Discussion Paper No. 114

THE AGGREGATION OF ENVIRONMENTAL BENEFIT VALUES:
WELFARE MEASURES, DISTANCE DECAY AND TOTAL WTP

Ian Bateman[1]
Brett H. Day[1]
Stavros Georgiou[2]
and
Iain Lake[3]

September 2006

[1]   Corresponding author: Ian J. Bateman, Centre for Social and Economic Research on the Global Environment (CSERGE), University of East Anglia, Norwich, NR4 7TJ, United Kingdom, e-mail i.bateman@uea.ac.uk, tel: +44 (0)1603 593125, fax.: +44 (0)1603 507719.

[2]   Department of Economics, Brock University, Canada

[3]   Environment Policy Economics, Department for Environment, Food and Rural Affairs (DEFRA), London, UK

The aggregation of environmental benefit values:
Welfare measures, distance decay and total WTP

Ian J. Bateman
Brett H. Day
Stavros Georgiou
and
Iain Lake

Abstract

We review the literature regarding the aggregation of benefit value estimates for non-market goods. Two case studies are presented through which we develop an approach to aggregation which applies the spatial analytic capabilities of a geographical information system to combine geo-referenced physical, census and survey data to estimate a spatially sensitive valuation function. These case studies show that the common reliance upon political jurisdictions and the use of sample mean values within the aggregation process are liable to lead to significant errors in resultant values. We also highlight the fact that for resources with use values then we should expect overall values to reduce with increasing distance from such sites, but that changes in the choice of welfare measure will determine whether such 'distance decay' is to be expected within values stated by those who are presently non-users. The paper concludes by providing recommendations for future improvements to the methodology.

Keywords: *Valuation; aggregation; total willingness to pay; distance decay; selection bias; contingent valuation; water quality; wetlands; rivers.*

Contents

List of Tables

List of Figures

Prologue

"Once upon a time a group of eminent chefs were asked to prepare the finest Horse and Rabbit Stew for the King's Birthday. For days they toiled and argued about how the rabbit should be prepared. Should it be roasted first or simply boiled, or perhaps jugged or marinated in a fine wine sauce. Finally they came to an agreement and as the birthday dawned they prepared the finest rabbit ever tasted for the stew. Sadly all their efforts were completely overwhelmed by the subsequent addition of the horse".[1]

1.   Introduction

Cost benefit analysis requires the aggregation of individuals' benefits in order to compare with the total costs of a project or policy. Because the methods for measuring non-market benefit values are based on analyses of individual behaviour, there is a problem in knowing how changes in a resource will affect aggregate values. This will depend on both the benefits per person and the population of beneficiaries. It is thus necessary to identify the extent of the market and how this varies with changes in the good. As suggested by *The Tale of Horse and Rabbit Stew*, the extent of the market may well be more important in determining aggregate values than any changes related to the precision of the estimates of per-person values. This paper builds upon earlier work by Smith (1993), Loomis (2000) and others to argue that differences in measures of individual values are the haggled-over rabbits in the stew, whereas the aggregation process itself is the horse which can so pervasively dominate the outcome.

The issue of defining the extent of the market for a public good essentially concerns the question of how broadly should individuals' marginal benefit schedules be vertically summed. Whether this should be confined to those living in the close vicinity of the good or extended across the region, country or even further afield, has implications for the appropriate level of government financing and provision. This issue becomes further complicated when considering the distinction between user and non-user values.

In considering the extent of the market it is useful to distinguish between the *political jurisdiction*, concerning some administrative area, and the *economic jurisdiction* incorporating all those who hold economic values regarding a project. From an economic

---

[1] With thanks to Colin Green for *The Tale of Horse and Rabbit Stew*.

cost-benefit point of view, the notion of *accounting stance* (or who has *standing*) defines the relevant jurisdiction for including benefits and costs. Now in order to achieve efficiency in resource allocation, it is necessary that the accounting stance should be such that it captures all Pareto-relevant impacts. Where political and economic jurisdictions coincide there is said to be *fiscal equivalence* (Olsen, 1969; Cornes and Sandler, 1996). However, it seems likely that such equivalence will be the exception rather than the rule. Simple logic suggests that, where some resource site generates any element of use value, the density of such users will be higher near to that site. Furthermore, as users typically hold higher values than non-users then we would expect average values to decay with increasing distance from that site. It seems likely that the resultant economic jurisdiction will often (but, as discussed subsequently, not always) fail coincide with any political jurisdiction. Although there is an extensive literature examining variability in estimates of mean WTP (see, for example, Hanemann and Kanninen, 1997; Kerr, 1996; Ready and Hu, 1995), nevertheless Smith (1993) suggests that such variability will have far less impact on estimates of total WTP than will changes in the extent of the market arising from the use of different approaches to aggregation. He provides an example related to the estimation of natural resource damages, in which the differences in plaintiff's and defendant's estimates of per unit values were small but total damage assessment differences were huge as a result of each parties assumptions regarding the appropriate jurisdictions used to define the extent of the market.

The highest profile UK application of a political jurisdiction approach to aggregation occurred in 1998 when the Environment Agency (EA) refused a water company application to abstract water from the River Kennet in southern England (ENDS, 1998; Moran, 1999). The EA approach followed recommendations from FWR (1996) to use water company operations boundaries as the relevant area for aggregation. Total benefit value was then calculated by multiplying the population within this jurisdiction by a sample mean WTP estimate transferred from a contingent valuation (CV) study of the River Darrent (Willis and Garrod, 1995). Aside from concerns regarding differences in the river types, commentators criticised the use of a sample mean taken from one area and applied without adjustment to an entirely different area which not only had differing geographic size but also contained a much larger population with very different socioeconomic characteristics (Moran, 1999; Bateman et al., 2000). These academic reservations were reflected in judicial review through which the EA ruling was overturned after an appeal in which the plaintiff attacked both the valuation and aggregation procedure employed.

Now in principle the use of a sample mean WTP value need not necessarily lead to biased estimates of aggregate values. If a representative[2] sample is drawn from the entire economic jurisdiction (or indeed some larger sample area) then multiplying the sample mean WTP by the population of the sampled area should give an accurate estimate of aggregate values in that sampled area. Indeed even if the sample area is a subset of the economic jurisdiction then the sample mean approach to aggregation still gives unbiased estimates of total value within that subset area. However, problems may well arise where a mean value from some subset area is used as the basis for estimating aggregate values for the entire economic jurisdiction (or some non-coincident political jurisdiction, as in the River Kennet case). There are two sources of error here. First, as noted above, underlying values for changes to some spatially confined resource are likely to decay with increasing distance from that resource (an issue we discuss in greater detail with respect to both users and non-users, in Section 2 of this paper). Failure to account for this distance decay will lead to error if this mean value is used to aggregate values for a larger (or indeed smaller) area. This problem is exacerbated by common practice where valuation studies indeed do fail to sample entire economic jurisdictions, but instead focus survey effort upon areas around the resource in question where values are highest. However, this is where underlying values are highest. The application of consequent sample means in an unadjusted manner to some large (typically political) jurisdiction seems liable to generate over-estimates of aggregate values. A second problem arises when, as seems likely, the probability of responding to a valuation survey is positively related to underlying values. In such cases self-selection bias (Heckman, 1974) is likely to increase with distance from the resource. Thus, while underlying values should exhibit distance decay, sampled values will understate this as with relatively higher values self-select themselves into the sample at a greater rate than those with lower values. Aggregation procedures need to recognise and address both of these problems.

Clearly given unlimited survey resources one would ideally sample from the entire feasible economic jurisdiction. But even here we face the problem that, a-priori, the extent of this area is unknown. In this paper we argue that, given these challenges, a superior approach to aggregation is to use survey data to adjust for self selection bias, capture underlying distance decay and define the economic jurisdiction. This, we argue, is best achieved through the identification and estimation of a *spatially sensitive valuation function*. Such an approach

---

[2] The concept of representativeness is worthy of comment. Typically this refers to issues such as the representation of population socioeconomic and demographic characteristics within a sample (Morrison, 2000). We do not discount the importance of such issues, but effectively argue throughout this paper that the issue of spatial representativeness is also important, particularly when we are aggregating values for spatially confined resources.

explicitly addresses self-selection and incorporates distance decay relationships into defining the limits of the economic jurisdiction while allowing for variability in the socioeconomic characteristics of the encompassed population within the aggregation process. We use this approach to generate estimates of aggregate WTP and contrast this with measures based upon both the use of a political jurisdiction and the reliance upon sample means. The impact of using different aggregation approaches is then contrasted with the variability induced by uncertainty in the estimate of mean WTP to examine Smith's contention that such variability will be overwhelmed by aggregation impacts.

The following section considers expectations regarding the underlying decay of values across space, highlighting important interactions between choice of welfare measure and the expected response of users and non-users of a site. We utilise these expectations to guide a review of existing evidence regarding such distance decay. This highlights a divergence in expectations depending on whether we are estimating an equivalent loss measure (in which respondents are typically asked to state their WTP to preserve some resource such that the future level of resource quality stays the same as its present quality) or a compensating surplus measure (in which respondents are typically asked to state their WTP for some improvement in the resource such that the future level of resource quality exceeds the present quality). This divergence is reflected in the empirical part of this paper. Section 3 considers self selection bias and reports an aggregation exercise for an equivalent loss measure using sample mean and valuation function approaches applied to a case where political and economic jurisdictions coincide. In contrast Section 4 aggregates a compensating surplus measure and present aggregate values estimated across differing political and economic jurisdictions as calculated using both sample mean and valuation function approaches. Through these applications we develop a methodology based upon the spatial analytic capabilities of a geographic information system (GIS). Section 5 discusses results from both our studies and embeds them within the previously reviewed literature. Section 6 concludes by giving some pointers toward ways in which this methodology might usefully be extended.

2. Expectations and existing evidence regarding distance decay in underlying values.

We can begin by considering the spatial distribution of values for some open access, public goods resource at its present level of quality. For the sake of simplicity we can identify two types of individual; users of a resource, and non-users. Users will hold use and option-use values and may well hold non-use values. Non-users may hold option and non-use values or have negligible WTP[3]. It seems reasonable to assume therefore that, *ceteris paribus*, users will typically hold higher WTP values than non-users.

Should we expect that the values held by those who are presently users of a resource will decline as distance from that resource increases? There are a number of issues here. It is clearly true that, as distance and travel time increase so the costs of accessing a site also rise. Also it may well be that the availability of substitutes increases with distance, lowering their opportunity cost. But these are cost side issues and do not dictate a reduction in benefits. The benefits held by someone who travels halfway round the globe to visit the Grand Canyon are likely to be higher than the *average* resident of nearby Phoenix[4]. Now this lack of relation need not always hold. For example, Parsons' (1991) commentary upon the travel cost (TC) method notes that certain of those individuals with high use values for recreational resources will internalise some costs by choosing to live in locations near to those resources. However, this will be less of an issue for resources which have been degraded for long periods of time (as in the case of the River Tame discussed in Section 4). Therefore, it is not *a-priori* clear that the benefits of those who are already users will necessarily exhibit a strong distance decay (although rising costs mean that their *net* benefits do decay with increasing distance).

Turning to consider those who are non-users of a resource, again it is not clear that values should decay across space. After adjusting for income etc., why should non-use values for preserving the Amazon rainforest be higher in the USA than in more remote Europe?

---

[3] The possibility of negative WTP also arises although, as we argue subsequently, this does not seem credible for the resources considered in our case studies.

[4] This example does highlight a practical problem facing almost all aggregation exercises. As most benefit-cost studies extend at most to national borders, values beyond that extent are typically ignored. This problem occurs not only in mean value aggregations but also in the application of valuation functions as these omitted observations have no opportunity to influence the shape of the fitted curve. The extent of the consequent error is likely to directly depend upon the international prominence of the resource in question, with higher prominence resources (such as the Grand Canyon, or the internationally significant wetland of our first case study) suffering higher error than lower prominence good (such as the resource considered in our second case study). However, such cross-border values are typically ignored in national decisions.

Empirically there is some (limited) evidence that there may be a cultural identity or 'ownership' dimension to non-use values with, those who live closer (or even in the same country) as a resource expressing relatively higher non-use values (Hanley et al., 2003; Bateman *et al.,* 2005a). However, the theoretical basis for this association is unclear.

What is evident from the above is that, as distance from a resource increases and opportunity costs rise, so both the number of users and the proportion of users to non-users will decline. Given that users will generally hold higher values than non-users these trends will result in a distance decay in mean per household values. This will be compounded by both the 'Parsons effect' and any empirical reduction in non-user values across space.

Now all of the above holds for situations in which there is no potential change in resource quality. However, valuation studies invariably estimate welfare measures for situations in which there is some potential change in resource provision or quality. Two measures dominate the CV literature in this respect; equivalent loss and compensating surplus. Equivalent loss studies ask survey respondents to state their WTP to preserve some resource such that the future level of resource quality stays the same as its present quality. In such a situation, while we would still expect to observe distance decay in values for the overall sample (because users are clustered nearer to the site)[5], there seems no theoretical reason to expect such a trend to be seen within just the responses of those who are present non-users. If they were non-users prior to the project then, as quality has not changed, they should remain so after the project has completed[6]. This is not the case for compensating surplus studies where respondents are asked to state their WTP for some improvement in the resource such that the future level of resource quality exceeds the present quality. Here some present non-users may see themselves as likely to become users of the improved good. This trend will become more significant the larger is the proposed improvement in resource quality[7]. Importantly, this non-user to user conversion is more likely to occur for households nearer to the resource than those further away. Hence we would therefore expect to observe a distance decay effect not only across the overall sample of respondents, but also amongst those who are presently non-users.

---

[5] Note that the WTP of those that were already users will not change due to such a project.
[6] While this seems to be the economic-theoretic expectation, the very act of being surveyed might convert some prior non-users into higher value future users if it raises awareness of the resource. Another possibility that a 'focussing illusion' (Schkade and Kahneman, 1998; Loewenstein and Frederick, 1997; Loewenstein and Schkade, 1999; Kahneman and Sugden, 2005) may lead to non-users overstating WTP amounts although whether this would have a distance decay dimension is unclear.
[7] And, of course, this improvement in site quality may also increase user WTP although it is not obvious that there should be a spatial dimension to this particular increase.

We therefore have clear expectations of a difference in the nature of distance decay relationships according to the welfare measure under investigation. These expectations are borne out through inspection of those (relatively few) prior studies which have tested for distance decay in CV WTP responses. Equivalent loss studies examining values for the *preservation* of water quality (Sutherland & Walsh, 1985), conservation areas (Imber et al., 1991), endangered species (Loomis, 2000), wetlands (Bateman et al., 2000) and remote mountain lakes (Bateman et al., 2005a) reflect expectations of significant distance decay in overall WTP values but not in present non-users values. Conversely, but again in line with expectations, compensating surplus studies valuing *increases* in wetland or bird life (Pate and Loomis, 1997) or *improvements* in forests (Mouranaka, 2004), river flows (Hanley et al., 2003) or river water quality (Bateman et al., forthcoming), demonstrate distance decay not only in overall sample values but also in the values stated by present non-users.

Summarising the discussions of Sections 1 and 2; we have argued that in a substantial number of cases the reliance upon sample means will fail to yield accurate measures of aggregate WTP. As a more robust and generally applicable alternative we propose an approach based upon the estimation of a spatially sensitive valuation function which takes into account issues of self-selection, the expected distance decay of values and the impact of variation in the socioeconomic circumstances of the relevant aggregation population. With reference to existing literature we have also argued that we should expect that the distance decay characteristics of any such valuation function will depend in part upon the choice of welfare measure, with compensating surplus but not equivalent loss measures exhibiting distance decay amongst the values of present non-users although both measures should give distance decay across the overall sample. Furthermore, aggregate values are likely to vary considerably according to whether political or economic jurisdictions are used; a variability which has the potential to outstrip that caused by uncertainty in estimates of mean WTP measures. In the following two case studies we attempt to address each of the issues raised above. In so doing we develop a method for generating spatially sensitive value functions by harnessing the spatial analytic power of a geographic information system (GIS).

3. Case Study I: Aggregating an equivalent loss measure: WTP to preserve a National Park wetland

The first case study is based on data taken from a postal CV survey of WTP for schemes designed to preserve the Norfolk Broads National Park in Eastern England from the threat of saline flooding (i.e. an equivalent loss measure). Approximately one thousand questionnaires were sent to households across Great Britain and a GIS was used to calculate distance from each address to the Norfolk Broads. Following two reminders a sample of 310 completed questionnaires was collected. This relatively modest return rate was expected given that surveys were equally distributed throughout the country, rather than being concentrated around the resource site. This permitted us to examine the determinants of response rates given constant survey effort. The GIS was used in conjunction with census data to extract socioeconomic profiles for all areas. This information was combined with the distance estimates to permit the estimation of a survey response (self-selection) function. The decision to reply to the survey was found to be inversely related to the respondents distance from the Norfolk Broads but positively related to the socioeconomic status of the local area of the address to which the questionnaires were sent. Given our theoretical expectation of distance decay and the positive association of income with WTP we can see that such self selection bias will inflate the sample mean estimates of WTP and upwardly bias an aggregation based upon that mean. However, analysis of this sample self selection (reported in Bateman and Langford, 1997) allows us to incorporate this within the valuation function approach to aggregation and thus avoid this over-estimation problem. There remained the issue of how to treat non-respondents who may or may not hold values for flood prevention scheme. It was decided to follow the lead of Bishop and Boyle (1985) and adopt the simple lower bound assumption that all non-respondents have zero WTP.

Analysis of the WTP responses themselves showed that these were also negatively related to distance from the Norfolk Broads and positively related to the respondents' income[8]. In line with expectations regarding equivalent loss scenarios, no significant distance decay was observed within the responses of those who were present non-users of the resource.

Analysis of the value function for the overall sample gave non-zero value predictions throughout Great Britain. This is perhaps not surprising given the National Park status of the

---

[8] While we adopted a two stage modelling approach here, a more efficient approach would be to use the Heckman hurdle modelling method (Heckman, 1974; 1987) to jointly examine the self-selection decision and WTP response.

resource, the major losses envisioned in the scenario and the relatively small size of the country. Therefore, within this application the economic and political jurisdictions appear to coincide as being the entire country and any differences should be due to accounting for self selection bias inflating the sample mean[9].

An investigation of the impact of WTP uncertainty was undertaken by truncating the base data of the 2.5% highest and lowest valuation responses and recalculating all aggregate sums using both the sample mean and valuation function approaches. The robustness of the latter approach was further investigated by using two measures of distance; the first a cruder measure using just four distance bands to cover the entire country, such as might be used in a absence of a GIS; the second using a more accurate, GIS generated, estimate of distance.

Full details of the aggregation calculations are given in Bateman et al., (2000) from which the results presented in Table 1 are drawn. Here we see that the use of an unadjusted sample mean in approach (1) yields aggregate values which are both much higher than those generated from the value function approach (2) and more subject to the effects of truncating outlier WTP values. By contrast the value function measures appear highly stable both to such truncations and to cruder approximations of distance than those available using a GIS. The comparison between aggregate estimates is remarkable, showing that the self-selection bias inherent in the simple mean WTP aggregation approach has in this case resulted in values which are more than six times higher than those obtained using the value function approach.

---

[9] As discussed previously, the possibility that cross-border values exist and have been truncated is acknowledged.

Table 1

Aggregate benefits of preserving the present condition of the Norfolk Broads National Park (£ million/annum). Economic and political jurisdiction coincide.

Aggregation approach	Untruncated values	Upper and lower 2.5% of responses truncated
(1) Aggregation using sample mean WTP	159.7	98.4
(2) Aggregation using a spatially sensitive valuation function		
i. using four distance zones	27.3	25.3
ii. using more accurate distance measures	25.4	24.0

Notes: Estimates are in £, 1992 values.

4. Case Study II: Aggregating a compensating surplus measure: WTP to improve urban river water quality

The second case study extends the GIS-based valuation function approach to aggregation developed in the prior analysis. Data is taken from a face-to-face CV survey of WTP for water quality improvements to the River Tame (i.e. a compensating surplus measure), for which a sample of 675 responses was gathered in interviews conducted at the respondents home address. The river rises on the west of the City of Birmingham, in Central England, flows eastward through the city and turns north till it meets with other elements of the Trent catchment. It is therefore a relatively minor asset compared to the National Park of the previous application. Furthermore its urban course means that it has suffered long term degradation and is now classed as one of Britain's most polluted rivers with fish stocks virtually non-existent and plant growth, insects, bird and animal life all severely limited. However, the river does have ecological and recreational potential and passes through high density residential areas, playing fields and a country park[10].

---

[10] Full details of the case study area are given in Georgiou et al (2000) and Bateman et al., (2005b and forthcoming)

In order to assess the impact upon aggregate values of changes in the quality of a resource the study design presented respondents with three nested levels of river quality improvement (referred to here as the *Small*, *Medium* and *Large* improvements). The survey instrument described these changes in terms of consequent improvements to water quality, fishing, plants and wildlife, boating and swimming. Respondents were told that these three improvements were alternatives to each other, of which only one would be implemented, and therefore should be evaluated relative to a common baseline of the current situation, which was also described[11]. An open-ended elicitation format was applied and the multiple valuation design of the study meant that each respondent provided three values, one for each of the improvements. In order to allow for possible intra-respondent correlation of these values, responses were subsequently modelled using a generalised least squares random effects approach (Greene, 1990).

The GIS was used to calculate distances from each respondent's home address to the River Tame[12]. This was combined with survey data detailing individuals' knowledge of water quality in the River Tame, visitation patterns, attitudinal data, etc. However, two survey deficiencies should be acknowledged here. First, household income data was not consistently collected and had to be subsequently proxied from small area census data (a less than ideal situation but, we feel, an adequate response given our use of census data at the smallest 'output area' level)[13]. Second, the address of those who refused to be interviewed was not recorded[14]. Therefore, we cannot adequately incorporate self-selection bias into the present study and recognise that consequent aggregation estimates in both the sample mean and valuation function approaches are upwardly biased. While the relatively small economic jurisdiction of this good (discussed subsequently) reduces the absolute size of any error relative to our previous study, this is a deficiency which should be addressed in future applications. However, as this error affects both of our aggregation approaches it does

---

[11] Tests for ordering effects showed that these were insignificant, which we argue is a consequence of respondents' prior awareness of all schemes before valuations were elicited (Bateman et al., 2004).

[12] Bateman et al., (2005b) reports analyses for three alternative distance measures; a simple measures taken by hand from a paper map (to simulate analyses where a GIS is not available); distance to major access points along the river; and the Euclidian distance measure reported in the present paper. All three provide similar results (verified through Hausman tests) although the GIS measures were found to be somewhat superior to the simple paper map approach.

[13] While point data is more informative than census tract or other polygon location data, Case (1991) and Swinton (2002) discuss conditions under which one can proxy for the other. The former paper also discusses spatial components of the error term, an issue which is not considered in the present paper. This may have reduced the efficiency of our estimates.

[14] This was contrary to instructions to the survey company. In the event only a non-response tally was collected, showing that some 761 households refused to be interviewed.

(inadvertently) allow us to control for this factor when comparing the other treatments considered in this study.

Analysis indicated that the semi-logarithmic form of Equation (1) provided a best fit specification of the spatially sensitive valuation function:

$$\ln(WTP_i + 1) = \beta_0 - \beta_1 Distance_i + \sum_{k=1}^{K} \beta_k x_{ki} + \varepsilon_i \qquad (1)$$

where i denotes individuals and *Distance* and the other k covariates (x_{ki}) are defined as follows:

Variable	Definition
Ln(WTP+1)	Natural log of (willingness to pay + 1) (continuous)
Large	Large environmental improvement (1 = yes; 0 = otherwise)
Medium	Medium environmental improvement (1 = yes; 0 = otherwise)
Distance	Euclidian distance (in metres) from respondents home to the River Tame.
Env	Respondent's interest in environmental issues (1 = not interested at all; 5 = very interested)
Know	Did respondent already know the River Tame (1 = yes; 0 = otherwise)
Retire	Is respondent retired (1 = yes; 0 = otherwise)
Age	Respondent age (continuous)
Poor	Proxy low income variable (percentage of people aged 16-74 unemployed in the respondents local census output area)

Note that the dependent variable is calculated by first adding a value of one to the stated WTP in order not to lose any zero bidders in the subsequent log transformation process[15]. This specification also means that the semi-logarithmic approach is not incompatible with identifying the limits of the economic jurisdiction (where fully untransformed WTP is predicted to fall to zero)[16]. Table 2 estimates this model, reporting results for both the full sample of respondents and the sub-sample of present non-users.

---

[15] The use of Tobit analysis is only applicable in those cases where the latent variable can, in principle, take negative values and the observed zero values are a consequence of censoring and non-observability. Actually, in line with Randall et al, (2001), we do not believe that it is reasonable to suppose that individuals may lose welfare from the projects being valued in our research. Therefore, while for sensitivity purposes we report maximum likelihood random effects Tobit specifications in Bateman et al., (2005b), we believe these downwardly bias predictions of WTP.

[16] The geostatistics and GIS literatures also provide an alternative approach in the form of kriging for semivariograms with a nugget and a sill to identify areas beyond which there is no spatial effect.(see, for example, Isaaks and Srivastava, 1989; Oliver and Webster, 1990; Deutsch and Journel, 1992).

Table 2
Spatially sensitive valuation functions for both the full sample and for present non-users only.

	All respondents	Present non-users only
Distance	-0.0000771	-0.00008
	(2.97)**	(2.56)*
Large	0.623	0.676
	(14.69)**	(11.42)**
Medium	0.330	0.329
	(7.78)**	(5.56)**
ENV	0.219	0.298
	(2.90)**	(3.00)**
Know	0.356	0.503
	(2.36)*	(2.46)*
Retire	-0.601	-0.613
	(3.35)**	(2.45)*
Poor	-0.100	-0.100
	(3.06)**	(2.28)*
Constant	1.001	0.583
	(2.79)**	(1.22)
Observations	1179	660
Number of groups	393	220
sigma_u	1.41	1.41
sigma_e	0.60	0.62
Rho	0.85	0.84
Wald chi2	261.06	167.58
prob>chi2	0.000	0.000
r-squared	0.12	0.16
Max Distance: *Large*	27.75km	
Max Distance: *Medium*	23.94km	
Max Distance: *Small*	19.66km	

Dependent variable = *Ln(WTP+1)*
Absolute value of z statistics in parentheses
* significant at 5%; ** significant at 1%

Considering the model for all responses we find, as expected, highly significant distance decay in values. The two scope dummies, *Large* and *Medium* act as intercept shifters upon this distance decay of values and indicate not only the expected higher values for larger improvements, but also (given the semi-logarithmic form of this function) a somewhat more rapid decline for the latter[17] although values for larger schemes always remain above those which they nest. Other covariate effects are as expected with values being positively correlated with environmental concern and prior knowledge of the study area and negatively related to the constraints of retirement. The expected negative correlation with higher deprivation was also observed.

We now use our estimated value function (based on responses from the whole sample) to calculate the level of WTP for different areas progressing away from the study site. This is achieved by holding the non-distance individual level variables (*Env*, *Know*, *Retire* and *Age*) at their mean values, using the GIS to integrate with output area census data to generate values for the socioeconomic deprivation proxy (*Poor*), and for each level of improvement (*Large*, *Medium* and the base case of a *Small* improvement), calculating the predicted household WTP for each output area. Given the significant distance decay relationship estimated in this model it is clear that at some point values will decline to zero. This defines the economic jurisdiction within which there are positive WTP values. The maximum limits of this area for each level of improvement are given in the last three rows of Table 2. As would be expected we see that the limit for the *Medium* improvement is wider than that for the *Small* improvement, while limits for the *Large* improvement are greater than either of the nested schemes. However, it is interesting to note that the limit for even the largest scheme, at less than 30 km, is very much smaller than the economic jurisdiction observed in our previous case study concerning a National Park and reflects the smaller scale of benefits under consideration in the present application.

One of the advantages of a GIS-based methodology is the ability of the software to readily generate graphic representations of findings. After removing those areas that were beyond the boundaries of the economic jurisdiction estimated for the specific distance decay function being considered (i.e. those with negative WTP values), chloropleth maps were produced for our whole sample values of each of the improvements. These are illustrated in Figure 1 and show how the size of the economic jurisdiction varies according to the improvement scenario

---

[17] More rapid rates of change were investigated by adding two terms for the interaction between distance and the two scope variables. However, in this case both were found to be clearly statistically insignificant and so are omitted from the final model reported above.

considered, as well as how WTP varies in each output area census unit according to both its proximity to the River Tame and its socioeconomic characteristics. This latter facet of the results is amply illustrated by the contrast in values either side of the upper reaches of the Tame (its west to east reach). Here higher values are predicted for the more affluent north bank, with lower values estimated for the poorer central city area on the south bank.

Figure 1
Maps of estimated mean WTP (per household, per annum) of Census output areas for various water quality improvements
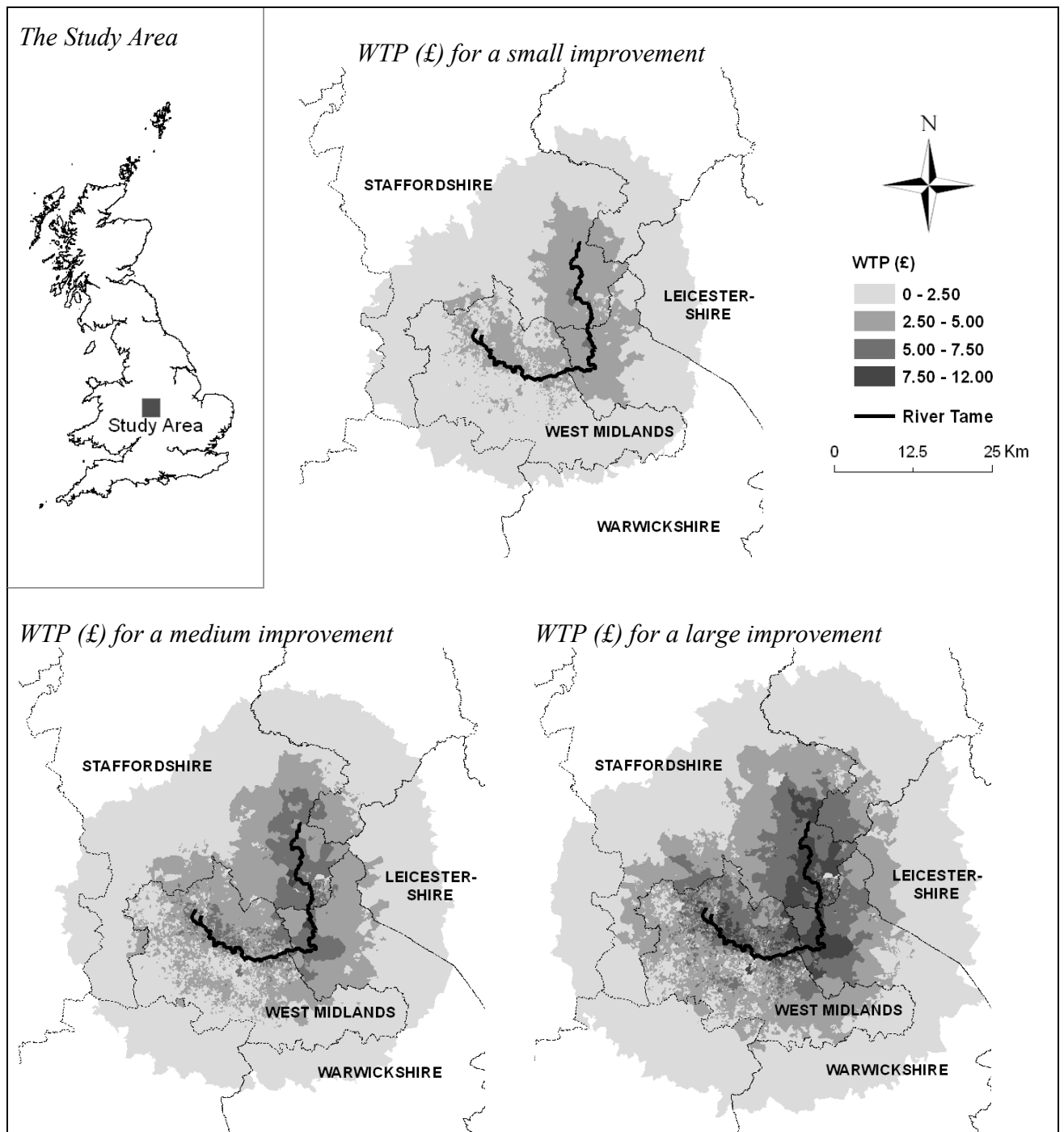
Turning to consider the model for present non-users values given in the final column of Table 2, a clear and highly significant distance decay effect is observed (p = 0.011). This result is characteristic of compensating surplus studies and stands in sharp contrast to the lack of a significant distance decay in present non-user values for the preceding equivalent loss case and prior studies. This contrast seems to support the contention that differences will arise across welfare measures because of their differing impact upon the quality of sites. Equivalent loss studies present scenarios in which final quality is maintained at initial levels. Irrespective of their distance to the site, this does not induce present non-users to become higher value users. However, compensating surplus measures present scenarios in which site quality increases. This induces some present non-users to convert to users, a conversion which is greater for those nearer to the site[18]. This in turn results in a distance decay in the values stated by those who are at present non-users of the resource, as observed here.

We can now use the above analysis to provide a spatially sensitive estimate of aggregate benefits for the economic jurisdiction and compare these with the standard approach to estimating aggregate benefits for a politically defined jurisdiction (such as the one adopted by the UK Environment Agency in the River Kennet enquiry). Under the latter approach we can define the aggregation population as simply those households which live within the relevant local water company area. Again following the EA approach we can then calculate aggregate benefit estimates by simply multiplying this population by the sample mean WTP. Resulting estimates are reported as the first column of results within Table 3. So as to allow comparison between aggregation errors (the 'horse' in the stew) and errors due to uncertainty regarding mean WTP (the 'rabbit'), we also report a confidence interval (CI) around these estimates based upon the 95% CI for the sample mean. The second column of results adopts the same approach to aggregation with the one refinement that the aggregation is now applied across the economic jurisdiction as defined by our spatially sensitive valuation function (as reported in Table 2). The population within the economic jurisdiction is substantially smaller than that of the political jurisdiction suggesting immediately that the latter is liable to lead to overestimation of aggregate benefits as it includes households for which WTP is at best zero (and arguably negative). Furthermore, as the scope of the good declines so the population within the economic jurisdiction becomes even smaller. This will progressively lead to greater error arising from reliance upon the EA political jurisdiction approach. Indeed the

---

[18] This spatial trend can be observed in the present distribution of users to non-users in the sample. This itself exhibits a highly significant distance decay (p<0.01) such that the proportion of users in the sample falls from nearly 50% near to the site to almost zero at a distance of 9 km. Further analysis of this trend is given in Bateman et al (2005b).

political jurisdiction method leads to estimates which are just over double that for the economic jurisdiction for the large improvement and more than two and a half times too high for the small improvement. These errors dwarf those due to uncertainty in the estimate of mean WTP which range from 17% for the large improvement to 20% for the small improvement.

Table 3
Aggregate benefits estimates based on sample mean and valuation function approaches

Quality change		Aggregation using sample mean WTP		Aggregation using WTP estimated from function
		Political Jurisdiction[1]	Economic Jurisdiction[2]	Economic Jurisdiction[2]
Large improvement[3]	Number of households	3,494,438	1,647,777	1,647,777
	Aggregate WTP	£82,049,404	£38,689,804	£5,040,526
	95% CI for aggregate WTP	(£68,001,763-£96,062,101)	(£32,065,740-£45,297,390)	
Medium improvement[4]	Number of households	3,494,438	1,486,415	1,486,415
	Aggregate WTP	£54,687,955	£23,262,395	£3,350,233
	95% CI for aggregate WTP	(£44,938,473-£64,437,437)	(£19,115,297-£27,409,493)	
Small improvement[5]	Number of households	3,494,438	1,336,736	1,336,736
	Aggregate WTP	£34,525,047	£13,206,952	£1,997,502
	95% CI for aggregate WTP	(£27,780,782-£41,269,313)	(£10,627,051-£15,786,852)	

Notes:
Estimates are in £, 1999 values.
1. Political Jurisdiction = local Water Utility Company area (Severn Trent and South Staffordshire Water Company Ltd.)
2. Area for which mean WTP > 0
3. Sample mean WTP for *Large* improvement = £23.48 (95% CI = £19.46: £27.49)
4. Sample mean WTP for *Medium* improvement = £15.65 (95% CI = £12.86: £18.44)
5. Sample mean WTP for *Small* improvement = £9.88 (95% CI = £7.95: £11.81)

While the change from political to economic jurisdiction substantially alters aggregation estimates, the final column of results shows that, at least in this case, an even greater source of error arises from reliance upon sample means within the aggregation process. As shown in Table 2, WTP values decline significantly across space such that, unless samples are fully representative of the underlying population mean values can be poor indicators of value for that population. Given that, ahead of any valuation survey, we are unlikely to know the extent of the economic jurisdiction, ensuring sample representativeness of this *a-priori* uncertain area can be a difficult if not impossible matter to assess. Application of the valuation function allows us to estimate how household WTP varies across the economic jurisdiction, here calculating values for each census output area, taking into account its distance from the site and those area characteristics included within the model (other variables being held at their mean values). Resulting values, shown in the final column of Table 3 are, we contend superior to those given elsewhere in the table as they are both based upon the economic jurisdiction and best capture the variability of values across that area. Comparison with other estimates is revealing. In particular when compared with the political jurisdiction and sample mean aggregation approach used in the EA River Kennet study we see that the latter are more than 16 times too high.

5.   Conclusion

This paper has considered some of the various factors which can influence the calculation of aggregate WTP estimates. Our analysis confirms the findings of Smith (1993) and Loomis (2000) that the choice of whether to aggregate across a politically defined or economic jurisdiction can have a very substantial impact upon estimates of aggregate value. Similarly we have shown that the use of simple approaches such as aggregation via sample means can severely bias such estimates, and is very likely to occur given that the survey analyst is very unlikely to have prior knowledge of the correct area over which to aggregate. As an alternative to such over-simplified approaches we have argued for the use of a spatially sensitive valuation function, explicitly incorporating sample self-selection and expected distance decay in values to both define the limits of the economic jurisdiction and investigate how values vary within that area.

In considering expectations of distance decay, we note that, while a mixed sample of users and non-users is liable to exhibit distance decay in mean household values, patterns of decay for non-users are dependent upon the chosen welfare measure. In particular for equivalent

loss scenarios a lack of change in resource quality means that we should not expect distance decay of the values stated by present non-users (except for any effect due to cultural affiliation with a particular resource). Conversely, for compensating surplus scenarios, the postulated increase in resource quality should induce some present non-users nearer to the site to become users, thereby raising their values and resulting in distance decay effects being observed. Table 4 collates results from the literature and supplements those with findings from the present study. As can be seen, this literature appears to bear out these expectations. As such, inspection of distance decay trends amongst users and present non-users could be seen as a further test of the theoretical validity of aggregation exercises.

Table 4
Distance decay in overall and present non-user WTP responses

Welfare measure →		Equivalent loss (WTP to avoid loss: final quality = present quality)		Compensating surplus (WTP for gain: final quality > present quality)	
Study	Good	All responses	Present non-users only	All responses	Present non-users only
Sutherland & Walsh (1985)	Preserving water quality	✓	✗	-	-
Imber et al., (1991)	Preserving Kakadu Conservation Zone	n/r	✗	-	-
Loomis (2000)	Preserving endangered species	✓	n/r	-	-
Norfolk Broads[1]	Preserving wetlands from saline flooding	✓	✗	-	-
Bateman et al (2005a)	Preserve remote mountain lakes	✓	✗	-	-
Pate & Loomis (1997)	Increasing the area of wetlands	-	-	✓	n/r
Pate & Loomis (1997)	Increasing bird numbers	-	-	✓	n/r
Hanley et al (2003)	Improving river flows	-	-	✓	✓
Mouranaka (2004)	Improving forests	-	-	✓	n/r
River Tame study[2]	Improving a river	-	-	✓	✓

Notes:
1. Case study 1 in this paper
2. Case study 2 in this paper
- = not applicable
n/r = not reported
✓ = significant distance decay
✗ = no significant distance decay

Our paper develops an approach to the estimation of spatially sensitive valuation functions for aggregation purposes based upon the spatial analytic power of a GIS. We believe that this represents a useful direction for future research which has the potential to substantially improve methodology in this area. However, we also recognise a number of significant limitations in our present case studies which require attention in future applications.

First, as demonstrated in the contrast between our first and second case study, it is vital that stated preference surveys collect data on non-response and the spatial distribution of such non-respondents[19]. Surveys are inherently liable to self-selection bias with higher value individuals being over-represented in any sample. Our first case study addresses this issue by identifying the location of non-respondents and observing a strong spatial dimension to their distribution. This is then used with lower bound assumptions to substantially adjust aggregate value estimates.

A second issue concerns the need to gather data to allow us to inspect the effect that changes in site quality are likely to have upon the number of visitors to a site. This effect is the driving force behind the significant distance decay within present non-users values under compensating surplus scenarios. Within stated preference studies this could be addressed through direct questioning of respondents regarding their expectations of future visit demand once site quality has been improved. However, this seems liable to error both because people may well be poor judges of future visit rates (in particular focussing illusion may lead to overestimation of future use) and because of possible strategic behaviour if respondents see this as a provision rather than valuation exercise (Carson et al., 1999). An alternative and, we feel, potentially preferable approach to this problem is to adopt revealed rather than stated preference methods. This allows inspection of the quality/visitation relationship through visitation data derived from multiple sites across which both quality and visitation varies.

An exciting possibility for future research might be to combine conventional RUM travel cost methods with the mixed individual/group demand modelling approaches recently pioneered for differentiated marketed goods by Petrin (2002), Berry et al. (2004) and others. Such

---

[19] Note that there may be a confidentiality issue here. While respondents give permission for their details to be incorporated within analyses, this is not the case for non-respondents. However, the use of details such as a non-respondents location is routinely incorporated within analysis of postal surveys and is a basis for self-selection modelling. That in-person surveys should collect such information seems objectively not to involve any further breech of confidentiality. In any sort of modelling there is the need to respect privacy by not reporting exact location details. Typically, such details are of no material interest to the analysis at hand and are not recoverable from reports of modelling exercises. However, the researchers should ensure that such conventions are respected at all times.

approaches would allow both individual and area data to be combined within models of demand in a manner which is both theoretically consistent and may allow us to address a range of wider policy issues such as area level impacts of policy options.

References

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

Bateman, I.J., Cole, M., Georgiou, S. and Hadley, D. (forthcoming). Comparing Contingent Valuation and Contingent Ranking: Valuing the Benefits of Urban River Water Quality Improvements, *Journal of Environmental Management*, in press.

Bateman, I.J., Georgiou, S. and Lake, I. (2005b). The Aggregation of Environmental Benefit Values: A Spatially Sensitive Valuation Function Approach, *CSERGE Working Paper series,* Centre for Social and Economic Research on the Global Environment, University of East Anglia.

Bateman, I.J. and Langford, I.H. (1997). Non-users' willingness to pay for a National Park: an application and critique of the contingent valuation method, *Regional Studies*, 31(6): 571-582.

Bateman, I.J., Langford, I.H. and Nishikawa, N. and Lake, I. (2000). The Axford debate revisited: A case study illustrating different approaches to the aggregation of benefits data, *Journal of Environmental Planning and Management*, 43(2), 291-302.

Bateman, I.J., Cooper, P., Georgiou, S., Navrud, S., Poe, G.L., Ready, R.C., Reira, P., Ryan, M. and Vossler, C.A. (2005a). Economic valuation of policies for managing acidity in remote mountain lakes: Examining validity through scope sensitivity testing, *Aquatic Sciences*, 67(3), 274-291.

Bishop, R.C. and Boyle, K.J. (1985). The economic value of Illinois Beach State Nature Preserve, *Final Report to the Illinois Department of Conservation*, HBRS Inc., Madison, WI.

Berry, S., J. Levinsohn and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market, *Journal of Political Economy*, 112(1), 68-105.

Carson, R.T., Groves, T. and Machina, M.J., (1999). Incentive and informational properties of preference questions, *Plenary Address, Ninth Annual Conference of the European Association of Environmental and Resource Economists (EAERE),* Oslo, Norway, June.

Case, A. C. (1991). Spatial patterns in household demand, *Econometrica*, 59:953-965.

Cornes, R. and Sandler, T. (1996). *The theory of externalities, public goods and club goods*, 2nd ed, Cambridge University Press, N.Y.

Deutsch, C.V. and Journel, A.G. (1992). GSLIB: Geostatistical Software Library and User's Guide. New York, Oxford Press.

ENDS (1998). Water abstraction decision deals savage blow to cost-befit analysis, *ENDS*, 278:16-18.

FWR (1996). Benefits assessments manual for surface water quality improvements, Foundation for Water Research, Marlow.

Georgiou, S. Bateman, I., Cole, M., and Hadley, D. (2000). Contingent ranking and valuation of river water quality improvements: testing for scope sensitivity, ordering and distance decay effects, *CSERGE working paper GEC 2000-18*, Centre for Social and Economic Research on the Global Environment, University of East Anglia and University College London, Norwich, UK.

Greene, W.H. (1990). *Econometric Analysis*, Macmillan, New York.

Hanley, N., Schläpfer, 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.

Hanemann, M. and Kanninen, B. (1997). The Statistical Analysis of Discrete Response Data. In: I. Bateman and K. Willis (eds.), *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EC and Developing Countries*, Oxford University Press, Oxford.

Heckman, J. (1974). Shadow Prices, Market Wages, and Labor Supply, *Econometrica*, 42(4), 679-94.

Heckman J.J. (1987). Selection Bias and Self-Selection, in P. Newman, M. Milgate and J. Eatwell (eds.), *The New Palgrave - A Dictionary of Economics*, Macmillan.

Imber, D. Stevenson, G. and Wilks, L (1991). *A contingent valuation survey of the Kakadu Conservation Zone*, Research Paper No.3, Resource Assessment Commission, Canberra.

Isaaks and Srivastava (1989). An Introduction to Applied Geostatistics, Oxford University Press, New York.

Kahneman, D. and Sugden, R. (2005). Experienced utility as a standard of policy evaluation, *Environmental and Resource Economics*, 32 (1): 161-81.

Kerr, G.N. (1996). Probability Distributions for Dichotomous Choice Contingent Valuation, CSERGE Working Paper GEC 96-08, Centre for Social and Economic Research on the Global Environment, University of East Anglia and University College London, Norwich, UK.

Loewenstein, G. and Frederick, S. (1997). Predicting Reactions to Environmental Change. In M. Bazerman, D. Messick, A. Tenbrunsel & K. Wade-Benzoni (Eds.), *Psychological Perspectives on the Environment*, Russell Sage Foundation Press.

Loewenstein, G., & Schkade, D.A. (1999). Wouldn't it be nice? Predicting future feelings. In D. Kahneman, E. Diener & N. Schwarz (Eds.) *Well-being: Foundations of hedonic psychology* (pp. 85 - 105). New York: Russell-Sage.

Loomis, J. B. (2000). Vertically summing public good demand curves: an empirical comparison of economic versus political jurisdictions, *Land Economics*, 76(2), 312-321.

Moran, D. (1999). Benefits transfer and low flow alleviation: what lessons for environmental valuation in the UK? *Journal of Environmental Planning and Management*, 42(3), 425-436.

Morrison, M. (2000). Aggregation Biases in Stated Preference Studies, *Australian Economic Papers, 39(2), 215-230.*

Mouranaka, A. (2004). Spatial economic evaluation of artificial Japanese cedar forest management as a countermeasure for Japanese cedar pollinosis: An analysis using a model of multizonal contingent markets with data from cities, towns and villages in Yamaguchi Prefecture, Japan, *Geographical Review of Japan*, 77 (13), 903-923.

Olsen, M. (1969). The principle of fiscal equivalence: The division of responsibilities among different levels of government, *American Economic Review*, 59, 479-87.

Parsons, G.R. (1991). A note on choice of residential location in travel cost demand models, Land Economics, 67: 360-364.

Pate, J. and Loomis, J.B. (1997). The effect of distance on willingness to pay values: A case study of wetlands and salmon in California, *Ecological Economics*, 20, 199-207.

Petrin, A., (2002). Quantifying the benefits of new products: The case of the minivan, *Journal of Political Economy*, 110(4), 705-729.

Randall, A., DeZoysa, D. and Yu, S. (2001). Ground Water, Surface Water, and Wetlands Valuation in Ohio, in J.C. Bergstrom, K.J. Boyle and G.L. Poe (eds.), *The Economic Value of Water Quality*, Edward Elgar, Cheltenham, UK.

Ready, R.C., and Hu, D. (1995). Statistical Approaches to the Fat Tail Problem for Dichotomous Contingent Valuation, *Land Economics,* 71, 491-9.

Schkade, D.A., & Kahneman, D. (1998). Does living in California make people happy? A focussing illusion in judgments of life satisfaction. *Psychological Science*, 9, 340 - 346.

Smith, V.K. (1993). Nonmarket Valuation of Environmental Resources: An interpretive Appraisal, L*and Economics*, 69(1), 1-26.

Sutherland R.J. and Walsh, R.G. (1985). Effects of distance on the preservation value of water quality, *Land Economics*, 61(3), 281-291.

Swinton, S. M. (2002) Capturing household-level spatial influence in agricultural management using random effects regression, *Agricultural Economics*, 27:371-381.

Willis, K.G. and Garrod, G.D. (1995). The benefits of alleviating low flow rivers, *Water Resources Development*, 11, 243-260.