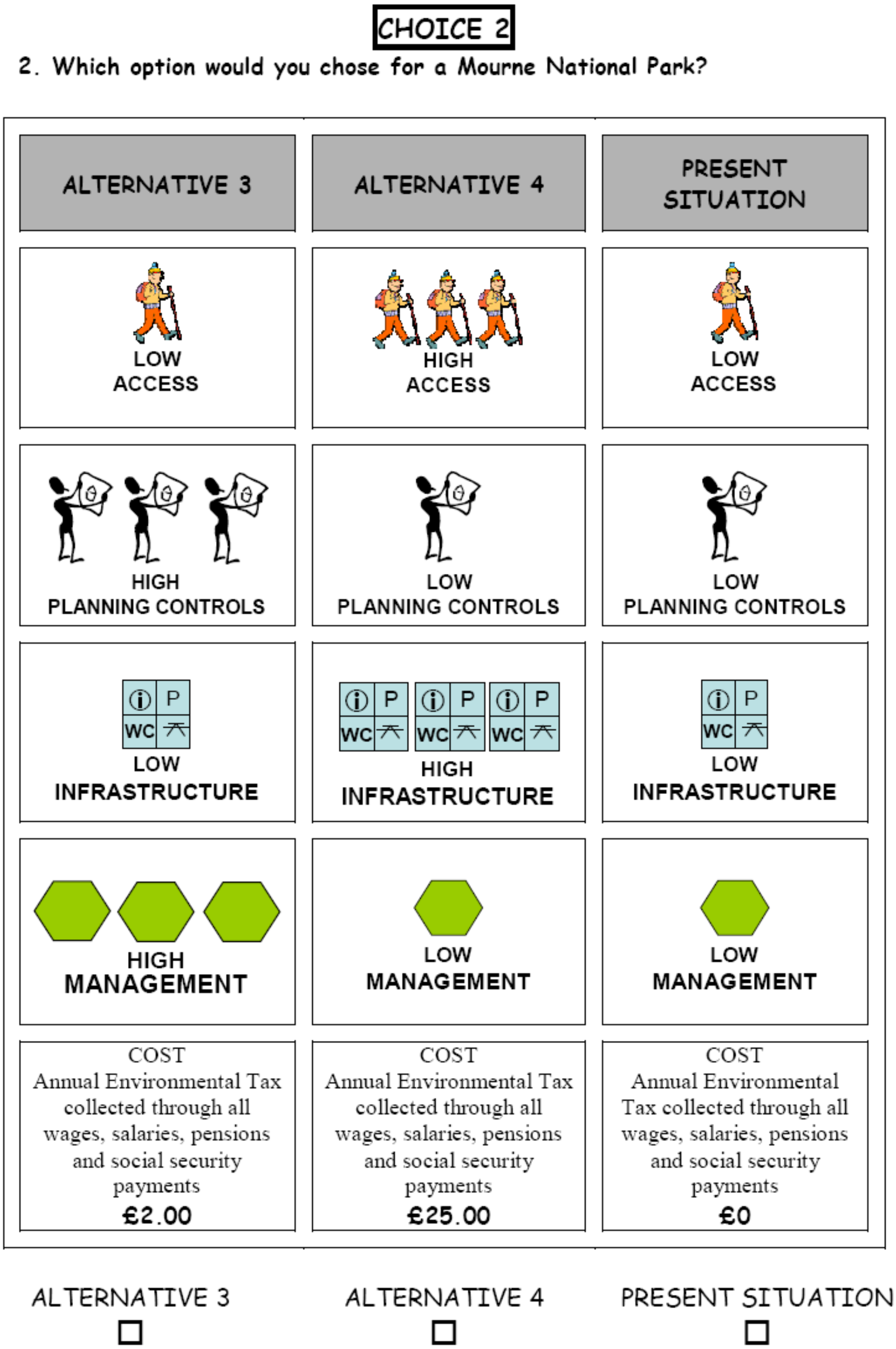


Yes ☐ No ☐ Don't know ☐

If you have answered "Yes", how many days would you visit the Mournes area in a 12 month period?

____

Figure 2. Example of CE question.



4. Results

Of the 4,507 surveys sent, we received 647 questionnaires back, for a response rate of 14.36%. On average respondents had spent 2.49 days to the Mournes mountains in the last 12 months. When asked how many days they would spend in the next 12 months at the current condition, they claimed to be willing to spend 5.07 days. This result is consistent with previous studies that found that respondents may overestimate the number of days they are willing to spend at a recreational site in the future under the status quo situation (Whitehead et al, 2000). Therefore, in the CB model, when we assess how changes in the levels of infrastructures, access, planning controls and management to the area affect the number of expected days compared to the current situation, we use the expected number of days that respondents expect to spend at the Mournes in the next 12 months and not the number of days they spent in the past 12 months. Table 1 reports the descriptive statistics of the sample, and compares them with the population's characteristics.

Table 1. Descriptive statistics

Variable	2001 Census	Our sample
Male	48.47%	41.05%
Single	29.67%	16.88%
Married	51.47%	60.83%
Unemployed	4.14%	4.45%
Retired	10.98%	22.10%
Working full time	37.55%	43.72%
Working part time	9.945	5.56%
Permanently sick/disabled	9.33%	13.99%
Looking after the home	5.70%	6.67%
Median age	34	49

Table 2 reports the results from a Random Effects Poisson model. The model shows that, as predicted by economic theory, the coefficient for the travel cost to the Mournes (COST_M) is negative and significant, suggesting that as the travel cost to visit the Mournes increases, the number of trips taken to the site decrease. We used as a substitute site the Sperrins mountains; the coefficient of COST_S, the travel cost to the substitute site, is negative and significant, as expected, suggesting that the number of trips to the Mournes increase as the travel cost to the Sperrins increase. Our respondents report that they expect to visit the Mournes more often with improvements in the attributes used in the CB questions: the sign of the coefficients for improvement in infrastructures (CBINFRA), in access levels (CBACCESS), in planning permissions (CBPLANNI) and in management of the area (CBMANAGE) are positive and significant, except for CBMANAGE which is not significant.

Table 2. Random Effects Poisson model for the demand for recreation.

Contingent Behaviour	Random Effects Poisson Model	
	Coeff	t-stat
Constant	-0.6953	-0.81
COST_M	-0.0929	-9.09
COST_S	0.0955	9.31
CBINFRA	0.0722	3.37
CBACCESS	0.1029	6.45
CBPLANNI	0.0471	2.53
CBMANAGE	0.0266	1.53
Age	0.0338	1.16
Age squared	-0.0003	-1.29
Male	0.1706	1.11
Wage1 (£60/2)	0.0392	2.91
Wage2 (£190 + £60 / 2)	0.0076	5.00
Wage3 (£580 + £290 / 2)	0.0027	4.18
Wage4 (£580 + £960 / 2)	0.0014	3.32
Wage5 (£3000/2)	0.0007	2.80
Single	-0.3784	-1.89
Cohab	0.0929	0.24
Divorced	-0.1317	-0.40
Separated	0.3990	1.21
Widower	0.1835	0.71
Degree	-0.1800	-0.69
Highers	0.6976	2.60
A Level	-0.1670	-0.63
Nationals	-0.1124	-0.29
GCSEs	-0.2873	-1.03
CSEs	-0.3035	-0.84
Other education	-0.4145	-0.95
Alpha	1.9830	19.16
Observations	3235	
Log Likelihood Function	-6325.214	

The focus of this paper is to check whether the results from this model that explains the demand for recreation to the Mournes can help better explain the answers that our respondents gave to the CE questions, where respondents were asked to choose their most favourite program for improving the Mournes.

In table 3 we report the outcome from two Mixed Logit Models (2) and (3) that explain the answers to the CE questions. Both specifications use as explanatory variables the status quo (current situation), and the attributes used in the CE questions, assuming a random parameters specification, where the random part of the coefficients are distributed as a triangular distribution with mean and spread constrained to be equal. The difference between specifications (2) and (3) is given by an additional variable in (3). Model (3), in fact, uses the residuals calculated from model (1) as an additional explanatory variable.

Table 3. Mixed Logit Models for the CE data.

Random Parameters Logit Model, triangular distribution with the mean and spread constrained to be equal, 500 replications				
	Model 2		Model 3	
		Parameters		
	Coeff	t-stat	Coeff	t-stat
Current situation	0.7754	3.95	0.7527	3.83
Access	0.7910	5.07	0.7786	5.01
Planning	0.7503	5.56	0.7563	5.60
Infrastructure	1.2003	7.90	1.1956	7.86
Management	0.6658	5.76	0.6639	5.70
Tax	-0.0817	-7.15	-0.0827	-7.17
CB Residuals			0.0561	3.02
		Derived Standard Deviations of parameters distributions		
Access	0.7910	5.07	0.7786	5.01
Planning	0.7503	5.56	0.7563	5.60
Infrastructure	1.2003	7.90	1.1956	7.86
Management	0.6658	5.76	0.6639	5.70
Tax	0.0817	7.15	0.0827	7.17
CB Residuals			0.0561	3.02
Loglikelihood function		-1099.5		-1074.04
Observations		1053		1031

The results show that improvements in the levels of access, planning permissions, infrastructure and management at the Mournes make respondents more likely to select a project. The cost (annual tax) is negative and significant, as predicted by economic theory, suggesting that respondents shy away from more expensive projects, all else being the same. The coefficient of CB Residuals in (3) is positive and significant. The part use the 'use value' utility from the model studying the demand for recreation helps to improve the estimates from the CE model. As the coefficient is positive, it suggests that the unobserved part of the 'use value' utility from the model for the demand for recreation has a positive impact on the probability that a respondent would choose a program in the CE exercises, thus implying that specification (2) underestimates the 'use value' component of the total economic value of a program to improve the Mournes mountains. In addition, the estimates of the coefficients for the attributes change from specification (2) to (3). Model (2) overestimates the effects of access, infrastructure, and management, but underestimates the effect of planning when modelling respondents' choices to the CE questions. We conclude that specification (2) presents biased coefficients estimates compared to specification (3).

5. Conclusions

This paper has investigated whether CE used to assess the WTP for improvements at recreational sites suffer from an omitted variable problem. When respondents select their most preferred option in CE questions, the importance of the use value component arising from the expected number of trips that respondents would take under each hypothetical scenario is underestimated if a logit model that explains the CE questions only uses the attributes of the CE as explanatory variables. Using a control function approach we improved the estimates of the logit model explaining the CE data. The implications of our results are important, as failing to consider this omitted variable problem leads to biased estimates, hence welfare biased estimates.

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, Vol. 26, pp 271-292.

Adamowicz, W., Swait, J., Boxall, P., Louviere, J., and Williams, M. (1997). Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation. Journal of Environmental Economics and Management, Vol. 32, pp 65-84.

Adamowicz, W., Boxall, P., Williams, M./Louviere, J. (1998) Stated preference approaches for measuring passive use values: Choice experiments and Contingent Valuation. American Journal of Agri Economics, Vol. 8, No. 1, p64-75.

Alberini, A. and Longo, A. (2006) Combining the travel cost and contingent behaviour methods to value cultural heritage sites: Evidence from Armenia. Journal of Cultural Economics. Vol. 30, p287-304.

Alberini, A., Zanatta, V. and Rosato, P. (2007) Combining actual and contingent behaviour to estimate the value of sports fishing in the Lagoon of Venice. Ecological Economics, Vol. 61, 530-541.

Bateman, I., Willis, K and Garrod, G. (1994). Consistency between Contingent Valuation Estimates: A comparison of Two Studies in UK National Parks. Regional Studies, Vol. 28, No. 5, pp 457 – 474.

Bhat, C. R., and Gossen, R. (2004). A mixed multinomial Logit model analysis of weekend recreational episode type choice. Transportation Research Part B, Vol. 38, pp767-787.

Bhat, M. G. (2003). Application of non-market valuation to the Florida Keys marine reserve management. Journal of Environmental Management, Vol. 67, pp 315-325.

Cameron, T. A., Shaw, W. D., Ragland, S. E., Mac Callaway, J., and Keefe, S. (1996). Using Actual and Contingent Behaviour Data with Differing Levels of Time Aggregation to Model Recreation Demand. Journal of Agricultural and Resource Economics, Vol. 1, No. 1: 130-149.

Cameron, T.A. (1992). Combining Contingent Valuation and Travel Cost Data for the Valuation of Non market Goods. Land Economics, Vol. 68, No. 3, p302-317.

Christie, M., Hanley, N., Hynes, S. (2007). Valuing enhancements to forest recreation using choice experiment and contingent behaviour methods. Journal of Forest Economics, Vol. 13, p75-102.

Clawson, M, and Knetsch, J. L. (1966). Economics of Outdoor Recreation, Washington, DC. Resources of the Future.

EHS (NI). (2004). A Mourne National Park? EHS (NI).

Englin, J. and Cameron, T. (1996). Augmenting Travel Cost Models with Contingent Behaviour Data: Poisson Regression analyses with Individual Panel data. Environmental and Resource Economics, Vol. p, p133-147.

Eom Y., and Larson, D. M. (2006). Improving environmental valuation estimates through consistent use of revealed and stated preference information. Journal of Environmental Economics and Management, Vol. 52, pp 501-516.

Europarc Consulting, (2002). Special Places need Special Care: A Report on designating AONBs of Northern Ireland as National Parks. Europarc Consulting, Göttingen, Germany.

Guevara, C.A. and M. Ben-Akiva (2006) "Endogeneity In Residential Location Choice Models", *Transport Research Board*, 1977(10): 60-66.

Hanley, N., Bell, D. and Alvarez-Farizo, B. (2003). Valuing the benefits of coastal water quality improvements using contingent and real behaviour. Environment and Resource Economics, Vol. 24, No. 3, p273-285.

Huang, J., Haab, T. C., and Whitehead, J. C. (1997). Willingness to Pay for Quality Improvement: Should Revealed and Stated Preference Data be Combined? Journal of Environmental Economics and Management, Vol. 34, pp 240-255.

Inform Communications. (2007). Report to the Mourne National Park Working Party on the outcome of their Public Consultation Process "Your opportunity to have your say". Belfast.

Johnson, F.R., Kanninen, B., Bingham, M., and Ozdemir, S. (2007). "Experimental Design for Stated Choice Studies," in Kanninen, B. (ed) "Valuing Environmental Amenities Using Stated Choice Studies" Springer, Dordrecht, The Netherlands.

Kirk, D. (2002). The Mountains of Mourne – A celebration of a Place apart. Appletree Press, Belfast.

Louviere, J., and Timmermans, H. (1990). Stated Preference and Choice Models Applied to Recreation: A Review. Leisure Sciences, Vol. 12, pp9-32.

Luce, R. D. (1959). Individual Choice Behaviour: A theoretical Analysis. John Wiley & Sons, New York.

McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behaviour. In Zarembka, P. (ed.), Frontiers in Econometrics. Academic Press, New York.

Northern Ireland Planning Advisory Board (1947). The Ulster Countryside: Report on amenities in Northern Ireland. HMSO, Belfast.

Petrin, A., and K. Train (2003). Omitted Product Attributes in Discrete Choice Models. NBER Working Paper No. W9452 National Bureau of Economic Research.

Whitehead, J. C., Haab, T. C., and Huang, J. (2000). Measuring recreation benefits of quality improvements with revealed and stated behaviour data. Resource and Energy Economics, Vol. 22, pp 339-354.

Whitehead, J. (2005). Combining willingness to pay and behavior data with limited information. Resource and Energy Economics, Vol. 27, o143-155.

Whitehead, Pattanayak, S. K., Van Houtven, G. L. and Gelso, B. R. (2007). Combining Revealed and Stated Preference data to estimate the Nonmarket Value of Ecological Services: As Assessment of the State of Science. Working Paper, 05-19, Department of Economics, Appalachian State University.

Whitehead, J. C., Dumas, C. F., Herstine, J., Hill, J. and Buerger, B. (2008). Valuing Beach Access and Width with Revealed and Stated Preference Data. Working Papers 06-15, Department of Economics, Appalachian State University.