2

Recreation: valuation methods

Introduction

At the heart of cost-benefit analysis (CBA) theory lie two basic principles (Pearce, 1986; Hanley and Spash, 1993): first that, as far as possible, all the costs and benefits arising from a project should be assessed; and, second, that they should be measured using the common unit of money. While these seem common-sense precepts, in application both principles raise highly complex problems. The issue of complete appraisal is, when taken to the extreme, ultimately insoluble in a world ruled by the laws of thermodynamics where, as noted by commentators such as Price (1987a, 2000) and Young (1992), everything affects everything else. For real-world decision-making, practical rules regarding the limits of appraisal are needed. Such rules are the stuff of numerous project appraisal guidelines, for example the Treasury's 'Green Book' (H.M. Treasury, 1991), whereas the research described here focuses on the second principle – of monetary valuation.

In discussing approaches to the monetary evaluation of environmental preferences we can first identify a wider global family of monetary assessment methods (see Figure 2.1). This comprises both the formal 'valuation' (or demand curve) methods discussed below and a quite separate group of *ad hoc* environmental 'pricing' techniques (see the review in Bateman, 1999). In theoretical terms valuation and pricing approaches are quite distinct. Whereas the former are based upon individuals' preferences and yield conventional, neoclassical, welfare measures (hence the term 'valuation methods'), the pricing techniques are much more akin to market-price observations. For example, the shadow project pricing approach uses the costs of hypothetical environmental asset replacement, restoration or transplantation schemes (Buckley, 1989) to yield prices for the environmental costs of a proposed project. While it has been argued that such methods provide useful information for the appraisal of projects, policies or courses of action (Turner *et al.*, 1992), pricing techniques reflect the costs of protecting or providing environmental

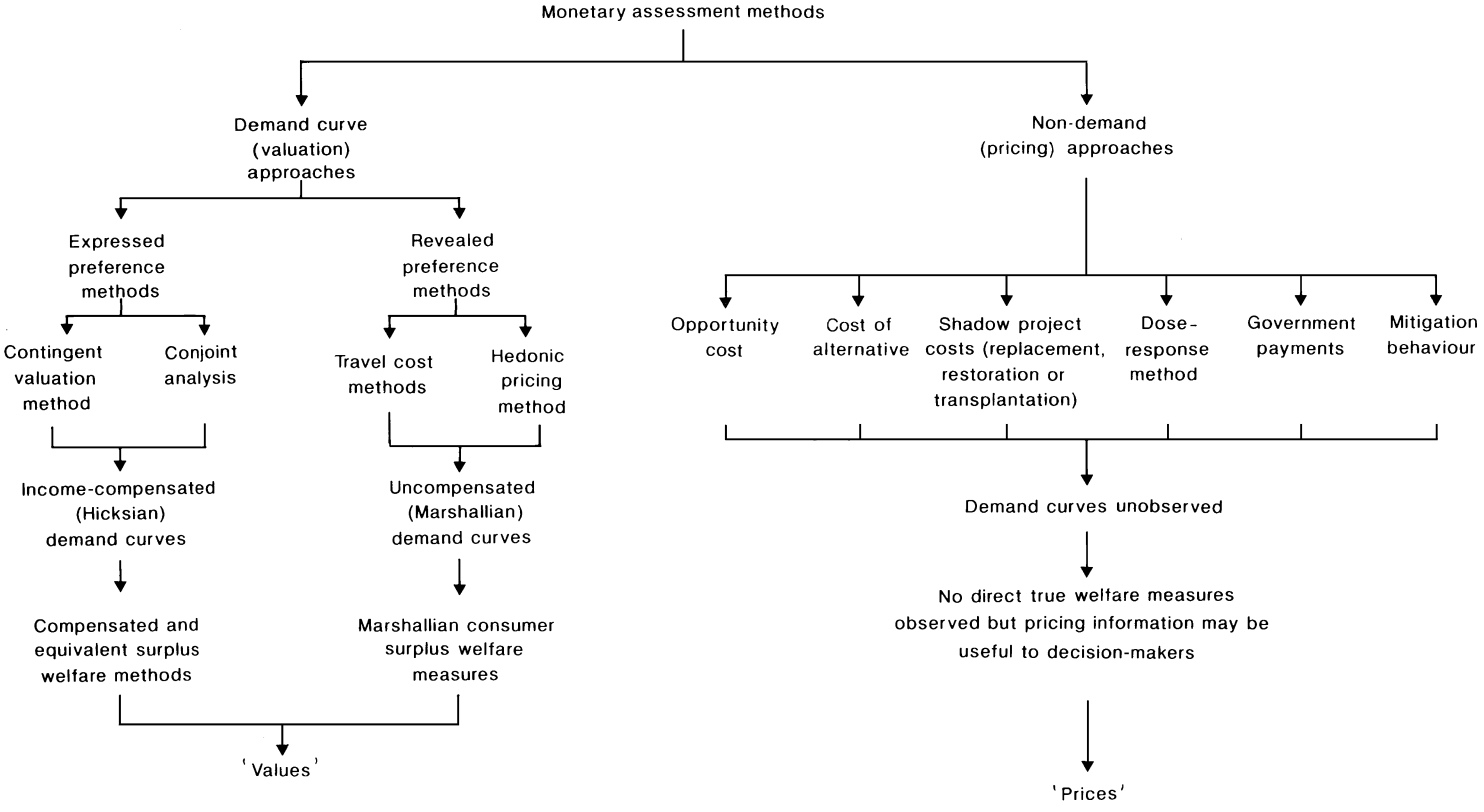

Figure 2.1. Methods for the monetary assessment of non-market and environmental goods. (*Source*: Bateman, 1999.)

assets but not the benefits of doing so. In considering only prices rather than values, decision-makers are in danger of making incorrect choices.[1] Certainly such information is insufficient for adequate CBA appraisals. We therefore reject the use of pricing techniques and turn to consider the more theoretically rigorous valuation methods.

The valuation methods all ultimately rely upon individual preferences. However, within this genre two distinct categories of approach can be defined: methods based upon preferences which are revealed through purchases by individuals of market-priced allied goods; and methods which rely upon expressed preferences elicited through questionnaire surveys. Both of these variants provide measures of value which are valid according to economic theory. However, the same theory shows that these measures need not be identical even when the same change in provision of a non-market good is considered (further discussion of this issue is provided in any basic microeconomics text, for example Laidler and Estrin, 1989; a simple overview is given here).

Revealed preference techniques typically cannot be applied directly to the valuation of environmental goods because of the lack of an observable market price. One solution is to investigate a surrogate market and this approach is adopted by the travel cost (TC) method. Here the costs of a visit to a recreation site are calculated as some combination of any entry charge (typically zero for UK forests), travel expenditure (e.g. petrol costs) and the opportunity cost of travel time (i.e. the value of the time devoted to travelling to the site; this might be wages forgone or the lost opportunity to enjoy some other activity during that time).[2] By comparing these travel costs with the number of visits, we observe that as costs increase (e.g. the further an individual has to travel to a wood), fewer visits are made. This negative relationship maps out a 'demand curve', the area under which provides an estimate of the value of visits to the site which is known as the 'consumer surplus'.[3] While this is a useful measure it is in fact the sum of two components: the substitution effect (which measures the increased consumption of any good when its price falls) and the income effect (which shows the change in consumption due to the increase in purchasing power or 'real' income which occurs when the price of a good falls). While the substitution effect is positive[4] for a reduction in travel costs, the income

[1] As an interesting example of how pricing methods may give little practical guidance to a decision, Medley (1992) refers to the Department of Transport's pricing of a motorway tunnel to avoid a cutting through the Twyford Down Site of Special Scientific Interest in Hampshire. At £70 million this was considered too expensive and abandoned without any appraisal of the benefits of such an alternative being undertaken.

[2] Brief discussion of how these travel costs are estimated is provided subsequently.

[3] In essence the reader can think of the consumer surplus value being estimated as the sum of what the individual visitor would pay, if required, for each of the visits to a woodland. In an attempt to widen readership we have avoided various technicalities in this and subsequent descriptions. References to further reading are provided below.

[4] Strictly speaking this effect is non-negative rather than absolutely positive.

Table 2.1. *Welfare change measures obtained from expressed preference measures*

	Change in provision	
	Gain	Loss
WTP measure	WTP to ensure that the proposed gain occurs	WTP to avoid the proposed loss occurring
WTA measure	WTA compensation if proposed gain does not occur	WTA compensation if proposed loss does occur

effect of such a cost reduction may be either positive or negative depending upon how the individual varies their consumption of the good as the purchasing power of their income changes. Given this uncertainty, consumer surplus might provide an imperfect estimate of the 'income-compensated' welfare provided by a given recreational site.

This problem is at least in theory addressed through the application of expressed preference approaches such as the contingent valuation (CV) method. Here respondents are directly asked to state the change in income which would just offset a proposed change in the provision of the good under investigation. Respondents might be asked to consider either a gain or loss over the present level of provision and in either case be asked questions concerning how much they might be willing to pay (WTP) or willing to accept in compensation (WTA) to just offset the relevant welfare change. Table 2.1 illustrates the four welfare measures so defined.[5]

The income-compensated values estimated by the expressed preference methods can therefore claim some theoretical superiority as welfare measures compared to the consumer surplus estimates provided by revealed preference approaches. However, expressed preference methods have been the subject of considerable criticism regarding the ability of respondents to articulate values for complex goods such as those provided by the environment (Kahneman and Knetsch, 1992; Hausman, 1993; Diamond and Hausman, 1994). In practice there is evidence that both TC-based consumer surplus estimates and CV-based WTP values are reasonably similar,[6] and our research uses both methods, as they are, respectively, the most commonly applied revealed and expressed preference techniques for valuing woodland recreation benefits.[7]

[5] For further discussion see Just *et al*. (1982), Johansson (1987) or any similar intermediate microeconomics text. For an empirical comparison of all four measures, see Bateman *et al*. (2000a).

[6] Carson *et al*. (1996) review 83 studies from which 616 comparisons of CV to revealed preference (RP) estimates are drawn, yielding a whole sample mean CV:RP ratio of 0.89 (95 per cent confidence interval = 0.81 to 0.96). This suggests that, while statistically different from each other (and, as we will see subsequently, on occasion strongly dissimilar), revealed and expressed preference measures do on average produce estimates which fall within the same broad range.

[7] For applications of the hedonic pricing revealed preference method to the valuation of woodland landscape amenity, see Garrod and Willis (1992a). In our own recent research we have examined the potential for improving

The remainder of this chapter presents brief reviews of the CV and TC methods, concentrating on areas of particular interest to this study. Given the focus of this research, these reviews are far from exhaustive and are deliberately written in a non-technical and introductory style. For further reading concerning the CV method, see Mitchell and Carson (1989), Bjornstad and Kahn (1996), Bateman and Willis (1999) and Bateman *et al*. (2002), while for the TC method, see Hufschmidt *et al*. (1983), Bockstael *et al*. (1991), Freeman (1993) and Herriges and Kling (1999); an introduction to all non-market valuation techniques is given in Champ *et al*. (forthcoming).

The contingent valuation method

Introduction: applying the CV method

The implementation of a CV study involves a number of distinct stages. In the first, preparatory, stage a 'hypothetical' or 'contingent' market is set up in which individuals are asked how much they are either WTP or WTA in respect of the proposed change in provision of the good under investigation. These questions may be framed using a variety of elicitation methods. In a WTP study the major alternatives are: (i) *open-ended* (OE), in which the respondent is asked 'how much are you willing to pay?', an approach which produces a bid response which is truncated at zero but is otherwise continuous;[8] (ii) *dichotomous choice* (DC), where respondents are asked 'are you willing to pay £*X*?', the amount *X* being systematically varied across the sample to test individuals' responses to different bid levels. This approach produces a discrete bid response variable and may be iterated using higher or lower bid amounts depending upon the respondents' replies to previous amounts;[9] (iii) *iterative bidding* (IB), in which a series of DC-type questions are followed by a final OE question; (iv) *payment card* (PC), in which respondents select their maximum WTP amount from a list of possible sums presented on a card to them.[10]

The respondent also requires information regarding the nature of the good under evaluation, the proposed quantity/quality change in provision of the good, who will pay for and who will use the good and how payment will be collected (the 'payment vehicle', for example higher taxes, entrance fees, donation to a charitable trust, etc.).

hedonic pricing models of landscape and noise disamenity values through the application of GIS techniques (see Lake *et al*., 1998, 2000a,b; Bateman *et al*., 2001a). Expressed preference methods other than CV (such as choice experiments, contingent ranking, etc.; see Champ *et al*., forthcoming) have not to date been widely applied to the study of woodland recreation values. An exception is provided by Hanley *et al*. (1998) who present a choice experiment study of forest landscape values in the UK.

[8] Bateman *et al*. (1995a) provide a comparison of OE, DC and IB formats.

[9] See, for example, Hanemann *et al*. (1991); Langford *et al*. (1996); Bateman *et al*. (2001b).

[10] See, for example, Rowe *et al*. (1996).

With the questionnaire complete, the process moves to the survey stage to obtain responses. In so doing the relevant population of either users or non-users, or a mix of the two, must be determined. User surveys may be conducted either on or off site while non-user surveys are restricted to the latter locations. In both cases either face-to-face or mail/telephone surveys may be used, each of which has its own merits and drawbacks.

Once responses have been collected, data analysis can commence. This has the dual objective of both obtaining the required welfare measures and assessing the validity of responses. Validation testing is complex and multifaceted (see discussions in Mitchell and Carson, 1989 and Bateman *et al.*, 2002); typically, however, considerable emphasis is placed upon the consistency of responses with theoretical expectations, this being assessed through the estimation of bid curves linking valuation responses to the characteristics of respondents (e.g. their income, use of the good, etc.), and upon assessing the extent to which CV estimates converge with those obtained by other valuation methods.

The final stage of the study is to derive aggregate welfare measures by linking sample responses to the relevant underlying population.[11] Providing that validity tests are satisfactory, these aggregate measures may then be incorporated within project appraisals.

Focal methodological issues

We now concentrate on topics which are central to our woodland recreation work.[12] The general issue under consideration here is the extent to which design issues affect elicited values. However, to set this in context, we begin by considering the process by which individuals form stated responses to CV questions. This discussion allows a consideration of the impacts which choices regarding survey design may have upon valuation responses. Areas highlighted for subsequent research include the effect of varying the elicitation method and changing the payment vehicle, the impact of asking respondents to consider budget constraints and the effect of varying the order of questions within a survey instrument. Some of these issues are tackled through non-woodland applications, results from which are presented in this chapter. Findings from woodland studies are presented in the following chapter.

[11] This is rarely as straightforward as it may appear. See discussions in Bateman *et al.* (2000b, 2002).

[12] This approach precludes discussion of a number of CV issues which we address in other contexts, including the ability of respondents to distinguish adequately between a conglomerate good (e.g. all natural areas) and its constituent parts (just one of those areas) (see Bateman *et al.*, 1997a) and the role of 'reference points' of prior provision in influencing the commonly observed asymmetry between WTP and WTA measures of the same change in provision (see Bateman *et al.*, 1997b, 2000a). This is of course far from an exhaustive list of current CV issues, for which the interested reader should consult the literature cited previously.

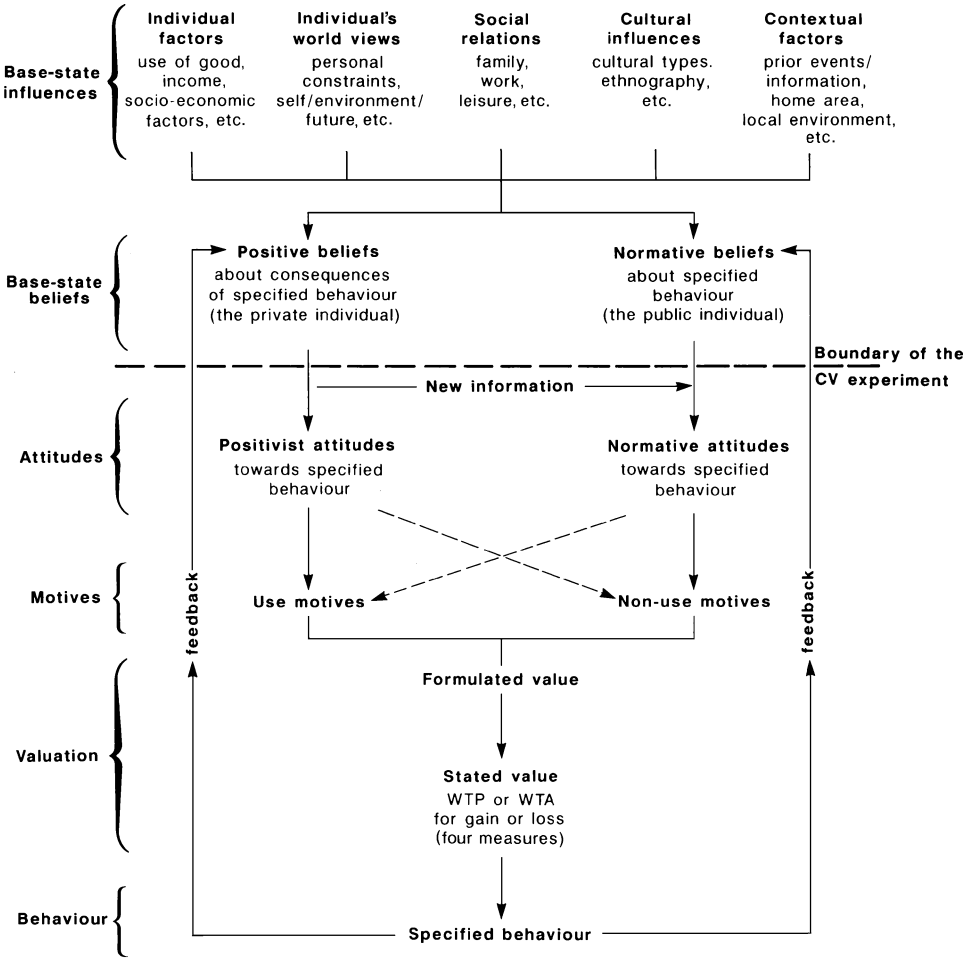

Figure 2.2. The value formation process.

The valuation process and its influences

Consideration of the process through which respondents derive valuation responses to CV questions can be traced back to the beliefs–attitudes–behaviour models proposed by Fishbein and Ajzen (1975) and Ajzen and Fishbein (1977). Recent research has suggested that this process may be highly complex, reaching beyond the somewhat simple models of self-interested rationality underpinning much economic theory. Figure 2.2 draws upon a number of sources to summarise recent thinking in this area.[13]

[13] This section draws upon a variety of sources including Fishbein and Ajzen (1975); Hoehn and Randall (1987); Brown and Slovic (1988); Mitchell and Carson (1989); Dake (1991); Harris and Brown (1992); Bateman and Turner (1993); Schkade and Payne (1994); Marris *et al.* (1996); Hanemann (1999).

The model presented in Figure 2.2 emphasises the pre-survey base-state as a vital determinant in the valuation process irrespective of the good under investigation. A variety of base-state influences are identified. These include the individual-level factors emphasised by traditional economic models of preferences, as well as world views (for example, whether individuals see the relationship between the economy and the environment as ultimately benign or degrading[14]), social factors (such as work and family influences), cultural influences (such as the typologies identified in recent empirical studies: Dake, 1991; Sjöberg, 1995; Marris *et al*., 1996; Langford *et al*., 2000) and contextual factors which may distinguish otherwise identical changes in provision of public goods. These elements combine to yield the base-state positive and normative beliefs which any individual brings to a valuation experiment.[15]

It is tempting to see the schism between so-called individual and citizen preferences (Blamey, 1995, 1996) as being reflected in differences between positive and normative beliefs and there is some evidence to support such a definition (Peterson *et al*., 1996). However, the complex and uncertain nature of this argument prohibits us from taking the matter further. Rather we can see these beliefs as the base-line points of reference from which the individual enters the CV experiment. Here the respondent is presented with new information which will be used to update the belief set. These beliefs will then form the individual's attitudes and norms concerning behaviour. Information, beliefs and attitudes all subsequently feed into motivation. It is arguable that non-use (existence and bequest) values arise from non-use motives such as altruism (Randall, 1987) drawing upon normative beliefs, whereas use values arise from positivist beliefs and attitudes. However, while these are likely to be the main routes of influence, we can also imagine norms concerning instrumental goods and positivist ideas concerning non-use values.

These use and non-use motives combine and are expressed as the WTP sum within the CV valuation process. This statement of value and the CV experience itself then feed back either via behaviour (an actual payment) or, more usually, directly into the individual's positive and normative beliefs, such that values for the same good may change if a CV study is iteratively repeated using the same respondents (see Coursey *et al*., 1986). As a simple investigation of the impact of use and non-use motives upon stated values it was decided that an initial and preliminary objective of our empirical woodland research would be to examine variations in WTP values between users and non-users of woodland recreation as well as to examine the WTA compensation levels demanded by the potential providers of

[14] These views relate, respectively, to O'Riordan's (1976) technocentric man and ecocentric man.

[15] Interestingly, Spash (1997) argues that individuals may also hold beliefs about the valuation process itself and that these may result in the preferences of those opposed to the valuation process being systematically under-represented.

woodland recreation (farmers) who were together the target of the wider research described in this volume.

The transition from formulated to stated value is the subject of theoretical analysis by Hoehn and Randall (1987) and Carson *et al*. (1999). Here the CV respondent is seen as undertaking a two-stage task of (i) value formulation and (ii) value statement. In moving from formulated to stated value the respondent has the opportunity to engage in a variety of strategic behaviours including both understatement and overstatement of formulated WTP. Hoehn and Randall and Carson *et al*. see these various strategies as being chosen according to the elicitation method being used (OE, DC, etc.) and so we consider this issue in some detail.

Elicitation effects

Different elicitation methods may either be neutral (i.e. they have no impact upon stated WTP) or lead to either under- or overstatement of values. Reasons for such effects are diverse and there is considerable debate regarding the impact which the differing strategic incentives and psychological effects of each elicitation method have upon stated values and the consequent validity of those values.[16] If methods are neutral in their effect, the elicitation issue can be ignored. However, both theory and empirical investigation suggest this is not generally the case and so we begin this consideration of potential bias by first considering issues surrounding value understatement, then overstatement, after which empirical evidence is examined and some conclusions drawn. For convenience we shall consider WTP measures throughout the following discussion.

Understatement of WTP

If an individual feels that a good will be provided irrespective of his response to a WTP question, or that the payments of others will be sufficient to secure provision, then, given the ability to freely vary his stated valuation (e.g. in an OE elicitation format), the individual will 'pretend to have less interest in a given collective activity than he really has' (Samuelson, 1954) and will understate his WTP for that good, i.e. he will 'free-ride' (Marwell and Ames, 1981; Brubaker, 1982). A similar result will be obtained when respondents feel that actual payments will (or should) be related to cost shares rather than to WTP (Hoehn and Randall, 1987). Here respondents will state the expected cost, if this is less than WTP, or zero otherwise.

Mitchell and Carson (1989) review a variety of studies for priced goods in which hypothetical OE bids were subsequently compared to actual prices paid. These studies indicated that, where a relatively weak free-rider incentive existed (e.g. where respondents were informed that a group threshold WTP was required in order to

[16] See, for example, Carson *et al*. (1999) and Bateman *et al*. (2001b, 2002).

secure provision), then a reasonably close correspondence between hypothetical and actual payments was found (OE bids were between 74 per cent and 96 per cent of actual payments in the studies cited). However, where a strong free-riding incentive existed (e.g. by guaranteeing provision as long as respondents stated some non-zero sum), divergence was consistently greater (OE bids being 61 per cent to 71 per cent of actual payments).

Overstatement of WTP

Bateman *et al*. (1995a) identify five factors which may induce a respondent to overstate WTP in a CV experiment, each of which we discuss below:

(i) Strategic overbidding (all elicitation formats)
(ii) 'Good respondents' (all elicitation formats)
(iii) Upward rounding (DC formats)
(iv) Anchoring (DC formats)
(v) Starting point effects (IB formats).

(i) Strategic overbidding. In an important empirical paper, Bohm (1972) argues that, contrary to the prediction of free-riding, respondents may overstate their WTP in hypothetical markets. Such 'strategic overbidding' may occur where respondents feel that the amount they will actually have to pay will be related to some sample measure, such as mean WTP, rather than their own statements. In such a case, if formulated WTP exceeds expected mean WTP, the respondent may inflate stated WTP (up to the expected mean) in an effort to improve the probability of provision.

Carson *et al*. (1999) extend this theoretical analysis in the context of OE format responses noting that in a case where there is uncertainty over the provision of a good, individuals have a strategic incentive to overstate their willingness to contribute to subsequent costs of provision, as the very nature of an OE response tells respondents that these amounts are unlikely to bind them subsequently. Empirical support for such a model is provided by Foster *et al*. (1997) who compare CV responses to actual donations for public goods, in this case the preservation of various UK bird habitats. This study found that while OE CV bids were on average not significantly different from those which could be expected in the real payment context, individuals presented with a hypothetical market were significantly less likely to opt out of making a bid than those faced with making real donations.

Contrasting these findings with the evidence for understatement in OE WTP responses (reviewed above) suggests that, in practice, some people respond to the OE elicitation method by free-riding, while others strategically overstate.[17]

[17] Further evidence for such a view can be gleaned from the relatively low degree of fit attained by statistical models of OE CV responses. If individuals are responding in diametrically opposed ways to these questions, then models are inevitably going to struggle to explain such data.

However, meta-analyses of CV studies commonly report that OE-format analyses record significantly lower WTP amounts than do those using other elicitation methods (see, for example, Brouwer *et al*., 1999, and the study presented subsequently in this volume). Therefore, on balance, OE formats appear to result in under- rather than overstatement and may, in the absence of superior measures, be justified as providing conservative estimates of underlying values.

(ii) 'Good respondents'. Orne (1962) points out that the relationship between analyst and respondent is an interactive process with the interviewee seeking clues as to the purpose of the experiment. If this purpose is inadequately conveyed then the respondent may react in one of two ways: either she will not give the questions due consideration or she will attempt to guess the 'correct' answers, i.e. she will try to be a 'good respondent' and give the answers which she feels that the analyst wants. The problem of limited involvement may be assessed by recording and analysing both the numbers of respondents who refuse to take part in the survey and the length of interview. The good respondent problem may be exacerbated where the interviewer is held in high esteem by the respondent (Harris *et al*., 1989), resulting in responses which differ from true WTP. Desvousges *et al*. (1983) found little evidence of such a bias but it should be noted that this study employed professional interviewers, a potential solution to such problems. Tunstall *et al*. (1988) further recommend that interviewers follow the wording of the questionnaire exactly and that respondents be presented with a choice of prepared responses so as to minimise over- or understatement of true evaluations.

In our own empirical work considerable emphasis has been placed upon minimising such sources of bias at the design stage. Experienced practitioners (including several of those cited above) were consulted regarding the construction of questionnaires and execution of surveys.

(iii) Upward rounding. Bateman *et al*. (1993b) argue that respondents in a DC format survey may have an incentive to accept bids which are in excess of true WTP if the difference between the two amounts is relatively small. The deviation caused by such an effect will only operate in an upward manner, i.e. the respondent will not refuse to pay a bid level which is just below their true WTP. However, provided that the respondent believes in the payment obligation (i.e. she does not engage in strategic overbidding) this should be a relatively minor effect.

(iv) Anchoring. Kahneman *et al*. (1982), among others, have argued that respondents faced with an unfamiliar situation (particularly where the good is also poorly described) will interpret the DC bid level to be indicative of the true value of the good in question (Kahneman and Tversky, 1982; Roberts *et al*., 1985; Kahneman, 1986; Harris *et al*., 1989; Green *et al*., 1998). Here the introduction of a specific bid level raises the probability of the respondent accepting that bid. Proponents of this idea argue that this 'anchoring' effect may occur where a respondent has not

previously considered her WTP for a resource (which is likely with regard to public or quasi-public goods) and/or is unclear in her own mind about the true valuation. In such cases the proposed bid level may provide the most readily available point of reference onto which the respondent latches. There is no *a priori* presumption about the direction of such an anchoring effect.[18]

(v) Starting point effects. Several studies have suggested that the use of an initial starting point in iterative bidding (IB) games may significantly influence the final bid; for example, the choice of a low (high) starting point leads to a low (high) mean WTP (see Desvousges *et al*., 1983; Roberts *et al*., 1985; Boyle *et al*., 1985; Navrud, 1989a; Green *et al*., 1990; Green and Tunstall, 1991). While the use of starting points may reduce non-response and variance, commentators argue that such an approach may lead respondents to take cognitive short-cuts to arrive at a decision rather than thinking seriously about their true WTP (Cummings *et al*., 1986; Mitchell and Carson, 1989; Loomis, 1990). It has also been noted that informing respondents as to the construction costs associated with a proposed environmental change may affect resultant bids (Cronin and Herzeg, 1982). One approach to this problem is to allow the respondent to choose a bid from a range shown on a payment card (Rowe *et al*., 1996). However, in some instances the choice of payment range on a card may affect reported WTP bids (for example, if respondents assume that such a range implies information about the 'correct' valuation response; see discussions in Kahneman and Tversky, 1982; Roberts and Thompson, 1983; Kahneman, 1986; Harris *et al*., 1989; Dubourg *et al*., 1997).

In summary, we can see that different elicitation formats may in theory result in either understatement or overstatement of values. We now consider empirical evidence concerning elicitation effects.

Elicitation effects: empirical evidence

Our own studies of elicitation effects have been conducted for both woodland and other resources. Three elicitation methods (OE, DC and IB) were assessed[19] in a CV study examining users' WTP for environmental preservation in the Norfolk Broads, a unique wetland area located in East Anglia, UK (Bateman *et al*., 1995a, 1999b). Following a pilot survey (discussed subsequently in relation to payment vehicle effects), a main survey sample of about 3,000 visitors was obtained through face-to-face on-site interviews. This sample was divided up to permit sufficient

[18] A related problem in DC (and potentially other) formats is the phenomenon of 'yea-saying' or 'nay-saying', whereby the respondent decides *ex ante* to answer positively or negatively irrespective of the actual bid presented (see Kanninen, 1995; Alberini and Carson, 2001).

[19] An iterated exercise, in which an initial DC question was supplemented with two follow-up amounts, was also conducted; see Langford *et al*. (1996) and Bateman *et al*. (2001b). This latter analysis supports the existence of anchoring effects within iterated DC designs.

Table 2.2. *WTP for preservation of the Norfolk Broads using various elicitation methods*

Elicitation method	Mean WTP (£ per annum)	95% confidence interval (£ per annum)
DC	144	75–261
IB	75	70–81
OE	68	60–75

responses to test the various elicitation formats under investigation. WTP results from these formats are summarised in Table 2.2.

Inspection of Table 2.2 shows that the results conform to our prior expectations concerning potential elicitation effects. The overstatement and understatement incentives inherent in the DC and OE formats, respectively, seem to be reflected in the ordering of derived valuation measures with mean DC WTP being more than twice the OE estimate. Of course, it could be that while OE responses were downwardly biased by free-riding (of which some, although not pervasive, evidence was found), DC responses were unbiased. However, statistical analysis of the determinants of WTP amounts provided evidence of a strong link between the starting point in the IB bidding game and the final valuation amount stated. This in turn suggests that DC bid levels might well be interpreted in a similar fashion, i.e. as anchoring points which respondents used as heuristic indicators of the 'correct' valuation of the good under investigation.

This analysis shows that we cannot reject the hypothesis that all elicitation formats are, in one way or another, biased instruments for obtaining WTP values. Which, if any, should be used in our subsequent research on woodland values? The answer to such a question is still unclear and the subject of considerable ongoing research within the CV community world-wide. For the purposes of the research described here we have adopted a simple rule that, wherever possible, we should employ lower-bound assumptions and conservative techniques, thus enhancing the robust nature of derived results.[20] As a consequence, in most of the woodland evaluation research presented in Chapter 3, we adopt an OE elicitation format for our CV studies (with a comparison against a payment card approach in one study) on the grounds that such a choice is likely to produce conservative estimates of WTP.

[20] Such an objective accords well with the practice of H.M. Treasury with regard to its evaluations of non-market woodland recreation values. The guidelines for best practice in CV studies given by the US NOAA Blue Ribbon Panel (Arrow *et al*., 1993) also emphasise conservation design although notably they recommend the use of DC-style referendum elicitation formats because of their desirable theoretical incentive properties.

Table 2.3. *Payment vehicle analysis results*

Payment vehicle	Sample size	Zero WTP (%)	Mean WTP (£)	S.E. mean	Median WTP (£)	Coeff. of variation (%)
DONATE	157	46.5	25.60	3.18	10.00	156
FUND	65	23.1	47.60	17.40	10.00	296
TAX	211	11.8	89.22	9.98	40.00	162

Payment vehicle effects

The idea that the way in which a payment is made is liable to affect an individual's willingness to make that payment (and implicitly the size of payment) is self-evident from the expansion of credit and payment options schemes within modern Western society. That such payment vehicle effects should arise in the purchase of public goods is therefore not surprising and indeed can be seen as evidence that CV respondents act as if hypothetical markets are binding. Payment vehicles can usually be described in terms of two characteristics: collection mechanism and temporal extent.

Considering first the commonly adopted approach of asking survey respondents annual payment questions a number of tax-based and donation-based collection mechanisms appear in the CV literature. The impact of varying these was studied through an earlier survey of visitors to the Norfolk Broads (Bateman *et al*., 1993b). Here a sample of over 400 respondents were presented with one of three payment vehicles: (i) an unspecified charitable donation (the DONATE vehicle); (ii) a payment to a hypothetical charitable fund specifically set up to facilitate flood defence work in the Norfolk Broads (FUND); and (iii) payments made via direct taxation (TAX).

All payment vehicles were applied using an OE WTP elicitation method. Results, which are detailed in Table 2.3, show that both the DONATE and FUND vehicles elicited large numbers of zero WTP bids (46.5 per cent and 23.1 per cent respectively), which contradicts prior expectations regarding a sample of visitors to the Broads who are expected to derive considerable value from the area. In contrast the TAX vehicle produced by far the lowest zero-bid rate (11.8 per cent) and also performed better in terms of bid variability than the FUND vehicle, and about as well as the DONATE vehicle. As no vehicle produced excessive evidence of strategic bidding (large numbers of unreasonably high bids) this was not deemed a problem.

All respondents in the Broads study were asked why they had responded in the way they had. Many of those presented with the FUND and (especially) DONATE vehicles commented that they were unhappy that such a vehicle would not be binding upon all and that they were not confident that payments via such vehicles

would be fully channelled towards preservation work (trust funds were not to be trusted!). Conversely, many of those responding to the TAX vehicle commented that, while they disliked paying extra taxes, they had confidence that such money would be spent efficiently upon any flood defence scheme.

In both the Norfolk Broads and woodlands studies a tax-based vehicle also has the advantage of being the most likely method by which changes in the provision of these quasi-public goods would be funded. This and the advantages outlined above made such a payment vehicle the preferred approach for our subsequent woodland studies. However, a remaining issue concerned the advantages of local taxes relative to national ones, a topic which is discussed in Chapter 3.

Turning to consider variation in the temporal extent of payment vehicles, a number of researchers have experimented with per visit measures for which entrance fees had been used. Several studies have noted differences in implicit values when both per annum taxation and per visit entrance fee measures were obtained for the same good (see, for example, Rowe *et al*., 1980, on landscape values; also Desvousges *et al*., 1983; Brookshire and Coursey, 1987; Navrud, 1989b). Given this result we felt it would be interesting to examine whether such a result was obtained when applying the same design to UK woodlands, and if so why. Consequently, such a comparison was made a further objective of our woodland valuation research.

Questionnaire design impacts: budget constraint and ordering effects

An area of particular interest was the impact which changes in the questionnaire might have upon stated values when the valuation question itself was kept constant. This was assessed through a joint consideration of two design issues: (i) the inclusion or exclusion of a question (prior to the valuation question) asking respondents to calculate their relevant annual recreational budget; (ii) the impact of changing the order in which per annum and per visit valuation questions were presented to respondents.

The relevant economic theory concerning the budget constraint issue is presented in the mental accounting literature (Deaton and Muellbauer, 1980; Tversky and Kahneman, 1981; Kahneman and Tversky, 1984) where total income is initially allocated to various broad categories of expenditure (e.g. housing, food, recreation, etc.), and then, in a second stage, subdivided among the specific items which constitute each category (e.g. the recreation category budget is allocated among forest recreation, water recreation, etc.). Because of the hypothetical nature of the CV market a potential problem may arise if respondents fail to consider all relevant material such as the relevant category budget. Evidence on the impact of explicitly asking respondents to consider income constraints prior to stating WTP sums is mixed (Burness *et al*., 1983; Schulze *et al*., 1983; Willis and Garrod, 1993; Loomis

et al., 1994) and consequently it was decided to make this an objective of subsequent empirical investigation.

Question-ordering effects are considered by Brookshire *et al*. (1981), Tolley and Randall (1983) and Hoevenagel (1990) who note that a good will elicit a higher WTP response if placed at the top of a list of goods to be evaluated than if it is positioned lower in the order. Similar evidence is presented by Kahneman and Knetsch (1992) as part of a series of tests examining the extent to which CV responses are the product of moral satisfaction (i.e. the 'warm glow' of contributing to a good cause) rather than being linked to the characteristics of the good under evaluation. This paper has triggered a wide empirical debate and stimulated theoretical research arguing that, as the consumption of a given good in isolation is not identical to the consumption of the same good as part of a larger set (because the other goods in the set may be substitutes for, or complements to, the good under question), then this phenomenon need not violate economic theory (see Carson *et al*., 1992, 1998; Randall and Hoehn, 1992, 1996; Carson and Mitchell, 1995; Rollins and Lyke, 1998). While not attempting to establish whether variation is due to moral satisfaction, warm glow or theoretically expected effects,[21] we do use ordering effect tests and simpler sensitivity analyses to establish the extent of such variation in values. Furthermore, by combining the budget constraint and question-ordering investigations within a split-sample design, a further analysis of the interactions of these effects could be undertaken to see whether design effects might multiply through a study.

Summary of woodland CV research objectives

CV is a widely applicable and widely applied monetary evaluation method with a consistent basis in economic theory. Given the breadth of the current research debate, we have deliberately focused our empirical investigations on a subset of related issues which together examine the impact of differing designs upon elicited values. The issues addressed in our subsequent woodland CV studies are:

(i) variations in WTP values between users and non-users of woodland recreation
(ii) the WTA compensation levels demanded by farmers for providing woodland recreation opportunities on their land
(iii) elicitation effects (specifically, a comparison between OE and payment card approaches)
(iv) the choice of payment vehicle (in terms of both local versus national tax payment collection mechanisms and annual versus per visit temporal extent of payments)
(v) budget constraint impacts
(vi) question-ordering effects.

[21] Our recent research on the relationship between study design and such effects is reported in Bateman *et al*. (2001c).

Taken together it was intended that these studies would provide some insight into the variability of valuation responses with changes in CV study design. These values could then be compared to those derived from the travel cost method, to which we now turn.

The travel cost method

Introduction

Like the CV method, the travel cost (TC) approach relies upon a survey to gather data. However, whereas a CV survey can, in principle, be applied in almost any situation, a TC survey must involve at least a high proportion of users of the recreational asset in question. Most typically this involves on-site surveys in which a questionnaire is used to collect data on users' place of residence; necessary demographic and attitudinal information; frequency of visit to this and other sites; trip information such as purpose, length, associated costs, etc. From these data, visit costs can be calculated and related, with other relevant factors, to the frequency of visits in a 'trip generation function' (TGF) from which a demand relationship may be established (for details see Freeman, 1993; Champ *et al*., forthcoming).[22]

As discussed by Bockstael *et al.* (1991), the literature can be divided into random utility models (RUM), which examine the probability of a visit to a site given information on all possible visit sites, and more basic TC models which predict visits to a given site by utilising data collected from a survey of visitors at that site. While theoretically more elegant, RUMs require more data than were available to this research and so the more basic approach is employed here.

Two variants of this style of TC model can be identified depending on the definition of the site visits variable (Bateman, 1993). The 'individual travel cost' (ITC) method focuses on the number of site visits made by each visitor over a specific period, say one year. The 'zonal travel cost' (ZTC) method, on the other hand, partitions the entire area from which visitors originate into a set of visitor zones and then defines the dependent variable as the visitor rate (i.e. the number of visits made from a particular zone in a period divided by the population of that zone). In both cases an uncompensated demand curve can be derived and consumer surplus estimates of recreational value obtained.

The UK woodland recreation literature includes examples of both ITC and ZTC applications which we review in Chapter 3. Results obtained from applying the two

[22] Note that Randall (1994) provides a fundamental caution to those who assume that the revealed preference nature of the TC approach gives it automatic ascendancy over expressed preference methods such as CV. He notes that the TC method relies upon researcher-assigned visitation cost estimates rather than observable visit prices and argues that these are inherently subjective, such that the method yields only ordinally measurable welfare estimates. In essence, while the CV method at least presents respondents with hypothetical costs, visitors never see the implicit travel costs used to calculate consumer surplus estimates in TC studies. An empirical assessment of 'Randall's Difficulty' is given by Common *et al*. (1999).

Table 2.4. *ZTC/ITC consumer surplus estimates for six UK forests*

	ZTC		ITC		
Forest	Travel cost coefficient	CS/visitor (£)	Travel cost coefficient	CS/visitor (£)	CS ratio: ZTC/ITC
Brecon	−0.384	2.60	−0.358	1.40	1.86
Buchan	−0.444	2.26	−0.996	0.50	4.52
Cheshire	−0.525	1.91	−1.259	0.40	4.78
Lorne	−0.694	1.44	−0.327	1.53	0.94
New Forest	−0.702	1.43	−0.215	2.32	0.62
Ruthin	−0.396	2.52	−0.386	1.29	1.95

Notes: All coefficients produced via OLS techniques and significant at 5% level; travel cost defined as full running costs; consumer surplus estimates at 1988 prices; $n = 21$ for all forests.
Sources: Garrod and Willis, 1991; Willis and Garrod, 1991a.

variants to the same data have been shown to be substantially different. Table 2.4 illustrates this point with regard to a joint ZTC/ITC study of six UK forest sites. Using the same estimation procedure and cost assumptions throughout,[23] estimates of consumer surplus produced by ZTC ranged from almost 40 per cent smaller to almost five times larger than those produced by ITC. As all the cost coefficients produced by both methods are statistically significant this indicates some serious problems for one or both of these approaches.

One limitation of the ZTC approach is the difficulty associated with the use of an average value as a dependent variable. Employing a zonal visitor rate means that it is impossible to use individual-specific explanatory variables. For example, membership of an environmental or outdoor pursuits association may well be a highly significant predictor of recreational visits. However, information on such individual characteristics cannot be used in the ZTC approach and a constructed zonal average for such variables is likely to be highly inefficient (Brown and Nawas, 1973). Similarly, intrazonal variation is to a considerable degree lost in the ZTC approach, as interzonal average effects dominate in curve-fitting. An extreme case of this occurs where concentric, circular travel time zones are used with no distinction being made within the resultant circles for other variables such as socio-economic or substitute availability measures.[24]

While concentric zones are common in earlier ZTC applications,[25] other approaches to zonal definition are perfectly feasible. The definition of the width

[23] See table notes for details. Cost definition and estimation issues are discussed briefly later.
[24] For an illustration, see Bateman (1993).
[25] Furthermore, zones may be cut off at some finite distance although the outer band may be infinite. Englin and Mendelsohn (1991), in their study of rainforest tourism, analyse visits from all countries.

and number of zones is typically either arbitrary or influenced by the availability of demographic data; for example, Böjo (1985) uses county boundaries to define zones. In effect, each possible definition of zones implies a different aggregation of population and, in practice, almost certainly a different visitor rate. This, in turn, will imply changes in the estimated demand curve and thereby different consumer surplus estimates. Therefore, in reality, it is almost certain that an analyst could respecify zones so as to either inflate or reduce valuation estimates as required. This is an example of the more general phenomenon known as the modifiable areal unit problem (MAUP) (Openshaw, 1984). The extent to which valuation results may alter is uncertain[26] and there is active research into statistical aspects of the MAUP issue (e.g. Batty and Longley, 1996).

A further problem for ZTC, which again does not afflict the ITC method, is that R^2 statistics will always be upwardly biased. This arises as a natural consequence of aggregating individual responses across zones and so reducing the number of curve-fitting points to the number of zones. Consequently the very high R^2 values recorded in many ZTC studies should be treated with extreme caution. Their only real validity is as indicators of which model has relatively higher explanatory power within any particular functional form; their absolute value should be disregarded (and even not reported as it may well be misleading). This criticism does not apply to the ITC for which goodness-of-fit statistics are, in this respect, unbiased.

Given these problems, Brown and Nawas (1973) argue that the ZTC method is inefficient and therefore prefer the use of ITC, a sentiment echoed in early applications by Gum and Martin (1975) and Bowes and Loomis (1980). Indeed the US literature over the past two decades has slowly moved from the use of ZTC to employing ITC. However, the ITC approach is not without problems.

Dobbs (1991) points out that most ITC studies to date have incorrectly estimated consumer surplus, in that they have ignored the inherently discrete nature of the dependent variable. In such cases the integration of a smooth demand function may lead to significant bias in consumer surplus estimates. However, Dobbs develops a programmable approach to the computation of discrete dependent variable benefits which overcomes this problem.

A more fundamental problem with ITC occurs where a high proportion of visitors make only one visit per annum or are first-time visitors (Freeman, 1979; Bowes and Loomis, 1980). In such cases, statistical techniques commonly used in ITC analyses may not have a sufficient spread of observations to make the approach operational. In recent work we have addressed this problem through the application of Poisson distribution models,[27] details of which we present in Chapter 4.

[26] See also Christensen (1985) and Price *et al*. (1986).
[27] However, a Poisson regression will have problems of underdispersion with large numbers of low counts.

In conclusion, the decision to use either zonal or individual TC approaches may have a substantial impact upon the results obtained. While there are a number of methodological problems associated with the application of both, these seem more tractable in the case of the ITC approach, which also has theoretical advantages over ZTC (Bockstael *et al*., 1991). Consequently we adopt the ITC method for use in the valuation studies presented in Chapter 3. However, the recently developed literature on benefits transfer has used the ZTC approach as a readily tractable technique for estimating the numbers of visitors arriving at a given site (Loomis *et al*., 1995). The area visit rates used in the ZTC provide as much information regarding which areas do not yield visitors (e.g. those which are distant from the site, are socio-economically disadvantaged, etc.) as those which do. Therefore the technique yields demand functions which can readily be applied across a study area of which the site distance, socio-economic and other characteristics are known to yield defensible estimates of the number of arrivals expected at a site.

As a consequence, while we use the ITC approach to estimate the value of a recreational trip to a woodland, our model of the number of trips made to woods and the latent demand for trips to potential new woodlands (presented in Chapter 4) owes more to the zonal-based approach of the ZTC method.

Focal methodological issues

As before, our methodological review focuses exclusively upon those issues which are addressed in our subsequent empirical work, with the interested reader being referred to the previously cited literature for wider discussions.

Calculating travel costs

Travel costs are composed of two principal elements: direct travel expenditure (e.g. petrol costs) and the opportunity cost of time.

Travel expenditure

Two issues are pertinent here: measurement and valuation. Accurate measurement is a vital ingredient of valid welfare estimation. However, we have shown elsewhere that a number of questionable simplifications are commonly adopted in the distance calculations underpinning expenditure estimates (Bateman *et al*., 1996a, 1999a; Brainard *et al*., 1999). For example, rather than using the actual point from which a visitor starts their journey, many ITC studies use centre points or 'centroids' of cities (Rosenthal *et al*., 1986) or counties (Mendelsohn *et al*., 1992) as outset origins.[28] This may cause a systematic error given that the very basis of the TC method

[28] Note that we are referring here to a problem with ITC studies, although all ZTC studies, by their very nature, also use zonal outset areas.

is that individuals are mindful of costs in determining their choice of recreation, i.e. we would expect, *ceteris paribus*, that within any area there would be more visitors from outset locations nearer to the study site than from further away. The use of a centroid will partly mask that variation as all visitors within the boundary of the outset area will be assumed to travel from the common central point. This should, on average, lead to an overestimation of the travel costs faced by visitors from a given area as, within that area, most visitors come from locations which are closer to the study site than is the centroid point. The larger the outset area used, the greater we would expect any resultant error to be. A further measurement issue concerns assumptions regarding routing. The use of constant road speeds or straight-line distances ignores the extent and quality of the road network which underpins true travel distances and times (Rosenthal *et al*., 1986). We address all of these aspects of the measurement issue directly in our empirical studies through the application of GIS techniques.

Turning to the valuation issue, a variety of alternative approaches can be identified; for example Bojö (1985) simply refers to the economy class rail fare. However, such a simple approach is less applicable to car travel, where three cost calculation options exist:

(i) petrol costs only (marginal costs)
(ii) full car costs: petrol, insurance, maintenance costs, etc.
(iii) perceived costs as estimated by respondents.

Clearly, using option (ii) will raise visit costs above that of option (i) and ultimately increase consumer surplus estimates, a result confirmed in comparisons of these approaches undertaken by both Hanley and Common (1987) and Willis and Benson (1988). Price (1983) and Christensen (1985) argue that the correct cost measure is that which visitors perceive as relevant to the visit. It may well be that visitors are poor at perceiving daily insurance and maintenance cost equivalents or that they see these as sunk costs which do not enter the TGF, i.e. they only consider the marginal cost of a visit, equating this with marginal utility. As a result of this apparent conflict we adopt a sensitivity analysis approach in our empirical work, testing all three of the above cost definitions.

Time costs

Time enters the visit cost function through the travel time and on-site time variables. However, theoretical analysis (McConnell, 1975, 1999; Freeman, 1979; Wilman, 1980; Johannson, 1987; Shaw and Feather, 1999a, 1999b; Berman and Kim, 1999) shows that the relevant opportunity costs per hour need not be the same for these two items. Furthermore, determination of these opportunity costs raises considerable problems.

Travel time values are particularly difficult to analyse in that, as noted previously, we have no definite *a priori* notion of whether travel time utility is positive or negative. If travel time has positive utility, then using some general travel time cost as a price will overestimate the consumer surplus of a visit. This will be the case for 'meanderers' who gain utility primarily from the journey itself (Cheshire and Stabler, 1976). Bojö (1985) does not include a travel time cost (i.e. implicitly he gives such time an opportunity cost of zero) on the grounds that 80 per cent of survey respondents expressed a positive utility for travel time to the site under analysis. This approach assumes that ignoring residual travel time costs only leads to a minor underestimate of the true consumer surplus.[29] However, this approach is far from standard. Indeed static optimisation of any conventional utility function (subject to income and time constraints) would indicate that the marginal rate of substitution between labour and leisure (i.e. the value of recreational travel time) is equal to the wage rate. However, when individuals are not able completely to vary the number of hours worked the substitution of time for money becomes constrained and the direct relation between the value of time and the wage rate breaks down (Johnson, 1966; McConnell, 1975).

Early applied investigations of the relationship of wages to travel time were undertaken by Cesario (1976) and Cesario and Knetsch (1970, 1976). These papers examined commuters' choice of transport to and from work (and relevant costs) to estimate an implicit value of travel time. Cesario (1976) concludes that, 'on the basis of evidence collected to date, the value of time with respect to nonwork travel is between one quarter and one half of the (individual's) wage rate' and subsequently uses a value of one-third the wage rate to price travel time.[30] However, this analysis only considers commuting time and there is no necessary reason why the marginal utility obtained should be applicable to recreation travel time.

Common (1973) and McConnell and Strand (1981) use an iterative process whereby successive time values are substituted into the TGF, the final choice being determined where the explanatory power (R^2) of the model was maximised. Desvousges *et al*. (1983) apply the value of time results of Cesario (1976), McConnell and Strand (1981) and a full wage rate assumption to an ITC model of individual visitation patterns at twenty-three water recreation sites in the USA. Testing at the 10 per cent confidence level, Desvousges *et al*. (1983) reject the McConnell and Strand (1981) approach, while the Cesario (1976) and full wage assumptions perform equally well, both being rejected in approximately seven of the twenty-three cases. On the basis of these results Smith and Desvousges (1986)

[29] Johansson (1987) points out that if time costs are ignored, then 'the estimated curve will be located inside and be less steep than the "true" one, except possibly for those living very close to the recreation site, since the underestimation of costs increases in relation to distance from the visitor's zone of origin'.

[30] An alternative approach is that of Nelson (1977) who calculates a marginal implicit price of proximity to the central business district with housing data for Washington, D.C., from which he derives a value of time which, when related to wage rates, falls within the Cesario range.

conclude that 'for practical purposes, there is no clear-cut alternative to . . . using the full wage rate as a measure of the opportunity cost. Even though it may overstate the opportunity costs . . . none of the simple adaptations are superior.'

Similar results are obtained in a completely different cultural setting by Whittington *et al*. (1990) in a study of the value of time spent collecting water in Kenya. Here two separate approaches are employed, both of which indicate a value of time approximately equivalent to the wage rate for unskilled labour. However, activities such as collecting water are qualitatively different from those associated with recreation. In their TC study of UK forest recreation, Benson and Willis (1992) employ three wage-rate-based value of time assumptions:

(i) 0, which assumes that visitors would not benefit from some alternative recreation activity
(ii) 25 per cent, the UK Department of Transport's value of non-working time used in CBA assessments of road proposals up to 1987
(iii) 43 per cent, the value of time used by the UK Department of Transport following its review of non-work time in 1987 (Department of Transport, 1987).

While the Cesario approach is, on the surface, theoretically and practically appealing, a deeper analysis of the complexities of the work/leisure relationship highlights some important problems. In a thorough analysis, Bockstael *et al.* (1987) note two major issues: (i) wage rates may vary with work hours; for example, a second job may pay a lower rate than the primary one; (ii) individuals face uneven time constraints, i.e. they may be restricted to working specific hours in particular jobs. As a result the wage rate may be an appropriate measure of time costs for those (at interior solutions) who can fully vary their work hours, but it will be inappropriate for those who cannot (at corner solutions). While Bockstael *et al*. provide a theoretically plausible approach to the valuation problem by incorporating time and income constraints into a utility function, the empirical application of such a technique is problematic. In particular the data requirements of such a model, including information regarding each individual's time constraints, are exacting. For these reasons such complex approaches have not been widely adopted and no published UK study has attempted such an analysis.

Shaw (1992) provides a number of suggestions as to how the value of time problem might be addressed in a practical study. One suggestion is to use CV-type questions to elicit WTP for recreation time,[31] while another is to accept that there is likely to be some rather unclear link with wage rate and therefore to use a sensitivity analysis approach with a wide range of wage fractions.

[31] We have employed a similar approach in a TC study of the Norfolk Broads (unpublished). Here respondents were asked WTP to reduce travel time. However, many gave a zero response indicating that the journey contributed positively to trip utility. Further direct questions confirmed this finding.

As far as the unit value of on-site time is concerned, if the length of time spent on site were a constant for all visits to a particular site, then such costs could effectively be ignored as they would imply only an increase in absolute visit costs but not a change in marginal relationships. Furthermore, in an empirical analysis, Bojö (1985) finds no evidence to refute an assumption of constant on-site time costs, while Bockstael *et al*. (1987) omit on-site time from their empirical analysis because of its potentially ambiguous effect upon demand arising from its inclusion within both the utility function and the constraints.

Summary: treatment of travel costs

The treatment of travel and time costs within the TGF is one of the most crucial issues in operationalising the TC method. The approach we have adopted in this study is as follows.

Measurement. One fundamental issue concerns the measurement of linear and temporal distance. We believe that the use of GIS to analyse digital road networks (incorporating road length, quality and average travel time by individual road segment) in certain of our TC studies considerably enhances the accuracy of measurement compared to that in most other published research.

Travel expenditure. Following the above review we adopt three definitions of monetary travel costs: petrol only; petrol plus standing charges (insurance, depreciation, etc.); and respondents' perceived travel cost.

Time costs. We adopt the suggestion of Shaw (1992) and perform a wage rate sensitivity analysis upon travel time. Four wage rate values are employed: 0 (following the argument of Benson and Willis, 1992); 43 per cent (the UK Department of Transport's value of time); 100 per cent (following the empirical findings of Smith and Desvousges, 1986); and the variable wage rate percentage which provides the best fit to the data (our preferred option). We recognise the limitations of such an approach and that the labour supply method of Bockstael *et al*. (1987) is theoretically superior. However, such an analysis is both complex and demanding in terms of data requirements. Given limited resources our approach should provide a reasonable approximation, while yielding an analysis which is more rigorous than other contemporary UK studies. In line with such research, we have omitted on-site time from the cost function (although such data were collected and analysed), following the argument that this may not significantly affect consumer surplus.[32]

Total travel costs. Given that travel and time costs are both functions of distance, their inclusion together within the TGF is likely to create significant problems of multicollinearity. Accordingly (and for additional reasons reviewed subsequently) we use the common approach seen in studies from Cesario and Knetsch (1970)

[32] Following the analysis of McConnell (1992a), who shows how on-site time may, in certain circumstances, be a significant factor (and proposes a solution to its treatment), we intend to incorporate this into future studies.

to the present day by adding together travel and time costs to produce total visit costs.

Where pertinent we then multiply the total visit cost by the respondent's stated proportion of the total day's enjoyment attributable to the site in question, thereby allowing for that share of the day's utility derived from other sites and the journey itself. This adjusted visit cost is then entered as an explanatory variable within the TGF.

Other explanatory variables

Demand for site visits is likely to be a response to the quality and attributes of a site, yet multicollinearity problems may make the incorporation of numerous such attributes within a single function difficult. Early TC studies tackled this issue through the construction of single-variable quality indices (Ravenscraft and Dwyer, 1978; Talheim, 1978). However, such an index cannot be adequately defined without full knowledge of the functional relationship between demand and site attributes. As this relationship is dictated by individual preferences for different attributes, the creation of a truly representative index is impractical.

Subsequent research has attempted to tackle the problem via multisite studies (Vaughan and Russell, 1982) or by adopting two-stage estimation procedures in which collinear quality attributes are omitted from the first stage (an otherwise conventional TGF) but then used as explanatory variables in a series of second-stage models which predict each of the independent variables used in the initial analysis (Smith and Desvousges, 1986). Such a two-stage procedure is modified for use in our models of agricultural value presented in Chapter 8.

In our own analyses we initially omit consideration of site quality impacts, concentrating instead on the development of improved measures of the principal explanatory variable, travel cost, through use of GIS techniques. However, in our discussion of ongoing work presented in Chapter 4, we detail recently developed models which use these same GIS techniques to incorporate detailed site quality variables into our TC models. A similar approach is taken to the issue of substitute sites, consideration of which is omitted from the models presented in Chapter 3 but included in Chapter 4, where we show how GIS techniques can produce highly detailed variables quantifying accessibility to alternative sites, measures which are readily incorporated into TC models.

Functional form

Analysts are faced with a variety of functional forms under which the TGF can be specified (typically linear, quadratic, semi-log, double-log and Box–Cox). None of these has strong theoretical ascendancy over the others. However, specification of a linear form produces a first derivative which will be a constant and is therefore

theoretically problematic, implying as it does non-diminishing marginal utility for additional trips to a site and thus that the individual cannot decide how many trips to make in total. Log forms may be useful for elasticity estimates and have the advantage of avoiding negative values for the dependent variable.[33] However, the double log form may also be criticised on theoretical grounds as its asymptotic properties imply infinite visits at zero cost, an attribute which is particularly unlikely for demand curves for on-site experience (see Everett, 1979).

An altered functional form (even if it has similar explanatory power) can have a highly significant impact upon the demand curve and resultant consumer surplus estimates. In an early TC study of recreational fishing in Grafham Reservoir (UK), Smith and Kavanagh (1969) found that both semi-log (dependent variable) and double-log functions fitted the data very well ($R^2 = 0.91$ and 0.97 respectively).[34] However, when the resultant demand curves were examined it was found, at a zero admission price, that the semi-log form predicted 54,000 annual visits while the double-log form estimated over 1,052,000, with obvious consequences for consumer surplus estimates. Subsequent re-estimation made little difference to this divergence.

From a statistical viewpoint the most appropriate functional form may be evaluated by examining relative degrees of explanation. However, R^2 tests are strictly non-comparable where the dependent variable changes. A more valid test is to compare visitor rates predicted by the model with observed visitor rates using either a large sample, Wilcoxon signed rank test[35] or a Mann–Whitney U test[36] as appropriate.[37]

Because of its large potential for disturbing consumer surplus estimates, we see the functional form issue as one of the most serious problems affecting the TC approach (as pointed out, it may potentially have far more impact than substitute site or congestion effects). Consequently we have made this a priority issue in our applied research. We investigate a variety of functional forms[38] and estimation procedures (see below) with regard to both the valuation models detailed in Chapter 3 and the prediction of arrival numbers discussed in Chapter 4.

Estimation procedure

Pearce and Markandya (1989) point out that a truncation bias may be introduced where ordinary least squares (OLS) estimation techniques are employed with ITC

[33] See, for example, Ziemer *et al*. (1980); Vaughan *et al*. (1982); Desvousges *et al*. (1983); Smith and Desvousges (1986); Hanley (1989); Benson and Willis (1990).
[34] This was a ZTC study, for which R^2 figures are, as previously discussed, upwardly biased.
[35] Wilcoxon (1945); see Mendenhall *et al*. (1986: p. 806).
[36] Mann and Whitney (1947); see Kazmier and Pohl (1987: p. 496).
[37] Box–Cox approaches to fitting functional forms are arguably superior to standard form approaches.
[38] The use of various functional forms such as log models also partially addresses the issue of heteroscedasticity (Maddala, 1988).

data. The normal error distribution inherent in this technique does not allow for the fact that in such studies the dependent variable can only take positive values. This problem has been tackled through the use of procedures such as maximum likelihood (ML) estimation, where the function can be specified so as to explicitly allow for this truncation.

Empirical studies come to differing conclusions regarding the extent of variance between OLS (truncated) and ML (non-truncated) estimates of consumer surplus. While some find relatively small differences (Balkan and Kahn, 1988), others find that benefit estimates differ substantially (Smith and Desvousges, 1986; Garrod and Willis, 1991). Given this debate we have employed both OLS and ML estimation techniques in our ITC studies although valuation estimates from ML models are preferred and only these are used in the CBA presented at the end of this volume.

Although the theoretical case against OLS methods still applies for ZTC models, in practice such an approach should produce accurate results where the definition of zones is such that all have a substantial positive visitor rate (e.g. when relatively few, often large, zones are used). However, if this is not the case then truncation effects will again make OLS techniques inappropriate (e.g. where many, often small, zones are used, some having zero visit rates). While we do not include a ZTC model in our valuation studies, such an approach is applied to our models of visitor arrivals at unsurveyed sites (Chapter 4), with OLS techniques being adopted in a study with relatively few large zones and a Poisson regression model (allowing for truncation) being implemented in a study with many small zones.

Summary of woodland TC research objectives

The TC method is a potentially useful valuation tool producing uncompensated consumer surplus estimates of use value. While the zonal (ZTC) approach is seen as providing a useful basis for prediction of the number of individuals expected to visit an existing or proposed woodland site (an approach which is developed in the potential demand models presented in Chapter 4), a review of the literature indicates an increasing preference for the use of the ITC variant for valuation purposes. Consequently the TC-based valuation studies presented in Chapter 3 use the ITC method. The above review has identified a number of research objectives for these studies which we summarise as follows:

(i) to investigate the impact of different strategies for measuring travel time and travel distance upon resultant consumer surplus estimates; in particular, utilising the analytical capabilities of a GIS, we examine the impact of improving the resolution of the defined journey outset location and the effect of moving from simple to sophisticated approaches for modelling journey routing

(ii) to conduct a sensitivity analysis across a variety of definitions of travel expenditure and time cost
(iii) to examine the impact of various estimation procedures and functional forms upon resulting consumer surplus estimates.

The research objectives outlined above are in harmony with those defined previously for our CV applications in that all of these analyses essentially examine the impact of varying study design and execution upon derived values. Convergent validity testing via comparison of CV and TC results provides a further research objective for our valuation studies.