

**Department of Economics and Marketing  
Discussion Paper No.28**

**The Impact of Journey Origin  
Specification and Other Assumptions  
Upon Travel Cost Estimates of Consumer  
Surplus: A Geographical  
Information Systems Analysis**

*Ian J Bateman, Julii S Brainard  
Guy D Garrod & Andrew A Lovett*

**January 1997**

Department of Economics and Marketing  
PO Box 84  
Lincoln University  
CANTERBURY

Telephone No: (64) (3) 325 2811  
Fax No: (64) (3) 325 3847  
E-mail: [Kawa@lincoln.ac.nz](mailto:Kawa@lincoln.ac.nz)

ISSN 1173-0854  
ISBN 0-877176-05-2

## *Abstract*

This paper presents a simple application of the travel cost method conducted using geographical information system software. This permits analysis of the impact of various assumptions concerning the definition of visitor outset origins and routing to recreation sites. Results suggest that varying these assumptions could lead to substantial impacts upon central estimates of consumer surplus.

*Ian J Bateman, Julii S Brainard and Andrew A Lovett, School of Environmental Sciences, University of East Anglia and Centre for Social and Economic Research on the Global Environment (CSERGE), University of East Anglia and University College, London.*

*Guy D Garrod, Centre for Rural Economy (CRE), Department of Agricultural Economics and Food Marketing, University of Newcastle upon Tyne.*

## **Acknowledgements**

This research was funded primarily by the Economic and Social Research Council (grant no. L320223002) and by the Nature Conservancy Council for England (English Nature: grant no. FIN/NC10/01). Assistance was also provided by the USE/School of Environmental Sciences Research Promotion Fund and CSERGE. This research was completed while Ian Bateman was on study leave at the Department of Resource Management, P.O. Box 56, Lincoln University, Canterbury, New Zealand, to whom he is very grateful.

## *Contents*

List of Tables	(i)
List of Figures (i)	
1. Introduction	1
2. Methodology	2
3. Case Study	3
4. Results	6
5. Conclusions	7
References	10

## *List of Tables*

1. Results of Individual TC Model for Different Specification of Time and Distance in the Travel-Cost Variable 9

## *List of Figures*

1. Comparison of Journey Origins with Geographical and Population Weighted Country Centroids 8

# 1. Introduction

Calculation of accurate travel times and distances supplies the basic information for travel cost (TC) analyses of the value of open access recreational sites. In practice, however, analysts are often forced to adopt simplifying assumptions with regard to these key measurements. For example, rather than using the actual origins from which visitors began their journey, researchers are often only able to use centroids from the areas within which these origins are located. Both Mendelsohn et al. (1992) and Loomis et al. (1995), for instance, use US county centroids as origin locations. This is often a substantial simplification. For example in the Loomis et al. study median County size ranged from 1181 km<sup>2</sup> to 3925 km<sup>2</sup> (calculated from US Census Bureau, 1995) across the various states considered.

The use of single outset origins for such large zones seems undesirable. However, the situation may be exacerbated if the population within a large area is unevenly distributed, for example, if the majority of people in a coastal county live near the sea<sup>1</sup>. Here a simple geographic centroid may significantly differ from one which is weighted by population. Finally a further complication may arise when researchers assume constant road speeds or straight line distances thereby ignoring the extent and quality of the road network which underpins true travel times and distances (see, for example, Rosenthal et al., 1986).

We can speculate upon the possible consequences of these various assumptions for TC estimates of consumer surplus. The impact of using straight lines rather than road distances would seem to be a straightforward reduction in the travel distance and time measures underpinning the travel cost variable. However, it may be that this reduction is not uniform across all visitors in that the journeys of those coming from nearby origins may be relatively more circuitous than those of individuals travelling from more remote origins. Such factors would result in straight line approximations giving biased estimates of consumer surplus. The effect of using geographic rather than population weighted centroid origins is less deterministic and will vary from case to case depending upon the distribution of population within chosen catchment areas. However, it is the choice of the size of the area for each centroid which is perhaps of most interest. When these are relatively small, centroids should provide a good estimate of true journey origin. However, as area size increases, this will only

---

<sup>1</sup> The Loomis et al. (1995) study considers a number of Californian counties in which this may be the case.

remain true if origins are randomly distributed across the centroid. This is unlikely to be the case even if the population is evenly spread across the area (i.e. where geographic and population weighted centroids coincide). A central tenant of the travel cost method is that, *ceteris paribus*, the lower the cost (i.e. the closer an individual lives to a site) the more trips will be made. Therefore, in any such catchment area, more visits will be made from origins nearer to the site than from those further away. This means that the centroid will systematically overstate the travel cost which visitors from that catchment are prepared to bear.

Furthermore, in relative terms, this overstatement will be greater for areas nearer to the site than for those further away. Consider a visitor whose true journey origin is 10km from the site but who lives in an area whose centroid is 20km from the site. Here we have a 100% error due to use of the centroid. However, a second visitor has an origin some 100km from the site which is in a similarly sized area with a centroid some 110km from the site. The absolute error is identical but the relative error is only 10%. This situation will result in a systematic bias to the estimated demand curve relative to the true relationship (based on actual origins). Here at lower travel costs we substantially overpredict the number of visits, i.e. the slope of the function becomes steeper and our consumer surplus estimate is biased upwards. The impact of this effect will be directly related to the size of catchment area adopted, i.e. larger areas should lead to larger (more biased) estimates of consumer surplus.

## 2. Methodology

Discussion with a number of recognised<sup>2</sup> experts in the field of TC research suggests that a principal cause of such simplifying assumptions being adopted is limitations in the software packages used to calculate travel times and distances<sup>3</sup>. In order to avoid the restrictions regarding centroid definition imposed by such off-the-shelf packages a TC methodology was developed employing the spatial analytic flexibility of a geographic information system (GIS)<sup>4</sup>, a software package specifically designed to manipulate, integrate and interrogate spatially referenced data at almost any desired

---

<sup>2</sup> Including Nancy Bockstael (University of Maryland), Nick Hanley (University of Stirling) and Ken Willis (University of Newcastle upon Tyne) to whom we are very grateful.

<sup>3</sup> In the USA a common package is PC Miler (ALK Associates, 1992) while in the UK Routefinder (SIA, 1992) has been used.

<sup>4</sup> Specifically Arc/Info Version 7.0.1 (ESRI, 1994) which directly interfaces with certain versions of the Splus statistical package (Stat Sci., 1993).

resolution, thus making it a very suitable medium for implementing TC studies. Further details regarding the development of our GIS-based TC methodology are presented in Bateman et al. (1996). However, with regard to the present study the methodology allows the following flexibility:

- i. Precise journey origins (accurate to 1km) may be specified<sup>5</sup>;
- ii. Alternatively centroid journey origins may be used with any catchment area being specified, these areas are not confined to existing administrative boundaries and may be user defined;
- iii. Centroids may either be geographic or population weighted;
- iv. Travel distance and travel time may either be calculated using straight lines or by reference to a digital road map. Where the latter approach is used, information on road quality and corresponding road speeds can also be incorporated to provide more accurate measures of travel distance and time<sup>6</sup>.

We now present an application of our methodology utilising a simple TC model<sup>7</sup> to illustrate the impact of each of these factors upon resultant consumer surplus estimates.

### **3. Case Study**

During March and April 1993 a face-to-face survey of visitors was undertaken at Lynford Stag, a typical open-access woodland recreation site, location within Thetford Forest, East Anglia. In total 351 parties of visitors were interviewed with respondents being asked a variety of questions including outset origin. Other questions collected data on a variety of socio-economic, activity,

---

<sup>5</sup> This of course depends upon the resolution of the data collected. Current work examines the feasibility and reliability of a 100m referencing system.

<sup>6</sup> Bateman et al. (1996) discuss this facet of the model at length.

<sup>7</sup> This is a simple, single site, TC study, undertaken so as to demonstrate the magnitude of the measurement effect under consideration. In particular it implicitly assumes that substitute sites are randomly distributed (for discussion of a more detailed model, with the exception of the measurement issues under analysis here, see Bockstael et al., 1991). Given this the absolute magnitude of benefit estimates produced should be treated with some caution, however, it is their relative size which is of interest here.

purpose and attitudinal variables thought likely to influence the individual's trip generation function (TGF).

The Ordnance Survey Gazette of Great Britain (Ordnance Survey, 1987) was consulted to identify 1km grid references for the stated outset origins. As expected these were clustered around the site with roughly 90% of origins within 100km of the site. The 1km origins form the base and most accurate estimates of journey outset from which welfare measures can be calculated. However, to address the issues under consideration we also defined a series of alternative centroid origins based on progressively larger catchment areas. The smallest of these was the ward - the basic reporting unit of the UK Census. This varies in size according to population density with rural wards, generally larger than their urban counterparts<sup>8</sup>. However, wards are typically relatively small areas of between 2-4km across. A larger catchment area was provided by UK district boundaries. These are substantial administrative areas which are generally of the order of the smallest of the US counties considered in the Loomis et al. (1995) study discussed above. Finally, we also used the catchment areas provided by UK counties, which compare with the largest US counties considered by Loomis et al.

We therefore have four resolutions of catchment area: 1km; ward; district; and county. The GIS was then used to calculate both geographic and population weighted centroids for each catchment area at each level of resolution with the exception of the 1km origins where data on population distribution were unavailable so that only geographic centroids could be calculated.

For each origin at each resolution travel costs were calculated first by using road distances (using the information on road availability and quality and road speeds mentioned above) and secondly by using straight line distances.

The various travel cost measures obtained from all these permutations were then entered into a series of trip generation functions. Statistical tests indicated that a natural log dependent variable (lnVISIT) would fit the data best producing a semi-log form typical of those used throughout the US and UK literature. To ensure comparability across results this functional form and the non-travel

---

<sup>8</sup> This does raise the possibility of heteroskedasticity problems, however, these were not central to the research question in hand.

cost explanatory variables<sup>9</sup> (derived through standard exploratory tests) were kept constant across these analyses.

Following theoretical and empirical arguments (Smith and Desvousges, 1986; Balkan and Khan, 1988; Willis and Garrod, 1991a, 1991b) trip generation functions were estimated using limited-dependent variable, maximum likelihood techniques (Maddala, 1983) thereby allowing explicit modelling of the truncation of zero and negative visits. Here we write our general trip generation function as per equation (1):

$$\ln\text{VISIT}_i = \beta X_i + e_i \quad (1)$$

where:  $i$  indexes individuals;  $X_i$  is our vector of independent explanatory variables (as defined previously) with coefficient vector  $\beta$ ; and  $e_i$  are disturbances assumed to be independent, identically distributed  $N(0, \sigma^2)$ . Given this model, the ML estimator is based on the density function of  $\ln\text{VISIT}_i$  which is truncated normal as given in (2):

$$f(\ln\text{VISIT}_i) = \begin{cases} \frac{(1/\sigma)\phi[(\ln\text{VISIT}_i - \beta X_i)/\sigma]}{(1 - \Phi[-\beta X_i/\sigma])} & \text{if } \text{VISIT}_i > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Goodness of fit measures were given by log likelihood values, while household consumer surplus estimates for  $Q$  visits was estimated using equation (3):<sup>10</sup>

$$\text{CS} = \frac{[\ln(Q+1) - Q]}{b} \quad (3)$$

where:  $Q$  = number of visits made per annum;  
 $b$  = coefficient on the travel cost variable.

Standard errors were used to construct 95% confidence intervals for the travel cost coefficient and the upper and lower limits were then used to estimate 95% confidence bounds for consumer surplus.

---

<sup>9</sup> These were as follows: Household size; whether respondent is on holiday; whether respondent is working; whether respondent lives near site; respondents rating of the scenery; whether respondent is a taxpayer; whether respondent is a member of the National Trust; whether the main reason for visit is dog walking. All variables were significant at the 5% level. For further information upon the model see Bateman et al. (1996).

<sup>10</sup> A formal proof of (3) is given in Bateman et al (1995).

The large standard errors generally found in travel-cost models meant that confidence intervals were wide.

## 4. Results

Figure 1 illustrates some of the graphical output which can be produced by the GIS and demonstrates the impact of adopting large catchment areas. Here we can see the 1km outset origins derived from visitors responses together with both the geographical and population weighted centroids obtained when county level catchment areas are defined. Inspection of those counties in the immediate vicinity of the site clearly shows that the majority of visitors set out from origins which are closer to the site than the centroids for their respective catchment areas. This is likely to be the case irrespective of the size or location of the area. However, the relative error caused by this effect is much greater for catchments close to the site than for more distant areas. This systematic bias will result in an overestimate of consumer surplus as discussed previously.

Full results from our analysis are presented in Table 1. Here we evaluate consumer surplus using equation (3) and setting  $Q$  equal to the annual mean number of visits (14.65) per household.

Consider first our analysis of the impact of using straight line as opposed to road based measures of travel cost. Here we can see that the straight line measure consistently produces lower estimates of consumer surplus. This is as expected and simply reflects the underestimate of true travel cost produced by straight line approximations. Nevertheless, the degree of difference, ranging up to 20% is substantial.

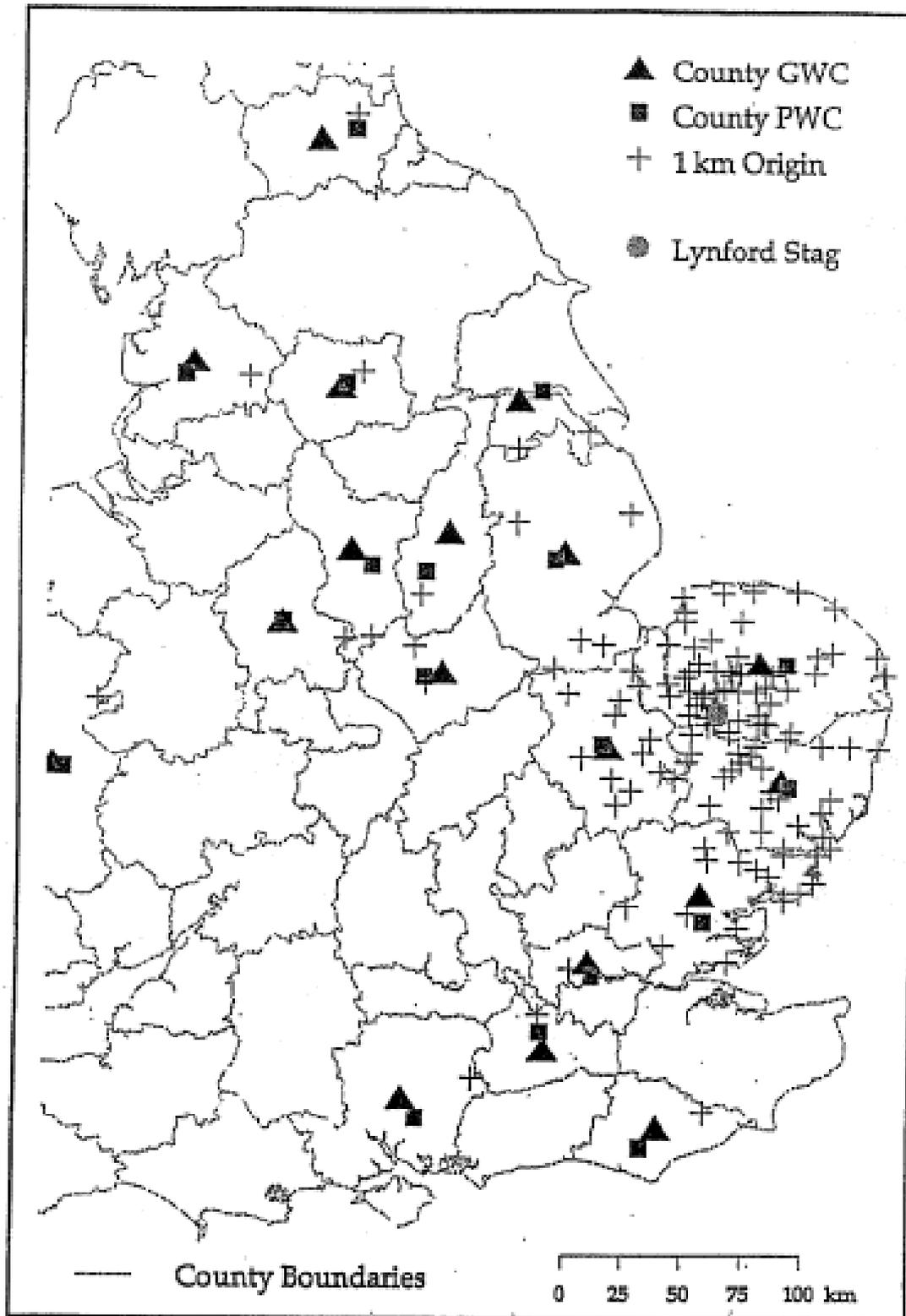
Turning to the impact of using geographical as opposed to population weighted centroids we can see that in this instance there is very little difference in the consumer surplus estimates. However, we would contend that this need not always be the case. Clearly if catchment areas are elongated or population is highly clustered in one area then geographical centroids may be misleading. Given this we feel that consideration of this issue is justified when researchers make choices in TC analyses.

Finally the effect of changing the size of catchment areas can be examined and here we can see the most substantial impacts of the various assumptions that can be made regarding centroids, with results fully in line with our prior expectations. The move from defining outset origin by 1km grid reference point to ward level centroid has virtually no impact upon welfare estimates. This is

unsurprising given that wards often cover just a few square kilometres. However, when we move to using district centroids the biases discussed with respect to Figure 1 begin to become noticeable with travel cost coefficients altering as expected and welfare estimates substantially increased. This effect becomes dominant when we change to our largest county level catchment areas with consumer surplus estimates increasing substantially to the point that our best estimate (the 1km resolution road distance based measure) is less than half of the comparable measure obtained using the county level centroid. It should be noted that the large standard errors on the travel cost coefficient (typical of this and other TC studies) gives rise to wide confidence intervals around welfare measures such that differences are not statistically significant. Nevertheless, given that most decision makers will be interested in best (central) estimates, these results do give some cause for concern. Furthermore, it can also be seen from the log likelihood results reported in Table 1 that the fit of the models declines as we increase catchment area.

## **5. Conclusions**

We have used a GIS-based TC model to examine the impact of certain common simplifying assumptions. Inspection of Table 1 shows that the assumptions regarding travel cost measurement and in particular the specification of catchment areas for journey origin centroids can have substantial impacts upon TC estimates of consumer surplus. The use of large catchment areas can lead to a substantial inflation of central welfare estimates. However, this study also gives a clear indication of best practice for such TC analyses while demonstrating the software capability for achieving this. Use of accurate journey origins and road rather than straight line distances produces improved trip generation functions and more defensible estimates of the welfare benefits of open-access recreational sites.



**Figure 1**  
**Comparison of Journey Origins with Geographical and Population Weighted County Centroids.**

**Table 1**  
**Results of Individual TC Model for Different Specification of Time and Distance in the Travel-Cost Variable**

Model (see key)			TC Coefficient	Standard Error	T-value	Log Likelihood	Annual CS/HH (£)	Annual CS/HH 95% UCL (£)	Annual CS/HH 95% LCL (£)
Area	PWC/GC	RD/SLD							
1km	GC	RD	-0.0281334	0.00914719	-3.08	-455.08	422.97	1166.06	258.34
		SLD	-0.0343321	0.0104276	-3.29	-454.32	346.60	856.45	217.26
Ward	PWC	RD	-0.0280613	0.00925375	-3.03	-455.21	424.06	1199.07	257.57
		SLD	-0.0343684	0.0105582	-3.26	-454.46	346.24	870.21	216.11
	GC	RD	-0.0284978	0.00923917	-3.08	-455.02	417.56	1145.39	255.32
		SLD	-0.0337608	0.0105273	-3.21	-454.61	352.47	906.47	218.76
District	PWC	RD	-0.0235753	0.00906603	-2.60	-456.70	504.75	2049.57	287.81
		SLD	-0.0268869	0.0104250	-2.58	-456.74	442.58	1843.77	251.47
	GC	RD	-0.0236146	0.00913093	-2.59	-456.78	503.91	2081.07	286.66
		SLD	-0.0280577	0.0104679	-2.68	-456.42	424.11	1578.06	244.97
County	PWC	RD	-0.0140468	0.00837676	-1.68	-459.12	847.13	5017.41	390.59
		SLD	-0.0141129	0.00941977	-1.49	-459.46	849.18	2674.14	366.41
	GC	RD	-0.0133633	0.00831721	-1.61	-459.23	890.46	4049.62	401.13
		SLD	-0.0144076	0.00937061	-1.54	-459.35	825.92	3005.85	363.08

**Key**

- Area = Indicates the relevant centroid used in each model. Each centroid is to 1 km resolution accuracy.
- PWC/GW = Indicates whether a population weighted centroid (PWC) or geographical centroid (GC) is used.
- RD/SLD = Indicates whether road distance (RD) or straight line distance (SLD) is used.
- CS/HH = Consumer surplus per household (UCL = upper confidence limit; LCL = lower confidence limit)

Note (all models): Petrol costed at 8p/km; Time costed at 43% of wage rate; Identical functional form and variable list. Sensitivity analysis on the 1km model is presented in Bateman *et al.* (1996).

## References

- ALK Associates (1992) *PC Miler*, ALK Associates, Princeton.
- Balkan, E. and Khan, J.R. (1988) The value of changes in deer hunting quality: a travel-cost approach, *Applied Economics*, 20:533-539.
- Bateman, I.J., Garrod, G.D., Brainard, J.S. and Lovett, A.A. (1995) Using geographical information systems to apply the travel cost method: A study of woodland recreation value, *Centre for Rural Economy Working Paper 16*, Department of Agricultural Economics and Food Marketing, University of Newcastle upon Tyne, pp28.
- Bateman, I.J., Garrod, G.D., Brainard, J.S. and Lovett, A.A. (1996) Measurement, valuation and estimation issues in the travel cost method: a geographical information systems approach, *Journal of Agricultural Economics*, 47(2): 191-205.
- Bockstael, N.E., McConnell, K.E. and Strand, I.E. (1991) Recreation, in Braden, J.B. and Kolstad, C.D. (eds.) *Measuring the Demand for Environmental Quality*, North-Holland, Elsevier Science Publishers, Amsterdam.
- ESRI (1994) *Understanding GIS: The Arc/Info Method*, Environmental Systems Research Institute, Redlands.
- Loomis, J.B., Roach, B., Ward, F. and Ready, R. (1995) Testing transferability of recreation demand models across regions: A study of Corps of Engineers reservoirs, *Water Resources Research*, 31: 721-730.
- Maddala, G.S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.
- Mendelsohn, R., Hof, J., Peterson, G. and Johnson, R. (1992) Measuring recreation values with multiple destination trips, *American Journal of Agricultural Economics*, 24(4): 926-933.
- Ordnance Survey (1987) *Gazetteer of Great Britain*, Macmillan, London.
- Rosenthal, D.H., Donnelly, D.M., Schiffhauer, M.B. and Brink, G.E. (1986) User's guide to RMTCM: software for travel cost analysis, *General Technical Report RM-132*, United States Department of Agriculture: Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.
- SIA (1992) *Routefinder User Documentation*, Service in Information and Analysis, London.
- Smith, V.K. and Desvousges, W.H. (1986) *Measuring Water Quality Benefits*, Kluwer-Nijhoff, Boston.
- StatSci (1993) *S-PLUS Guide to Statistical and Mathematical Analysis, Version 3.2*, Statistical Sciences, Seattle.
- U.S. Census Bureau (1995) U.S. Census Bureau, URL: <http://www.census.gov/>, May 1995.
- Willis, K.G. and Garrod, G.D. (1991a) An individual travel cost method of evaluating forest recreation, *Journal of Agricultural Economics*, 42: 33-42.
- Willis, K.G. and Garrod, G.D. (1991b) Valuing open access recreation on inland waterways: on-site recreation surveys and selection effects, *Regional Studies*, 25: 511-524.