3. Stated preference approaches to environmental valuation

This chapter introduces two methods of environmental valuation which rely on the *stated preferences* approach: that is, they rely on the researcher directly asking people about their willingness to pay or willingness to accept compensation for changes in environmental quality. These two methods are contingent valuation, and choice experiments (which are sometimes referred to as choice modelling, or conjoint analysis). In this chapter, we will:

- Provide an overview of the contingent valuation and the choice experiment methods.
- Explain the main problems faced in applying these methods and interpreting their results.
- Present some recent examples of the use of contingent valuation and choice experiments in environmental policy analysis.
- Explain the process of 'benefits transfer'.
- Finally, we briefly review how stated preference methods can be used to value changes in risks in terms of mortality and illness, since such benefits can be important aspects of a CBA applied to environmental legislation.

3.1 THE CONTINGENT VALUATION METHOD

As stated in the previous chapter, the basis for the economic valuation of a change in prices or the availability of a good is to enquire what is the most an individual is willing to pay (WTP) for that change, if it is desirable, or the minimum compensation they are willing to accept (WTA) to forgo the change. Contingent valuation does just this – it asks people what they are WTP for an improvement in environmental quality, or what they are WTA to go without this improvement. Alternatively, people can be asked their maximum WTP to avoid a decrease in environmental quality, or their minimum WTA to put up with this decrease.

The contingent valuation method (CVM) for the valuation of

environmental goods was first used by Davis in a study of hunters in Maine in 1963. However, it was not until the mid-1970s that the method's development began in earnest (Hammack and Brown, 1974; Brookshire et al., 1976; Randall et al., 1974). Since then, the method has become the most widely used (and perhaps most controversial) of all environmental valuation techniques. Much argument surrounded the application of CVM to controversial environmental management and litigation issues such as the protection of the Kakadu National Park in Australia, and the use of the method to estimate damages from a major accident involving the oil tanker *Exxon Valdez* in Alaska in 1989. This latter incident gave rise to the commissioning of an eminent group of economists to apply CVM to measure lost non-use values which Exxon could be sued for in the US courts. As a response, Exxon commissioned another eminent group to publish a critique of the method. The consequence was a US federal government enquiry into the method (the 'NOAA Blue Ribbon Panel'), whose report was a qualified endorsement of the technique (Arrow et al., 1993; Bateman and Willis, 1999). This shows the importance which has become attached to the method (in Exxon's case, a prospective damage claim of $2.8 billion.) Since then, debate has continued over the best way in which to apply CVM, and into how reliable the values it produces can be judged to be. A comprehensive account of the CVM method may be found in Bateman et al. (2002), whilst an early overview of the method was (influentially) provided by Mitchell and Carson (1989). In what follows, we first run through the stages of a CVM, then review some problems in applying CVM and interpreting the results from a CVM survey. Since a very large amount of literature now exists on CVM, we focus on a selection of issues only.

3.1.1 Basics of a CVM Exercise

Most CVM exercises can be split into five stages: (1) setting up the hypothetical market; (2) obtaining bids; (3) estimating mean WTP and/or WTA; (4) aggregating the data; and (5) carrying out validity checks.

Stage 1: the hypothetical market

The first step is to set up a hypothetical market for the environmental good in question. For example, take a policy to improve air quality in a city centre by changing from diesel-powered buses to electric-powered trams, and by converting taxis to run on hydrogen-powered fuel cells. A decision would be made about the relevant population to sample for the CVM – akin to decisions over the relevant population in CBA generally – and a random sample drawn from this population. The description of the 'hypothetical market' needs to include:

- what change in environmental quality is envisaged, and over what time period;
- who would pay for this change, and why;
- how they would pay for this change;
- what would happen if the policy is not introduced (the 'status quo').

In our example, respondents might be told that the local government could engage in such a policy, describe what the policy would consist of, and explain that the policy could only go ahead if extra funds are generated. This sets up a *reason for payment* for the change in environmental quality. How funds will be raised also needs to be described; the *bid vehicle* must be decided upon; for example, through an increase in local property taxes, local income taxes, or a tax on car drivers. In this example, the bid vehicle could be higher local property taxes. The survey instrument (questionnaire) should also describe whether all consumers will pay if the change goes ahead, and how the decision on whether to proceed with the project would be taken.

Good questionnaire design is absolutely vital to a good CVM exercise. The questionnaire should be developed using focus groups drawn from the relevant population, and then pre-tested before the main survey occurs. The information given to respondents about all aspects of the hypothetical market, together with such information as is provided on the good being valued (in this case, an improvement in urban air quality), constitute the 'framing' of the good.

Stage 2: obtaining bids

Once the questionnaire has been designed, the survey is carried out. This can be done by face-to-face interviewing (in people's homes, or at a recreational site), telephone interviewing, via the Internet, or by mail. Telephone interviews are probably the least-preferred method since conveying information about the good may be difficult over the telephone. Internet surveys are growing in popularity. Mail surveys are frequently used, but suffer from potential non-response bias and often from low response rates. Personal, face-to-face interviews offer the most scope for detailed questions and answers, but are relatively costly. Typically, a CVM survey will ask some general questions about environmental attitudes; test for knowledge of the good in question and provide information on the hypothetical scenario; collect WTP/WTA information; ask for socio-economic data on the respondent; and pose some 'de-briefing' questions such as how hard the respondent found the exercise. Box 3.1 contains excerpts from a recent CVM questionnaire by way of illustration.

BOX 3.1 AN EXAMPLE FROM A CONTINGENT VALUATION QUESTIONNAIRE

This questionnaire was part of a study by the consultancy firm Jacobs for the Scottish Executive in 2003, which estimated the benefits of designating Natura 2000 sites in Scotland under the EU Habitats Directive (Jacobs et al., 2004). Most of the benefits were thought to involve non-use values. The survey was conducted in people's houses, using a random sample of Scottish households. After some warm-up questions on attitudes to nature conservation, interviewers asked the following question:

READ OUT and show map of Scotland with Natura 2000 sites shown

Natura 2000 is a new European network of conservation sites containing a representative sample of animals, plants and wildlife habitats of European importance. Most sites have had some form of protection for many years. So far around 300 Natura 2000 sites, excluding marine sites, have been established throughout Scotland.

They cover about 11% of the land in Scotland, and contain some of the most important and unique wildlife habitats in Europe. If the sites are not fully protected, many of the habitats, animals and plants will be damaged and, eventually, lost over time.

Public funds currently available may not be enough to pay for the conservation of the 300 Scottish Natura 2000 sites.

In principle, is your household willing to contribute additional money through your tax bill to ensure that all 300 sites remain fully protected for their wildlife and landscape?

Yes __ No ___ Not sure/don't know __

If the respondent said 'yes' the interviewer then asked:

You have said you would be willing, in principle, to contribute towards the conservation of the Natura 2000 sites throughout Scotland. We are very interested to know *how much* extra you would be willing to pay to ensure their complete protection for the next 25 years.

SHOW PAYMENT CARD

Using this card to help, what is the maximum total amount that your household would be willing to pay *in additional taxes each year* for the next 25 years towards the complete protection of all 300 Natura sites?

Before you answer this question, please bear in mind:

- You will no longer be able to spend this money on other things.
- Other sites in Scotland may still provide some similar wildlife habitats, although not as important.

Individuals can be asked to state their WTP/WTA in a number of ways (in what follows, we focus on WTP alone for simplicity, and since that is what most studies estimate in practice):

- As a *payment card*. A range of values is presented on a card, and the respondent is asked to pick that which most closely matches their WTP. Payment ladders can also be used. Data from such modes can either be treated as continuous information on WTP (that is, if someone ticks the $5 box, we interpret this as showing their maximum WTP is $5) or, more correctly, as interval-type data (so if they tick the $5 box but do not tick the next highest one – say $15 – we know their maximum WTP is at least as big as $5, but smaller than $15). People can also be asked how sure they are that they would pay each amount on the card.
- As an *open-ended question*. Individuals are asked for their maximum WTP with no value being suggested to them.
- As a *single bounded dichotomous choice*: a single payment is suggested, to which respondents either agree or disagree (yes/no). This is rather like voting on the provision of a public good at a fixed price.
- As a *double-bounded dichotomous choice*. Those respondents who say 'no' to the first amount are then asked if they would pay a lower amount, whilst those respondents who say 'yes' to the first amount are asked if they would pay a higher amount. Other variants exist.

Stage 3: estimating WTP

For open-ended responses, calculating mean or median WTP is simple, although researchers must take care to separate out *protest responses* first: these are zero values for WTP given for reasons other than a zero value being placed on the environmental good in question. These might occur because an individual objects on moral grounds to paying for the environmental good, or finds the hypothetical scenario hard to believe, or does not trust the government to actually deliver the environmental improvement on offer. Mean WTP is the relevant value for use in cost–benefit analysis, although authors often focus on median WTP since it is less impacted by extreme values, and since it is meaningful from a political consensus viewpoint (if median WTP for the air quality improvement is £70/household/year, then at least 50 per cent of the population would vote 'yes' to a policy costing £70). Confidence intervals for WTP should also be reported. For payment card designs, mean WTP could be calculated from the maximum value that people say they are WTP. *Bid functions* are usually estimated to investigate the determinants of variations in WTP for open-ended and

payment card data. A bid function is a regression equation which relates WTP to those variables thought likely to influence it. For example, we could take the individual WTP statements from our study and regress them on variables measuring household income, age, health status and whether the respondent has children of school age:

$$WTP = f(\text{Income, Age, Health Status, Kids}) \qquad (3.1)$$

The intention is to see how much of the variation in WTP can be statistically explained, and to see whether variables are related to WTP in a intuitively-consistent manner. In the air pollution example, other things being equal, we might expect WTP to be positively related to household income, and to whether people have children of school age, since children may be thought particularly vulnerable to air pollution. Old people or people of poorer health status might also care more about air quality improvements. Often, though, it is not possible to form a firm prediction about the relationship between WTP and variables we may collect as part of the survey. For payment card designs, estimating equation (3.1) is complicated by the fact that we only know that the respondent's maximum WTP is at least as big as the value they choose on the card, but less than the next highest value (see Haab and McConnell, 2002, for details).

For dichotomous choice (DC) designs (single and double-bounded), the researcher must estimate WTP, since all the respondent reveals is whether she is willing to pay a given amount, not her maximum. Several approaches are available to do this, the most popular being Hanemann's 'utility difference approach', which we now explain (Hanemann, 1984). Full treatments of these issues raised here can be found in Hanemann and Kanninen (1999), and in Bateman et al. (2002). Let us focus on a single-bounded DC design, and assume that Sue derives utility from an environmental good q. Let's assume that Sue has a utility function $U(q, y)$, where y is income. Let us also assume that the researcher cannot observe all of the aspects of this utility function: for example, we may not be able to measure Sue's preferences very well. This idea is known as the *random utility model*, which underlies the DC version of CVM, as well as choice modelling and multiple-site travel cost models. The random utility model can be represented like this:

$$U_j = v(y_j, q) + \varepsilon_j \qquad (3.2)$$

This says that utility is composed of two bits, a deterministic part v and a random part ε, which are 'additively separable'. It is assumptions about the distribution of this random term, and about the functional form of v which will give rise to different models of WTP.

Imagine that Sue, as part of a CVM questionnaire, is offered the option that environmental quality will rise from $q0$ to $q1$, where $q1$ is better than $q0$. Sue is asked whether she will pay £A for this change. She will answer yes with probability:

$$\Pr(yes) = Pr\{v(q1, y - A, e) \geq v(q0, y, e)\} \tag{3.3}$$

and her maximum WTP for this change in q will be her compensating surplus C, defined as:

$$v(q1, y - C, e) = v(q0, y, e) \tag{3.4}$$

which means that (3.3) can be re-written as:

$$\Pr(yes) = \Pr\{C(q0, q1, y, e) \geq A\} \qquad (3.3)'$$

To continue, the researcher must now estimate a statistical model which relates Sue's response, and those of everyone else in the valuation survey, to both the amount A and, typically, people's socio-economic characteristics. How exactly to proceed will depend on a range of factors, notably (as mentioned above), what we assume about the nature of people's utility functions, and what we assume about the distribution of the random part of utility. Haab and McConnell (2002) provide an excellent technical guide to these issues. The simplest case they consider is where the utility function is linear. This implies that the deterministic part of utility looks like this:

$$v_j = \alpha Z_j + \beta(y_j) \tag{3.5}$$

where Z is a range of socio-economic characteristics and y is income for individual j. The deterministic part of the utility function for the hypothetical CVM scenario is given by the difference between utility with the project and income less the offer amount A $(y - A)$, and utility without the project and the original income y. We next need to choose a distribution for the random part of utility: the most common choices are that ε is distributed normally, which leads to a probit model, or logistically, which leads to the logit model. Using the latter assumption, the probability that someone will choose to say 'yes' in the CVM scenario to the offer amount A is:

$$\Pr(yes) = \frac{1}{(1 + \exp(-\alpha Z - \beta A))} \tag{3.6}$$

To estimate this equation, simply create the dependent variable 'response', coded as 1 = yes and 0 = no, then regress this on the socio-

economic variables Z and the offer amount A for each person, using the 'logit' command in a package such as STATA or LIMDEP.

We then want to calculate welfare measures, typically mean and median WTP. How this is done will again depend on what assumptions have been made about the functional form of v, and the distribution of ε. Again, Haab and McConnell (2002) give full details. For the simplest case of the linear utility function, then mean WTP is given by:

$$E(WTP) = \left(\frac{\alpha Z}{\beta}\right) \tag{3.7}$$

Median WTP can be calculated as the value of A that there is a 50–50 chance a randomly selected person would agree to pay.

An alternative way of calculating mean WTP from dichotomous choice CVM data has also emerged, known as the 'non-parametric' or 'distribution-free' approach. This emerged because of a basic problem with the parametric approach set out above, namely that the mean WTP estimate obtained from a given data set depends on what assumptions the researcher makes about the forms of v and ε. Full details on how to use a non-parametric approach to analysing CVM data is given in Haab and McConnell (2002). But we can summarize the main details here of what is referred to as the 'Turnbull method'. First, we observe that if Joe says 'no' to a bid of t_j (we use t instead of A here to make comparison with Haab and McConnell easier), then his maximum WTP must be less than t_j. If he says 'yes', then his WTP must be equal to or greater than this amount. Define F_j as the (unknown) probability that Joe, and anyone like him, will say 'no' to price t_j. It turns out that if we knew F_j, we could calculate mean WTP for our sample. A good estimate of F_j is the proportion of all respondents asked whether they would pay amount t_j who answered 'no'. This can be calculated for each amount asked. We would end up with something like the data in Table 3.1. This CVM data is 'well-behaved', since the value of F_j rises every time the price increases. If the raw data do not have this property, then to apply this non-parametric procedure the analyst has to merge neighbouring price bands together until the merged data do have the property.

A 'lower bound' on WTP can now be calculated, using the formula:

$$E(WTP) = \sum_{j=0}^{M} t_j.(F_j + 1 - F_j) \tag{3.8}$$

This means calculating the difference between the proportion of 'no' responses at a given price, and deducting from it the proportion of 'no' responses at the next lowest price; this gives the quantity $(F_{j+1} - F_j)$,

Table 3.1 Example data from a discrete choice contingent valuation study

Amount offered	Number of 'no' responses	Total number of people made this offer	F_j = (number of 'no' responses / number of people made the offer)
100	98	190	0.51
200	78	144	0.54
300	105	166	0.63
400	113	154	0.73

Table 3.2 Transformed discrete choice data for use of Turnbull Method

Amount offered, t	Number of 'no' responses	Total number of people made this offer	F_j = (number of 'no' responses / number of people made the offer)	$(F_{j+1} - F_j)$
100	98	190	0.51	0.51
200	78	144	0.54	0.03
300	105	166	0.63	0.09
400	113	154	0.73	0.1
400+			1	0.27

and then this is multiplied by the price. These amounts are then summed together. We assume that the probability of saying 'no' to a zero price is zero, and the probability of saying 'no' to some 'choke price' is one. This would give the data shown in Table 3.2.

The lower bound estimate on mean WTP, $E(WTP)$, would then be:

$$E(WTP) = (\$100 * 0.03) = (\$200 * 0.09) + (\$300 * 0.1) + (\$400 * 0.27) \\ = \$159$$

(notice that we ignore the first value for $(F_{j+1} - F_j)$ since this would be multiplied by zero). Haab and McConnell (2002) also give a formula for calculating the variance of WTP, so that a 95 per cent confidence interval for mean WTP can be worked out. For median WTP, we ask: 'at what value of t_j do just more than 50 per cent of people vote no?'. In the above data, this is at a price of \$100. Median WTP will lie between this value and the next highest price.

Whilst the non-parametric approach outlined above has many advantages (it is simple to use, it does not involve making assumptions about the distribution of 'true' WTP), there are also some problems with the method. The main problem is that it is hard to take account of the variables that might be driving WTP. Suppose we think that how long people have lived in an area might well determine how much they are willing to pay to protect a local beauty spot from destruction. The main way of investigating this with the non-parametric approach is to divide the sample into, say, those that have lived in the area more than five years, and those who have lived in the area less than five years, and then to calculate separate means for each group. But you can imagine that this procedure gets rather limiting if one wants to investigate the impacts of many variables on WTP. Another problem is that splitting the sample in this way reduces the number of observations available to calculate each mean, which means that the standard error of our WTP estimate will increase, leading to less precise estimates.

Stage 4: aggregating the data

Aggregation refers to the process whereby the mean bid or bids are converted to a population total value figure. Decisions over aggregation revolve around three issues. First is the choice of the relevant population. This should have been decided when constructing the sampling frame from which the sample was drawn. The aim is to identify either (a) all those whose utility will be significantly affected by the action or (b) (which is the same or a smaller group) all those within a relevant political boundary who will be affected by the action. A decision must be made over the criteria to be used in deciding on who counts in (a) or (b). This group might be the local population, the regional population, the population of Scotland, or the population of the UK, or the whole of Europe. Clearly, where significant non-use values are involved, this population of beneficiaries could be very large. The second issue is moving from the sample mean to a mean for the total population. Several alternatives have been proposed. The sample mean could be multiplied by the number of households in the population, N. However, the sample might be a biased reflection of the relevant population; for instance, it might have higher income levels or show a lower level of educational achievement. If these variables have been included in a bid curve, an estimated population mean bid can be derived by inserting population values for the relevant variables in the bid curve. This number could then be multiplied by N. The third issue is the choice of the time period over which benefits should be aggregated. This will depend on the setting within which the CVM exercise is being performed.

Stage 5: carrying out validity checks
How good are the CVM estimates which the analyst produces? This is clearly an important question from a policy perspective, and in terms of the credibility of environmental valuation. Several 'validity checks' have emerged. These are:

- scope tests;
- convergent validity;
- calibration factors;
- protest rates;
- construct validity.

Scope tests involve examining whether WTP varies significantly with the quantity of q on offer. A simple scope test would be to test the null hypothesis that WTP ($q2$) > WTP ($q1$), where we assume $q2 > q1$. For example, this could mean that WTP to protect all wetlands in one region of France was greater than WTP to protect a single wetland. Scope tests arose as a validity criterion because of a worry that the failure of WTP to show scope sensitivity would imply that a poor description of the environmental change/good in question had been provided, or that people's WTP amounts were largely symbolic donations which could not be interpreted as compensating surplus/equivalent surplus – although sometimes a CVM survey may fail a scope test due to a small sample size. For more discussion, see Heberlein et al. (2005).

Convergent validity is a test for whether WTP for a given environmental quality change estimated using CVM is significantly different from WTP for the same change using some other technique; for instance, comparing CVM and travel costs estimates for a day's fishing (see Chapter 4). This assumes that CVM and, in this instance, travel costs measure the same underlying value, which may not be true when non-use values are concerned.

Calibration factors address a fundamental weakness of CVM: that the values stated are hypothetical commitments, not real ones. A calibration factor is calculated by comparing a WTP value obtained from a CVM survey with a comparable real commitment – obtained, typically, through experimental economics methods (Fox et al., 1998), or occasionally by means of a comparison with actual voting behaviour (Schlapfer et al., 2004). If WTP (CVM) $\gg$ than WTP (real), then doubt is cast on the CVM estimate. We come back to the problem of hypothetical versus real WTP below in section 3.1.2. However, it is hard to calculate calibration factors for many environmental goods since the reason why we undertake CVM is precisely because some aspect of the good defies market valuation; this is especially true for non-use values. Many experimental studies have shown that stated WTP is bigger than actual WTP: could we therefore claim that CVM *always* produces numbers that are 'too big' by some fixed proportion? No: the current view is that the calibration factor varies according to the nature of the good and the nature of the valuation market, and does not lend itself to generalization.

Protest rates are another indicator of the quality of a CVM survey. The protest rate is defined as the percentage of responses which are protest bids (see above): too high a protest rate ('too high' is a subjective matter, but a protest rate of over 40 per cent would raise concerns) implies that there is something wrong with the design of the hypothetical market; for example, people did not find it believable, or found it morally objectionable. One useful exercise can be to try and statistically explain why some individuals protest and others do not. Finally, the worth of an individual CVM study can be assessed using the criterion of *construct validity*. This asks whether WTP varies in a manner which is consistent with theoretical expectations. Usually this question is addressed by estimating a bid function, and seeing whether parameter signs are in accord with a priori expectations (for example do people with more experience or knowledge of the good pay more? Does higher income boost WTP?), and also by considering what percentage of the variation in WTP can be explained statistically. However, for many variables it is hard to decide what the relationship with WTP should be (for example do we expect older people to value forest conservation more than young people? Do we expect locals to value it more than visitors?), whilst there are many different theoretically-consistent assumptions one could make about the nature of the underlying utility function. The construct validity notion is therefore not as useful as it first seems.

3.1.2 Some Problem Areas in Contingent Valuation

Hypothetical market bias
The most simple objection to CVM, as to any stated preference method, is that by asking a hypothetical question, one only receives a hypothetical answer. In other words, what people say they would pay in a CVM study for, say, a reduction in air pollution in their city, is more than they would actually pay if asked to do so. This tendency to overestimate true WTP – if we could observe it – has been called hypothetical market bias. The basic problem with addressing this issue is that we use CVM precisely because the market does not generate a price for many environmental goods – thus it is hard to know what 'true' WTP actually is for, say, an increase in biodiversity. Some authors have used experiments to compare stated with actual values for a range of goods. Harrison and Rustrom's

BOX 3.2 AN EXAMPLE OF A CVM STUDY: REDUCING ECOLOGICAL DAMAGES DUE TO ACID RAIN

Banzhaf et al. (2006) report on a survey carried out to estimate the benefits of reducing acid rain damages in the Adirondacks National Park in the US. Damages from acid rain in the Adirondacks have been important historically in terms of the development of air pollution policy in the US, as they are a well-known example of environmental damages from emissions of SO_2 and NOx. The health benefits of reducing SO_2 and NOx emissions have been widely studied, but no previous study had looked at the economic value of ecological benefits from avoided damages. Non-use values were thought, a priori, to be an important component of the Total Economic Value of reductions in acid rain emissions, thus a stated preference method was chosen by the analysts – in this case, contingent valuation. The sample population was composed of residents of New York State, and most responses were collected through an Internet panel. Considerable effort was made to translate current scientific understanding of how the ecology of the park would benefit from a reduction in acidification into a format which was capable of conveying this effectively to ordinary people: some 31 focus groups were used in survey development.

Two versions of the survey were used, which varied according to the extent of ecological damages under the 'policy off' or status quo scenario. The 'policy on' scenario referred to the use of liming (spreading lime by helicopter) to reduce acidification, rather than the reduction of emissions, since questionnaire pre-testing suggested that people would protest against taxes being used to pay for pollution reductions directly (since 'the polluter should pay'). Higher state taxes over a 10-year period were used as the bid vehicle using a dichotomous choice format. For the baseline case, mean WTP was between \$48–\$107 per annum, depending on how the data was analysed: this implied annual aggregate benefits of between \$336 million and \$1.1 billion. Interestingly, these ecological damage avoidance benefits were about one-third the size of the health benefits estimated for the policy change.

This case study is a good example of a large CVM survey which has been carefully analysed, and which relates to a specific policy question: are the benefits of the damage restoration programme bigger than the costs?

(2005) review of such work shows that 34 of 39 tests revealed hypothetical bias, ranging from 2 to 2600 per cent. Another recent review is provided by Murphy et al. (2005), who find a mean calibration factor of 1.35 (that is, stated values exceed actual monetary values by 35 per cent on average), although they note that for public goods, this hypothetical bias increases. These results reinforce the argument that people tend to overstate their actual WTP when confronted with hypothetical questions. Conversely, in

a study of 616 comparisons of contingent valuation results and estimates derived from actual markets via revealed preference methods, Carson et al. (1996) found that CVM estimates were on average *lower* than revealed preference estimates. List and Gallet (2001) review 174 sets of results from 29 papers, and find that the degree of hypothetical market bias seems to depend on certain characteristics of individual CVM studies, such as how the payment question is asked.

The extent of hypothetical market bias in any particular CVM study is thus hard to predict in any particular study, although a reasonable bet would be that true WTP is less than stated WTP. This is simply because the typical CVM study is 'non-consequential' for respondents: nobody is actually going to ask them to pay the amount they said they would be WTP, and environmental quality is unlikely to change directly as a consequence of their WTP statement. This brings us to a related issue, namely that of *incentive compatibility*. An incentive compatible CVM study would be one where for any respondent, their best bet is to truthfully reveal their exact maximum WTP. No actual CVM undertaken 'in the field' is likely to possess this characteristic. Instead, we can talk about how 'demand revealing' a particular CVM design is – how much of people's true WTP will be revealed by their WTP statement? In fact, this problem of incentive compatibility is not restricted to hypothetical markets, or to CVM. For example, when environmental charities ask for donations to meet a funding target for protecting a threatened habitat, an individual has an incentive to 'free ride' by offering to pay less than the true value. Why? Because if the benefits of the good – here, habitat conservation – are available to everyone regardless of whether they pay or not, then I can get a benefit even though I do not pay for it. This might be particularly true of non-use values for biodiversity or wilderness. For a recent overview of findings on hypothetical market bias (which includes a discussion of the importance of distinguishing between bias at the level of aggregate and individual responses), see Burton et al., 2007.

What can be done about hypothetical market bias? Besides testing for it, which is a rather hard thing to do in many contexts, one suggestion has been simply to tell respondents about the fact that, in a hypothetical survey, people tend to overstate their WTP, and then ask them not to! This is known as 'cheap talk' (Cummings and Taylor, 1999; Aadland and Caplan, 2003). A short version of this, used by Whitehead and Cherry (2007) reads: 'Now please think about the next question (the WTP question) just like it was a real decision. If you signed up for the program you would have *A* dollars less to spend on other things.' The evidence suggests that cheap talk can moderate hypothetical market bias, especially for those with higher WTP values.

Choice of response mode

One issue which has generated many articles in academic journals is which response mode should be used, and how data should be analysed. Open-ended CVM questions have been criticized for being too hard for respondents to complete, and for resulting in high-variance mean WTP distributions. However, the approach in principle tells us exactly what we want to know – the most someone is WTP for an environmental change, or the least they will accept in compensation. Open-ended designs thus continue to be used, although typically only for environmental goods that respondents are familiar with (for example fishing permits). Single-bounded DC formats became almost 'industry standard' following the 1993 National Oceanographic and Atmospheric Administration report into CVM produced by the US government, partly because it was alleged to be incentive-compatible – that is, that it would lead people to reveal their preferences truthfully, partly because it was argued to be more realistic (a fixed price for providing a public good), and because in the US respondents are familiar with voting on local public good issues.

However, single-bounded DC designs turned out to produce systematically-high mean WTP estimates ('yea-saying' being one explanation); required larger sample sizes because they are statistically inefficient; and produce mean WTP estimates which can be very sensitive to statistical assumptions about the functional form of WTP. Partly in response to these weaknesses, the double-bounded DC design was pioneered by Hanemann and Carson, and became widely used in the 1990s. But concerns arose over the effects of the size of the first bid on responses to the second bid (that is, whether both responses came from the same underlying distribution of WTP) (McLeod and Bergland, 1999). A further problem with the double-bounded DC design is that it typically fails to make the *decision* rule clear to respondents: will governments go ahead with a project if enough respondents vote 'yes' to the first amount, or to the second amount asked? Understanding what respondents believe about this would be important to understanding how much of their true WTP they will reveal. It has also been argued that single- and double-bounded DC formats do not encourage respondents to think carefully enough about the value they place on an environmental good, since 'yes' and 'no' are easy answers to give (Fror, 2008).

Alternative mechanisms are thus still widely used. Methods that have become popular include payment cards that allow respondents to say how sure they are they would pay the amount asked, over a series of amounts; and payment ladders which allow people to say the most they are sure they would pay, and the least they are sure they would *not* pay, thus typically identifying a range of uncertainty, given that people may be unsure

BOX 3.3 IS OUR ESTIMATE OF WILLINGNESS TO PAY SENSITIVE TO HOW WE ASK THE QUESTION? SOME EVIDENCE

As we noted above, there is a debate amongst CVM practioners about which format to use for WTP questions. Open-ended (OE), payment card (PC) and dichotomous choice (DC) formats all have advantages and disadvantages. But does it make a difference to our estimates of WTP, and if so, is this proof that hypothetical markets are somehow unreliable? Patricia Champ and Richard Bishop investigate this question using some rather unique data. As they show (2006, Table 1), many previous studies have compared mean WTP for DC, OE and PC formats. A typical finding is that WTP is sensitive to the choice of format, with DC designs usually giving higher WTP values. This sensitivity has been used to criticize contingent valuation, since the argument is that the underlying utility change should be invariant to how we try and measure it. However, Champ and Bishop show that this sensitivity also exists for actual payments for real goods. Their experiment involves customers of a Wisconsin power company being offered the chance to buy their electricity from renewable sources rather than from coal fired power stations. Respondents were told that renewable sources – in this case, wind power – had lower environmental costs than fossil fuel powered electricity, but that wind energy was more expensive. Consumers could thus opt, if they wanted, for more expensive, cleaner electricity. Two designs of the questionnaire were used, one with a DC format and one with a PC format.

Results showed that both the distribution of WTP and its mean value were different according to the format used, with the DC design giving higher WTP estimates. Since this was for real payments for an actual good, the authors concluded that the effect of format on WTP was nothing to do with hypothetical market problems! Rather, they suggest that different designs may convey different information about the good on offer to respondents, in that the payment format contains value 'clues' that cause people to respond differently. As the authors say, 'the bottom line is that, a priori, one elicitation format is not unequivocally better than the others'. All methods have advantages and disadvantages.

of their preferences for some environmental goods (Hanley et al., 2009). Non-parametric means of data analysis have also been introduced to try to get around sensitivities to distributional assumptions within the single- and double-bounded DC designs.

Information provision

An early concern in CVM was the sensitivity of WTP estimates to the amount and nature of information provided to respondents (see the survey in Munro and Hanley, 1999). For example, mean WTP for protecting a not very well known species of wildlife could depend on what people are told about this species as part of the CVM questionnaire process. In a sense, we would want this to be so, since the value of market goods depends on what people know about the characteristics of these goods (for example my maximum WTP for a motorbike will depend on what I can learn about its performance: if I am subsequently told that the reliability of the brand is questionable, my willingness to pay will fall). Yet especially where the analyst is dealing with unfamiliar environmental goods – such as biodiversity – providing adequate information about the good to be valued is crucial if we wish to elicit 'informed' preferences. But how best to do this? And what constitutes 'adequate' information?

One new concept which addresses this question is the 'valuation workshop' technique, as explained in MacMillan et al. (2006), where respondents meet together with 'experts' over a number of occasions, discuss the valuation problem with each other, and take time to think about their preferences. Finally, an interesting new angle on the information story is concerned with what people know about *why* environmental problems occur: there is now some evidence to suggest that people are willing to pay more to cure environmental problems that they believe to be caused by human actions than they are for identical problems due to 'the forces of nature' (Bulte et al., 2005).

Voluntary versus non-voluntary payments

In many cases, the use of a voluntary payment mechanism as the bid vehicle is the most realistic choice in designing a CVM study. For example, if one thinks about an increase in the protection of an endangered bird species in the UK, then asking people their maximum WTP in terms of contributions to an environmental charity which acts to buy up and safeguard this bird's habitat is both realistic and in line with people's experience. However, some researchers have recommended against using voluntary payment mechanisms, since they encourage free-riding. With free-riding, respondents take advantage of the fundamental non-excludability of public goods (see Chapter 2). They do this by stating a maximum WTP which is below their true value, since they know that so long as the good is provided for some, it will be available to them too. Stated WTP, obtained from a CVM exercise, will thus be an underestimate of true value. One way of dealing with this problem is the 'provision point mechanism', whereby respondents are told that a minimum level of aggregate contribution is required for the public good to be supplied at all. This may be reinforced by either a proportional rebate rule (all excess contributions are returned weighted by your WTP), or an extending benefits rule, whereby additional amounts of the public good are provided above the amount that has been set out,

should aggregate contributions exceed the minimum. Poe et al. (2002) show that this type of design can greatly improve the demand-revealing potential of voluntary contribution CV studies, by reducing free-riding. Stated WTP thus moves closer to true WTP.

3.2 THE CHOICE EXPERIMENT METHOD

3.2.1 Introduction

The choice experiment method is one method within a wider group of approaches known as choice modelling or conjoint analysis. The choice experiment method adopts a particular view on how the demand for the environment goods is best pictured, known as the *characteristics theory of value*. This states that the value of, say, a forest is best explained in terms of the characteristics or *attributes* of that forest. Different forests are actually different 'bundles' of attributes, and what people value is these bundles. Moreover, the value of any particular forest then can be broken down into the value of its different attributes. Using observations of people's choices between different bundles of attributes, the researcher can infer (i) which attributes significantly influence their choices; (ii) assuming price or cost is included as one attribute, what they are willing to pay for an increase in any other attribute; (iii) what they would be willing to pay for a policy that changed several attributes simultaneously.

The choice experiment (CE) method is becoming increasingly popular as a tool for estimating and indeed investigating environmental values. Policy makers have seen a powerful set of advantages for the CE method, in terms of being able to measure benefits for a wide range of policy changes. Bateman et al. (2002) give several examples of the use of the method in the policy process. For a very useful guide to the CE method, see Louviere et al. (2000) and Henscher et al. (2005).

3.2.2 How to Carry Out a Choice Experiment

In the choice experiment method, the researcher first of all identifies the main attributes that are relevant for describing the environmental good in question. This is done using focus groups, and by finding out from policy makers and administrators which aspects of the environmental good are likely to be affected by a policy action. For forests, the attributes might include species composition, age, type of felling regime, and the provision of recreational facilities. For a river, the attributes might be in-stream ecological quality, flow rates, and condition of the river banks. For a national park management problem, the attributes could be provision of guided walks, set-aside of conservation areas, traffic management, and management of agricultural areas. If the researcher wants to use the CE to measure economic values, then a price or cost attribute must also be included. For forest recreation, this could be the travel costs of a visit to the site; for river quality, it could be local water and sewerage rates; for a national park it could be a tourist tax. The researcher needs to be sure that the selected attributes are (i) likely to be relevant in terms of the preferences of the population to be surveyed; and (ii) likely to be amenable to change by environmental managers.

Different bundles of these attributes are then assembled, using experimental design principles. Software is available for this task (such as SAS), along with design catalogues. Bundles are then arranged in pairs, and respondents asked to choose between them and some status quo alternative; this is known as a 'choice set'. Typically, each individual might answer 4–8 choice sets. For example, a study by Morrison et al. (2002) looked at the benefits of protecting wetlands in Australia. Each respondent was asked to choose most preferred alternatives amongst pairs of different wetland management options, such as the choice set shown in Table 3.3 (this has been adapted a little from the original):

The questionnaire would be designed, piloted and implemented just like a contingent valuation study, as described in the previous section. Similar requirements exist for the description of the hypothetical market.

Table 3.3 Choice experiment for valuing Australian wetlands

Which option would you prefer that the government went ahead with? A, B or C?

	Management option A	Management option B	Management option C (status quo: no change on present)
Wetland area conserved	1000 ha	800 ha	700 ha
Bird species conserved (number)	40	30	25
Farm jobs protected	15	16	20
Cost to households in terms of increase in local taxes over next 5 years	\$30/hsld	\$15/hsld	\$0/hsld

Source: Adapted from Morrison et al. (2002).

Once questionnaires have been completed, the researcher now has data on which options individuals chose (option A, option B, the status quo), and she can relate these choices to the levels that the attributes took in these options. In this way, choices can be statistically related to attribute levels, including price. The usual statistical model employed is known as the conditional logit model. This means we can write down the probability that an individual *i* chose a particular option like this:

$$P_i(choose\,A) = \frac{\exp(\mu V_{iA})}{\sum_j \exp(\mu V_{iJ})} \tag{3.9}$$

where V is the 'observable' part of utility within a random utility model (as described briefly in section 3.1.1), μ is a 'scale parameter' which relates to the variance of the error component of the random utility model, and J are all the other options the individual could have chosen instead of A. A typical assumption is that V is a linear function of the choice attributes X:

$$V = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_n X_n + \beta_c C \tag{3.10}$$

We see that there are $(n + 1)$ attributes and that for each one, the model estimates a value β which shows the effect on utility of a change in the level of each attribute. Thus β_1 shows the effect of utility of a change in attribute X_1. The model also estimates a parameter β_c, which is the effect of a change (increase or decrease) in the price or cost of the option on the likelihood of choosing that option. Software packages such as *STATA* and *LIMDEP* can be used for this kind of estimation. Now knowing the β values is interesting, since now we know how much utility goes up or down when the attributes increase or decrease (albeit moderated by the scale parameter). These values tell us whether people prefer an increase or a decrease in each attribute; we can also see by looking at the *prob* or t-statistic values from the computer output whether these attributes are statistically significant or not. Box 3.4 shows the output from *LIMDEP* for one choice experiment, and how this is interpreted.

The final steps in a choice experiment are to calculate willingness-to-pay estimate, based on the β values already discussed. The β values show the effect on *utility* of changes in the attributes, but for cost–benefit analysis we need money-metric measures of willingness to pay. For a marginal change in an attribute, this WTP value is typically given by, for attribute X_1:

$$IP_{x_1} = \frac{\beta_{x_1}}{\beta_c} \tag{3.11}$$

BOX 3.4 LIMDEP OUTPUT FROM A CHOICE EXPERIMENT

In this choice experiment of water quality improvements on a rather polluted river, there were four attributes being used: price (PRICE, below), how much of the river was improved (RQ), the change in the number of days when the river smelled bad (ODOUR), and the improvement in ecological conditions (EC). We also collected data on a large number of socio-economic characteristics of respondents, such as age and highest level of education achieved: each socio-economic variable was interacted with the constant (K) to let it enter the model. Data on how far people lived from the river was also obtained (DIST).

```
+----------------------------------------------------------------+
| Discrete choice (multinomial logit) model                      |
| Maximum Likelihood Estimates                                   |
| Model estimated: Mar 06, 2008 at 07:07:04PM.                   |
| Dependent variable                              Choice         |
| Weighting variable                                None         |
| Number of observations                            3059         |
| Iterations completed                                 6         |
| Log likelihood function                      -2594.238         |
| Number of parameters                                16         |
| Info. Criterion: AIC =                         1.70660         |
| Finite Sample: AIC =                           1.70665         |
| Info. Criterion: BIC =                         1.73811         |
| Info. Criterion:HQIC =                         1.71792         |
| R2=1-LogL/LogL* Log-L fncn R-sqrd RsqAdj                       |
| Constants only   -3184.9785      .18548  .18334               |
| Response data are given as ind. choice.                        |
| Number of obs.= 3150, skipped 91 bad obs.                      |
+----------------------------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]
K	-5.46435700	.46862412	-11.660	.0000
RQ	-.00094221	.00513121	-.184	.8543
ODOUR	.00663149	.00665766	.996	.3192
EC	.60183041	.05772591	10.426	.0000
PRICE	-.09344849	.00473387	-19.740	.0000
RECREA	.74035732	.08679536	8.530	.0000
KNOW	.65302930	.14312736	4.563	.0000
DIST	1.17894781	.15266722	7.722	.0000
DIST2	-.09097955	.01760967	-5.166	.0000
AGE	-.13958176	.02960205	-4.715	.0000
EDU	.20341194	.07409420	2.745	.0060

If we look at the results, we can see that neither the RQ or ODOUR attributes had a significant effect on choices, since the prob value for these attributes is bigger than 0.05. But people did care about the improvements in ecological quality and the price of the option. We can also see that how far away people live from the river matters to their choices, but that this relationship is actually quadratic. Finally, we can see that age and education seem to affect people's choices, as did how many recreational visits they made to the river (RECREA) and how well informed they are about water quality in the river (KNOW).

This value for any attribute (other than price!) is called the *implicit price*, or *IP* in equation (3.11). For instance, in Table 3.3 one of the attributes was the number of bird species conserved. Dividing the β value for this attribute by the β value for the tax increase would show the (average) willingness to pay of people in the sample to increase the number of bird species conserved by one. However, often we wish to value multiple changes in attributes. For instance, a new policy on wetlands conservation could alter the area conserved (labelled A below), the numbers of bird species conserved (labelled B) and the provision of recreational trails, labelled R. The price for this would be an increase in local taxes, which are attribute c. The average willingness to pay for this suite of changes in attributes can be calculated using equations (3.12), (3.13) and (3.14) below:

$$CS = -\frac{1}{\beta_c}(V_1 - V_0) \tag{3.12}$$

$$V_0 = \alpha + \beta_A A_0 + \beta_B B_0 + \beta_R R_0 \tag{3.13}$$

$$V_1 = \alpha + \beta_A A_1 + \beta_B B_1 + \beta_R R_1 \tag{3.14}$$

This might look a bit complicated but is actually very easy, and calculated with an Excel spreadsheet once you have got your estimates from the choice model in equation (3.9). Equation (3.12) says that the Compensating Surplus (*CS*) from an improvement in wetlands conservation – that is, the average person's willingness to pay for this package of changes – is given by the difference between their (measurable) utility before the improvement goes ahead, given by V_0, and their measurable utility after the change, V_1, converted into monetary units using the coefficient on the tax or price attribute, β_c. In turn, utility in the 'before' and 'after' cases is given by the levels of the attributes in each case (so A_0, B_0 and R_0 in the 'before' case, A_1, B_1 and R_1 in the 'after' case), multiplied by the attribute coefficients, and including the term α. This was the constant in equation (3.10), and is usually referred to as the *Alternative Specific*

Constant. It shows the utility people get simply from either staying in the status quo or leaving it (depending on whether it is positive or negative), independently of the values taken by the attributes. By fixing the status quo utility (as in equation 3.13), and varying the levels of the attributes, compensating surplus figures can be produced for as many combinations of attributes and levels as the design makes possible: that is, for a wide range of policy outcomes. We illustrate this in Box 3.5 for a soil conservation programme in Spain. It is this flexibility of choice experiments which makes the method so popular.

BOX 3.5 A SPANISH SOIL EROSION STUDY

Colombo et al. (2005) use the choice experiment to estimate the benefits of reducing soil erosion in Andalusia, Spain. The study considers the reduction of the off-site impacts of soil erosion in two watersheds, the Genil and the Guadajoz. Due to soil and climatic conditions and the nature of current farming practices, soil erosion levels in these catchments are well in excess of national average levels, and are known to result in widespread environmental problems. Among the most important of these are increased desertification, the siltation of water bodies, and reductions in biodiversity. To reduce these impacts it is necessary to provide subsidies to farmers to encourage them to adopt soil conservation measures in their land management. These measures include sowing a grass cover in olive orchards and reforesting degraded hill and mountain slopes. The choice experiment used the following attributes:

- desertification in semi-arid areas;
- quality of surface and groundwater;
- effects on flora and fauna;
- agricultural jobs safeguarded;
- area of countryside covered by the measures;
- cost to households in the area of the policy.

Attribute levels were defined in a number of ways. For example, for desertification, respondents were told what the current situation was, then it was explained that policy could change this to a small improvement or a moderate improvement. In both cases, respondents received an explanation of what this would actually mean 'on the ground', using words and pictures. The results showed that respondents had a positive willingness to pay for improvements in all of the policy attributes. Implicit prices were calculated and gave the following results (all values are in euros per household per year):

- For a change in desertification from continuing degradation (the status quo) to a 'small improvement': 17.78 (95% confidence interval: 12.02–25.21).

- For a change in desertification from continuing degradation (the status quo) to a 'moderate improvement': 26.51 (95% confidence interval: 20.05–35.76).
- For a change in water quality from 'low' to 'medium' quality: 18.39 (95% confidence interval 12.67–25.96).
- For a change in water quality from 'low' to 'high' quality: 26.27 (95% confidence interval 20.10–34.67).

Finally, the Compensating Surplus for a number of policy scenarios was measured, using the formulae given in this chapter. For instance, for a policy which produced a big improvement in desertification, high levels of water quality, good (versus declining) species numbers, 150 farm jobs and which covered 500 hectares, the mean WTP was €40.98, with a 95 per cent confidence interval from €34 to €47 per household per year.

3.2.3 Problems with the Choice Experiment Method

Accommodating variation in preferences across people

The standard approach to choice experiments which was described above has one important feature that needs a comment. This is that, if we use a Conditional Logit model to represent the choices that people make – as in equation (3.9) and the example in Box 3.4 – then we are effectively assuming that each person in the sample places the same value on each attribute used in the design. In other words, we effectively assume that the marginal utility for Joe if attribute X_1 is increased – β_1 – is the same as the marginal utility for Jane, and that the marginal utility for Joe of an increase in attribute X_2, β_2, is the same as that for Jane. This is because we only estimate one value for β_1 and one value for β_2 in equation (3.10). Now, as the example in Box 3.4 shows, we can allow that the value of a change away from the status quo can depend for an individual on their age or education, since we interact the constant with these terms. But this is a very limited way of handling *preference heterogeneity*, whereby we actually expect that people will care to different degrees about the same attribute.[1] We could also split the sample according to what we imagine might be a reasonable grouping according to preferences (for example between old and young, between rich and poor, rural and urban), but again this requires us to know how best to do this.

Choice experiment practitioners have thus looked for alternative ways of modelling preference heterogeneity. This literature is rather technical, so cannot be expanded on in detail here. But two approaches can be mentioned. One is known as the *random parameters logit model*. This represents variations in preferences by including two terms for each attribute in equation (3.7): a mean effect, which represents average preferences, and a standard deviation term, which represents how much preferences in the sample vary around this mean. The second approach is called the *latent class model*. This takes a rather different approach: respondents are divided by an algorithm into latent (that is, unobservable) classes according to how they have responded to the choice questions, or according to their observable characteristics. A set of preference parameters, that is the β values in equation (3.7) are then estimated for each class. For an example of how both approaches can be used, the reader can consult Birol et al. (2006).

Issues with experimental design

Designing a choice experiment is almost an art form! Decisions must be taken on a great number of issues:

i. what attributes to include;
ii. how to describe them to respondents;
iii. what levels are to be used for each attribute;
iv. what price or cost term will be used;
v. how the attributes and levels are combined in choice sets;
vi. how many choice sets respondents can deal with;
vii. how many choice options are included in each choice set.

It is likely that the estimate we get for the willingness to pay of respondents for a change in any particular attribute, or how precise a measure we obtain of this, depends on what decisions are made above. The overall success of the choice experiment in terms of what it tells us about people's choices and values also depends on these steps. Many papers exist which investigate these issues, mostly in non-environmental applications of the method (for example in a transport, marketing and health context): lessons learnt can be found in the main choice experiment textbooks, such as Louviere et al. (2000) and Henscher et al. (2005). Suffice it to say that the best way of designing choice experiments is still an open question, partly because of the several ways in which 'best' can be interpreted. Moreover, choice experiments also depend, just as contingent valuation studies do, on the description of the hypothetical market and on sample selection.

Hypothetical market bias

Another parallel between choice experiments and contingent valuation is the possibility that responses in a hypothetical market setting will tell us little about how respondents would behave in a real market. This issue has been addressed in a couple of ways within the CE literature, comparing real with hypothetical responses in terms of (i) how well hypothetical choices predict real choices, and (ii) how close predicted WTP from hypothetical choices is to real WTP in an actual market. Of course, the same problem faces the CE practitioner as faces the CVM analyst, that for most environmental goods we cannot observe 'real' market prices – that is the problem the method tries to address! However, some findings exist which compare real with hypothetical choices where this problem can be got around: these suggest that the extent of hypothetical market bias might not be too extreme in CE (Blamey et al., 2001; Carlsson and Martinsson, 2001). More recent evidence is presented by List et al. (2006), who compared actual with hypothetical scenarios for two choice experiments. They argue that two tests are of interest – whether a hypothetical choice experiment overstates the extent to which people would actually pay for, say, wetland conservation and the differences, if any, in the marginal values of the attributes used in the choice experiment between 'real' and 'hypothetical' choices. They found no statistically-significant differences between hypothetical and real WTP, or between the marginal values of attributes, when a 'cheap talk' script was used as part of the choice experiment – that is, when respondents were explicitly told about the problem of hypothetical market bias, and asked to consider their responses carefully.[2] Finally, choice experiment responses are also known to be liable to a 'status quo bias' – a tendency for respondents to choose the 'do nothing, zero additional cost' option for reasons other than utility differences between this and the other choice options. This can be diagnosed by testing whether the parameter estimate for the Alternative Specific Constant for the status quo choice is statistically significant or not.

Is the value of the whole equal to the sum of the parts?

One of the advantages of choice experiments is that they enable the researcher to do two things: (i) estimate the value for each of the attributes of an environmental good; and (ii) estimate the value for a policy which changes many of these attributes simultaneously. Now imagine that we wish to use CE to value the protection of a forest threatened with felling. A CE study is undertaken which estimates values for five forest attributes, which includes 'loss of the forest' as a level for each. Can these values be added up to show the economic loss from the forest being felled? Or imagine a landscape valuation study which identified five landscape attributes, and is then used to predict the economic value of changes in landscape quality. Can the value of a future landscape be inferred from the sum of the characteristic values? This is an issue revolving around whether people think about environmental goods as bundles of attributes (this is what the theory assumes), and around whether the CE designer has done a good job in selecting the attributes. But in cases where we are more concerned with the 'value of the whole' rather than the 'value of the parts', it might be wise to undertake a contingent valuation study rather than a choice experiment.

3.3 BENEFITS TRANSFER

Benefits Transfer (BT) is the practice of extrapolating existing information on the non-market value of goods or services (Brouwer, 2000). Typically, the practice involves predicting compensating or equivalent surplus values for an environmental quality or access change at one site, based on data collected using either stated or revealed preference methods at another, similar site. Adjustments are often made for differences between the environmental characteristics of the site to which values are to be transferred (known as the 'policy site') and those of the site at which the original data was collected, known as the 'study site' (Downing and Ozuna, 1996). Differences in socio-economic characteristics of the affected population between the study and policy sites can also be allowed for (Morrison et al., 2002).

The aim of BT techniques is to provide decision makers with a monetary valuation of environmental goods and service in a cost-effective and timely manner, since original valuation studies are both expensive and time-consuming. Demands for environmental valuation estimates are rising in the policy community in both Europe and the US. In Europe, this is partly being driven by the introduction of the Water Framework Directive, which requires benefit–cost analysis of water quality improvements throughout the European Union, and by the greater emphasis on the application of cost–benefit principles to environmental policy design in the EU (European Commission, 2002). In the UK, widespread use of benefits transfer has already occurred within policy making and regulatory bodies, for instance in the setting of water quality targets for private water companies.

Papers investigating the use and accuracy of BT have become increasingly frequent since an initial set of papers on the subject appeared in a special issue of *Water Resources Research* in 1992. Recent applications of BT include Rozan (2004) on improved air quality in France and Germany, Muthke and Holm-Müller (2004) on national and international transfers of water quality improvement benefits, Jiang et al. (2005) on coastal land management, and Colombo and Hanley (2008) on agricultural landscapes.

Many early BT studies used the contingent valuation method to undertake benefit transfers. However, Morrison et al. (2002) pointed out that, within the field of stated preference methods, Choice Experiments are

arguably better suited to BT because it is possible to allow for differences in environmental improvements across sites as well as differences in socio-economics characteristics across impacted populations. Moreover, compensating surplus estimates for a wide range of potential policy scenarios can be calculated from the choice models estimated. Benefits transfer has also been investigated using the revealed preference methods of recreation demand modelling which we outline in the next chapter.

The accuracy of BT can be tested in a number of ways. Two main approaches have been followed in the literature. The first is the transfer of mean WTP values from the policy site to the study site. Transferring unadjusted mean values has been criticized since it does not take into account any possible differences between either the populations or the goods at the policy and study site. Because of that, an alternative adjusted mean value approach has developed, which adjusts mean WTP of the study site to account for differences in the environmental characteristics of the policy site and/or for differences in the socio-economic characteristics of the affected population between the two sites. In the case of unadjusted mean value transfer, the null hypothesis of benefits transferability is:

$$WTP_s = WTP_p \tag{3.15}$$

where WTP_s and WTP_p are the mean WTP at the study and policy sites measured from two different original studies. In the case of the adjusted value transfer, the WTP_s is adjusted using data on socio-economic and environmental characteristics of the policy site, before the comparison takes place. Such adjustments are, to a varying degree, somewhat ad hoc.

The second approach to BT is benefit function transfer, where the entire demand function (or choice equation, in a CE setting) estimated at the study site is transferred to the policy site. Values at the policy site are predicted using independent variables (such as household income) collected from secondary data at the policy site and parameter values estimated from the study site. In the benefit function transfer the regression parameters of the study site and the environmental and population characteristics of the policy site are used to test:

$$\text{predicted } WTP\ (\beta^s, X^p) = WTP^p \tag{3.16}$$

where predicted WTP (β^s, X^p) is the willingness to pay at the policy site estimated using the parameters of the benefit function of the study site (β^s) and the X values (site attributes, socio-economics characteristics and so on) of the policy site and WTP^p as defined above. An alternative test is the comparison of function parameters between the study and policy site:

socio-economic characteristics in separate sub-markets have the same preferences, we can look at the intersection of the *MV* curve with this second implicit price curve, *Ip2*. This defines another point of equilibrium for 'type a' buyers, and now we can identify the *MV* curve as being the solid line rather than the dashed line. This is the approach taken by Boyle et al. (1999). A second statistical problem relates to the fact that in choosing the quantities of each attribute to purchase when buying a house, people are also effectively determining the implicit price we observe. This gives rise to a problem of endogeneity of the regressors in the second stage demand equation, so that the researcher must try and implement an instrumental variables approach (Taylor, 2003).

Haab and McConnell (2002) conclude that 'with a few exceptions, researchers have abandoned attempts to (measure) preferences, and work instead with the hedonic price function'. Indeed, as McConnell and Walls (2005) note in their review of HP studies of the value of open space 'almost none of the studies attempt to carry out a second stage estimation of the demand function . . . most focus on the marginal price of an additional acre of open space'. This focus on implicit prices is common in the HP literature, and is due to the statistical problems of second stage demand estimation noted above. For examples of a study which does manage to estimate a second-stage demand curve, see the work by Day et al. (2007) on the value of peace and quiet in the housing market in Birmingham, UK; and the study by Boyle et al. (1999) on water quality in Maine, USA.

BOX 5.2 THE VALUE OF GREEN SPACE IN GUANGZHOU, CHINA

A study by Jim and Chen (2006) explores the amenity value provided by environmental and landscape attributes such as urban green spaces, water bodies and noise that influence residential housing prices in the city of Guangzhou, China. The environmental features included in the model specification were: window orientation, view of green spaces, traffic noise, nearby water bodies and presence of nearby wooded areas. Linear and semi-logarithmic functions were applied to test the relationship between the housing units' sale prices and their characteristics, including these environmental features. In both models, the number of bedrooms, exposure to traffic noise and proximity to wooded areas were statistically insignificant. The possession of southward and northward-facing windows had a positive effective on apartment selling price, as did proximity to water bodies and a 'green view', while increasing distance from the town centre lowered the selling price. Using the implicit prices obtained from the semi-log model, the authors conclude that 7.1 per cent of the selling price is due to a 'green space' view, whilst 13.2 per cent is contributed by proximity to water bodies.

BOX 5.3 ECO-LABELS AND THE CLOTHING INDUSTRY

Most of the discussion of the HP method in the main text is concerned with the relationship between environmental quality and house prices. But as hinted, the approach can be used to search for relationships between any marketed product and the characteristics of that product, which can include different indicators of environmental quality. Nimon and Beghin (1999) use HP to investigate whether consumers are willing to pay a price premium for 'environmentally-friendly' clothing. As the authors note, some manufacturers seek to differentiate their products on the basis that they are 'organic', or 'natural', based on the expectation that some consumers will be willing to pay extra for a shirt, say, which has been made from a production process (organic cotton growing) which, it is claimed, has lower environmental impacts than conventional cotton farming. Indeed, a glance at the Greenpeace clothing catalogue will show many examples of such differentiated goods on offer. People may also be willing to pay more for a product which is differentiated in terms of possible health impacts of its production: this was claimed to be the case for cotton clothing in terms of the nature of dyes used to colour material. The environmental and health attributes the authors consider are thus 'organic cotton', 'environmentally-friendly dyes' and 'no dyes'.

A semi-log functional form was used for the hedonic price equation, based on 750 products on sale. Results show that a large and statistically significant relationship exists between the organic cotton label and prices, with an average mark-up of 34 per cent being found for organic clothes. Interestingly, only 37 per cent of this mark-up is accounted for by the higher cost of producing organic cotton. No significant effects were found, however, for environmentally-friendly dye labels (interestingly, no-dye clothes sold for a discount). The authors note that a stronger relationship might be expected for health-related characteristics in the case of children's clothing, since parents might care more about their children's health than their own, but they find no evidence for this in the data.

5.4 PROBLEMS WITH THE HP METHOD

Omitted Variable Bias

If a variable that significantly affects house prices is omitted from the HP equation, which is in addition correlated with one of the included variables, then the coefficient on this included variable will be biased.

This might be a particular problem when 'emitters' cause more than one impact. For instance, paper mills will affect water quality, which impacts on house prices, but may also impose disutility due to odours. Traffic can cause disutility due to noise, but also due to dust and safety concerns. Including only either noise or water quality in the HP equation will result in biased estimates for the marginal values of noise/water quality (Leggett and Bockstael, 2000); this was suspected to be the case in the Copenhagen noise study referred to above. Leggett and Bockstael (2000) solve the problem by including separate variables for water quality levels and for the distance of houses from pollution sources; fortunately, due to the nature of the natural processes relating emissions to water quality levels, these two variables are not too correlated with each other.

Multi-collinearity

Some attributes in the hedonic price function may be highly correlated with each other. For example, houses close to a river may score highly in terms of both 'peace and quiet' and 'scenic quality of views'. Houses in deprived neighbourhoods may score badly both in terms of 'local crime rates' and 'quality of local amenities'. This means that the parameter estimates for implicit prices will be imprecise, and that the effects of attributes that are highly correlated with each other (for example two measures of air pollution in a city) will be difficult to disentangle from each other. The researcher might decide in such cases either to leave out some explanatory variables from the hedonic price equation, or to seek for alternative ways of representing their influence on house prices.

Choice of Functional Form for the Hedonic Price (HP) Function

Economic theory does not specify which functional form should be used for the HP equation, yet the choice of functional form will influence the value that implicit prices take. We can reasonably suggest that the functional form used should allow house prices to rise as more of a desirable attribute is supplied, and that linear models may be rather unrealistic, since they imply that the cost of buying cleaner air quality or more bedrooms does not vary with the quantity of these attributes purchased. Choice of which form to use will thus depend on econometric considerations, and flexible forms such as the Box–Cox have been suggested and used (Cropper et al., 1988). Semi-log forms where the natural log of house prices is used as the dependent variable are also popular, since they also allow for non-linear implicit prices, which can be calculated using a simple formula (for example Geoghegan et al., 2003). In Leggett and Bockstael (2000), results for four different functional forms are presented. For a comprehensive guide to how implicit prices will vary according to the functional form of the HP function, see Taylor (2003, p. 354).

BOX 5.4 THE VALUE OF OPEN SPACE

McConnell and Walls (2005) undertook a review of non-market valuation methods and the value of 'open space'. They review 40 studies published between 1967 and 2003, organized according to whether they were concerned with 'general open space, parks and natural areas', green-belts, wetlands and forests, and agricultural lands. Some of the implicit prices for open space were found to be negative, and some statistically insignificant, but in most cases proximity to open space is correlated with an increase in house prices. Some of the results surveyed are shown below:

Study	Type of open space	Marginal value in $ for living 200 metres closer
Anderson and West, 2003	State/regional parks, wildlife refuges	600
Schultz and King, 2001	Wildlife habitat	429
Doss and Taff, 1996	Open-water wetland	1980
Mahan et al, 2000	Wetland of any type	286
Smith et al, 2002	Public open space	–553
		Marginal value from conversion of 1 acre
Irwin, 2002	Conservation land	3307

They conclude that open space values seem to depend on location, type of open space and research methods.

Market Segmentation

The hedonic price function relates, in theory, to the equilibrium implicit prices for housing attributes in a single market. How big this market is in spatial terms can be difficult to assess. If we study the relationship between traffic noise and house prices in Glasgow, should we consider the whole of the city to be one housing market, or are there separate markets North and South of the river, with separate hedonic price functions for each? It can be hard to test for this market segmentation econometrically, since we are unsure about both functional form and market size (Palmquist, 2003). Michaels and Smith (1990) use definitions of separate markets from realtors (estate agents) to solve this problem. Geoghegan et al. (2003) estimate separate HP models for three neighbouring counties in Maryland to look at the effects of protecting open space (agricultural, forest, and parkland and golf courses) on property values. They found that the implicit price of open space varied a lot across these three counties: for one county (Carroll), open space had no significant effect on house prices, whilst for the other two, the effects on house values of increasing open land conservation by 1 per cent was much higher in Calvert County than in Howard County.

Expected or Perceived Versus Actual Characteristic Levels

House sales may be a function of expected future environmental conditions in addition to current observed conditions. For example, the implicit price for noise may also show what people expect to happen to noise levels in that part of town in the next 10 years, not just what noise levels are when the study is undertaken. Also, implicit prices for open space may depend on what people think will happen to this open space in the future (Smith et al., 2002). McCluskey and Rausser (2003) look at the effects over time of the discovery and eventual remediation of toxic wastes from an old lead smelter, which affected house values in Dallas County, Texas. The authors allow for the effects of distance from the smelter on housing values to vary with time. One can argue that one factor which varies with time is people's beliefs about the extent of the risk from wastes left behind by the smelter (which were used as part of landfill for construction), and the likelihood and extent of eventual remediation of the risks. Thus, people's beliefs about how risk levels were changing over time, as well as actions which actually reduced risks (such as the various stages of clean-up which occurred at the site), may be what drives house prices.

Another problem arises in that individuals' subjective values of such risks are likely to be either less than or greater than the scientific probability of health damages occurring. People often tend to overestimate the likelihood of low probability, high cost events (such as a plane crash) occurring, and underestimate the likelihood of high probability events happening. The implication is that hedonic prices may either overestimate or underestimate welfare changes according to whether a low or high objective probability event is being considered, and to the amount and quality of information available to individuals (Kask and Maani, 1992). Recent work on this issue has looked at how people learn about risks, and how this relates to behaviour in housing markets. For example, Hallstrom and Smith (2005) study the effects of a 'near miss' hurricane, Hurricane

Andrew, on housing markets in Lee County, Florida. The argument is that this near-miss caused local residents to re-evaluate the risks of living in a hurricane-prone part of the US, and that this re-appraisal of risks should be reflected in house prices. The authors indeed find that the near miss caused a fall in house prices due to a re-evaluation of risks, this fall being equivalent to about 3 per cent of average annual income in Lee County.

Spatial Auto-Correlation

Spatial auto-correlation refers to the phenomenon whereby certain factors influence house prices for all properties in a neighbourhood, but are not observable to the researcher. This means that the error term in equation (5.2) is correlated across neighbouring properties. The effects are to make the estimates of the hedonic price equation parameters inefficient, and to bias standard errors (making the associated t-statistics 'too big'). This means we might incorrectly infer that an attribute has a significant effect on house prices, when in fact it does not. Spatial auto-correlation can be tested for, and steps taken to remove its effects – see Geoghegan et al. (2003) for details.

Restrictive Assumptions

The HP gives an accurate estimate of the value of environmental quality only if (i) all buyers and sellers in the housing market are well informed of attribute levels at every possible housing location; (ii) all buyers in the market are able to move to utility-maximizing positions (otherwise, marginal cost is not equivalent to marginal WTP); (iii) the housing market is in equilibrium; the vector of implicit prices is such that the market clears at all times. Clearly, these assumptions will never fully describe reality. For example, buyers could be more poorly informed about the characteristics of certain houses than sellers. Pope (2008) notes that information disclosure laws in the US over house sales implies that the government indeed thinks that buyers do not 'know enough'. He looks at the effects of airport noise disclosures on house prices around Raleigh-Durham Airport in North Carolina. He found that disclosure laws increased the implicit price of aircraft noise by 37 per cent, leading him to conclude that the 'information environment' should be carefully considered when using HP to value amenities and disamenities. In the case of aircraft noise he looks at, one might say that a HP study carried out on the disamenity of aircraft noise prior to the implementation of information disclosure laws would have undervalued the costs of noise nuisance.

BOX 5.5 VALUING THE DISAMENITIES OF LANDFILL SITES

Landfill sites, whether for municipal solid waste or industrial wastes, have long been associated with impacts on house prices, since the assumption is that no one wants to live next to a landfill. Other things being equal, then, house prices will have to be lower, the closer one gets to a landfill site, to compensate buyers for the negative externalities of such facilities – noise, smell and seagulls! Several HP studies of landfill impacts on house prices can be found in the literature, including an interesting article by Hite et al. (2001). Hite et al. explain that both distance to a landfill site and the expected lifetime of that site can be expected to have an effect on house prices; whilst how well-informed house buyers are about landfill sites in an area could also matter to the implicit prices the analyst can uncover. They also allow for the fact that property taxes matter to the house buyer, and these depend both on public goods supplied in a neighbourhood (for example spending on schools) and on house prices.

The study is based on 2913 house sales in Franklin County in Ohio in 1990. House sales information was supplemented with data on household socio-economic characteristics for buyers. Environmental and neighbourhood characteristics data was also collected. Four landfill sites exist within the study area, and the distance of each house in the data base to each site was measured. Information was also included on how long these sites had left to operate (two had already closed in 1990). The authors find a significant effect of distance from all of the four landfill sites on property values, and that this effect persists even after a landfill has closed. The longer the lifespan a landfill site has at the time of sale, however, the lower the house price. They conclude that 'welfare losses from decreased property values near landfills can be of a significant magnitude'.

A similar study is that by Eshet et al. (2007) for waste transfer facilities in Israel. The authors use data from four cities to study the relationship between distance from the waste site and house prices. The data set consists of 9505 house sales located within 4 km of a waste site. Regression results using the quadratic model showed that the maximum distance affected by disamenities varied between the four cities from 2.29 to 3.29 km. Housing prices increase at a decreasing rate away from the transfer station: moving from the second to the third kilometre adds US$4460 to the price of an average house, whereas moving from the third to the fourth kilometre away adds only US$3150 to the price of an average house.

5.5 CONCLUSIONS

As has been pointed out in the preceding section, there are many problems associated with the HP technique. Perhaps the most important of these are the assumptions made about the related market (the housing market, in this chapter). Moreover, the method cannot be used to measure

non-use values, and is restricted in terms of the kinds of environmental goods to which it can be applied (some tie has to be found to marketed goods). However, the method does make use of data on actual behaviour, unlike the stated preference methods described in Chapter 3. Although this chapter on HP has concentrated on house prices and environmental quality levels, the technique is applicable to other goods. HP can be used to estimate the implicit price of any observable characteristic of any good, so long as adequate data is available. HP can therefore be used to estimate the value of the 'green premium' on environmentally-friendly consumer goods (see Box 5.3), or the value of environmental risks on human health through wage differentials.

How reliable are hedonic price estimates of environmental benefits? Smith and Huang (1993) conducted a meta-analysis of 37 HP studies, to see how well they could detect the influence of air pollution on house prices. The authors report that 74 per cent of the studies found a negative and significant relationship between measures of air quality and house prices. They find that, overall, 'there is a systematic relationship between the modelling decisions, the descriptions used to characterise air pollution, the condition of local housing markets, and the conclusions reached about the relationship between air quality and house prices' – see also their 1995 meta-analysis (Smith and Huang, 1995). Palmquist cautions in his review of the HP literature that 'there is still substantial room for improvement' (Palmquist, 2003, p. 64), but this comment could equally be applied to all valuation methods!

As a means of measuring marginal values for certain environmental goods, the hedonic price method has much to recommend it.

NOTES

1. We thank V. Kerry Smith for his extensive and very helpful comments on this chapter.
2. That is, the rate at which an individual is willing to exchange one good for another: the slope of an indifference curve.
3. As with the travel cost model.

REFERENCES

Anderson, R.J. and T.D. Crocker (1971), 'Air pollution and residential property values', *Urban Studies*, **8**, 171–80.

Bjorner, T., J. Kronbak and T. Lundhede (2003), *Valuation of Noise Reduction: Comparing Results from Hedonic Pricing and Contingent Valuation*, Copenhagen: AKF Forlaget.

Boyle, K., J. Poor and L. Taylor (1999), 'Estimating the demand for protecting freshwater lakes from eutrophication', *American Journal of Agricultural Economics*, **81**(5), 118–22.
Cropper, M., L. Deck and K. McConnell (1988), 'On the choice of functional form for hedonic price functions', *Review of Economics and Statistics*, **70**, 668–75.
Day, B., I. Bateman and I. Lake (2007), 'Beyond implicit prices: recovering theoretically consistent and transferable values for noise avoidance from a hedonic price model', *Environmental and Resource Economics*, **37**, 211–32.
Eshet, T., M.G. Baron, M. Shechter and O. Ayalon (2007), 'Measuring externalities of waste transfer stations in Israel using hedonic pricing', *Waste Management*, **27**, 614–25.
Freeman, A.M. (2003), *The Measurement of Environmental and Resource Values*, Washington, DC: RFF Press.
Garrod, G. and K.G. Willis (1992), 'The amenity value of woodland in Great Britain', *Environmental and Resource Economics*, **2**(4), 415–34.
Geoghegan J., L. Lynch and S. Bucholtz (2003), 'Capitalization of open spaces into housing values and the residential property tax revenue impacts of agricultural easement programs', *Agricultural and Resource Economics Review*, **32**(1), 33–45.
Haab, T. and K. McConnell (2002), *Valuing Environmental and Natural Resources: The Econometrics of Non-market Valuation*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Hallstrom, D.G. and V.K. Smith (2005), 'Market responses to hurricanes', *Journal of Environmental Economics and Management*, **50**(3), 541–61.
Hite, D., W. Chern, F. Hitzhusen and A. Randall (2001), 'Property value impacts of an environmental disamenity: the case of landfills', *Journal of Real Estate Finance and Economics*, **22**, 185–202.
Jim, C.Y. and W.Y. Chen (2006), 'Impacts of urban environmental elements on residential housing prices in Guangzhou (China)', *Landscape and Urban Planning*, **78**, 422–34.
Kask, S. and S. Maani (1992), 'Uncertainty, information and hedonic pricing', *Land Economics*, **68**(2), 170–84.
Lancaster, K.J. (1966), 'A new approach to consumer theory', *Journal of Political Economy*, **74**, 132–57.
Leggett, C.G. and N. Bockstael (2000), 'Evidence on the effects of water quality on residential land prices', *Journal of Environmental Economics and Management*, **39**, 121–44.
McConnell, V. and M. Walls (2005), 'The value of open space: evidence from studies of non-market benefits', Discussion Paper, Resources For the Future, Washington, DC, available from www.rff.org.
McClusky, J. and G. Rausser (2003), 'Hazardous waste sites and housing appreciation rates', *Journal of Environmental Economics and Management*, **45**(2), 166–76.
Michaels, R. and V. Smith (1990), 'Market segmentation and valuing amenities with hedonic models: the case of hazardous waste sites', *Journal of Urban Economics*, **28**, 232–42.
Nimon, W. and J. Beghin (1999), 'Are eco-labels valuable? Evidence from the apparel industry', *American Journal of Agricultural Economics*, **81**(4), 801–12.
O'Byrne, P., J. Nelson and J. Seneca (1985), 'Housing values, census estimates, disequilibrium and the environmental cost of airport noise', *Journal of Environmental Economics and Management*, **12**, 169–78.

Palmquist, R. (1991), 'Hedonic methods', in J. Braden and C. Kolstad (eds), *Measuring the Demand for Environmental Improvement*, Amsterdam: North Holland.
Palmquist, R. (2003), 'Property value models', in K.-G. Maler and J. Vincent (eds), *Handbook of Environmental Economics, Volume 2*, Amsterdam: North Holland.
Poor, P.J., K.L. Pessagno and R.W. Paul (2007), 'Exploring the hedonic value of ambient water quality: a local watershed-based study', *Ecological Economics*, **60**, 797–806.
Pope, J.C. (2008), 'Buyer information and the hedonic: the impact of seller disclosure on the implicit price of aircraft noise', *Journal of Urban Economics*, **63**(2), 498–516.
Rosen, S. (1974), 'Hedonic prices and implicit markets: product differentiation in pure competition', *Journal of Political Economy*, **82**, 34–55.
Smith, V.K. and Ji Chin Huang (1993), 'Hedonic models and air quality: 25 years and counting', *Environmental and Natural Resource Economics*, **3**(4), 381–94.
Smith, V.K. and Ji Chin Huang (1995), 'Can markets value air quality? A meta-analysis of hedonic property value models', *Journal of Political Economy*, **103**, 209–27.
Smith, V.K., C. Poulos and H. Kim (2002), 'Treating open space as an urban amenity', *Resource and Energy Economics*, **24**, 107–29.
Taylor, L. (2003), 'The hedonic method', in P. Champ, K. Boyle and T. Brown (eds), *A Primer on Non-Market Valuation*, Dordrecht, NL: Kluwer.