

# Cost-Benefit Analysis as Market Simulation

*A New Approach to the Problem  
of Anomalies in Environmental  
Evaluation*

---

Robert Sugden

The Frontiers of   
Environmental Economics

**Cost-Benefit Analysis as Market Simulation:  
A New Approach to the Problem of Anomalies in Environmental Evaluation**

**Robert Sugden**

School of Economics, University of East Anglia,

Norwich NR4 7TJ, England

email: r.sugden@uea.ac.uk

28 June 2007

**Keywords:** cost-benefit analysis, market simulation, willingness to pay, environmental evaluation

**JEL classifications:** D61 (allocative efficiency, cost-benefit analysis), Q51 (valuation of environmental effects)

**Acknowledgements**

A previous version of this paper was presented at the conference ‘Frontiers of Environmental Economics’, organised by Resources for the Future in Washington, D.C., 26-27 February 2007. I thank participants at this conference, particularly my discussant, Michael Hanemann, for valuable comments. I also thank Richard Thaler for advice on the interpretation of ‘libertarian paternalism’. My work was supported by the Economic and Social Research Council of the UK (award no. RES 051 27 0146).

## **Abstract**

Most normative economics assumes that individuals have coherent preferences. This paper responds to growing evidence of failures of this assumption, particularly in the context of stated-preference methods widely used in environmental policy analysis. I propose a criterion of consumer sovereignty which does not presuppose preference coherence, and which is satisfied by competitive markets. I then propose an approach to cost-benefit analysis that attempts to simulate consumer sovereignty in situations of market failure. The key idea is to use valuations revealed 'at the moment of consumption'. I argue that this principle is better implemented by hedonic pricing than by contingent valuation.

Most normative economic analysis assumes that individuals have coherent preferences over all relevant outcomes. More specifically, preferences are assumed to be stable (that is, not liable to sharp changes over short periods of time, and not subject to significant degrees of random variation), context-independent (not affected by variations in the ‘framing’ of what, in terms of economic theory, is the same decision problem), and internally consistent (satisfying conventional properties of consistency, such as completeness and transitivity). For many years, these assumptions were taken to be uncontroversial. When acting as policy advisers, economists used methodologies based on these assumptions. But recent developments in experimental economics, behavioural economics and stated-preference research has generated evidence which suggests that, in many of the contexts for which public policy decisions are required, individuals do *not* have coherent preferences. These findings cause particular difficulties for those forms of cost-benefit analysis, widely used in environmental policy appraisal, which rely on stated-preference methods. The aim of this paper is to consider how economics can make a useful contribution to normative policy analysis if and when the assumption of preference coherence fails to hold.

In Section 1, I introduce the problem of preference incoherence in general terms. Traditionally, normative economics has used a criterion of preference-satisfaction. In the rhetoric of the discipline, this criterion has been defended as an application of the principle of consumer sovereignty, and in contrast to paternalism. But preference-satisfaction, as traditionally understood, cannot be used as a decision criterion when preferences are incoherent. Some influential behavioural economists are now arguing that, in the light of the evidence of preference incoherence, economics needs to reconsider its traditional aversion to paternalism. In Section 2, I review this argument for ‘soft paternalism’. In Section 3, I present an opposing argument, based on a re-interpretation of the principle of consumer sovereignty. I propose an understanding of consumer sovereignty in which value is attached to the satisfaction of individuals’ demands, as and when they are expressed, and irrespective of whether they reveal coherent preferences. Drawing on analysis from an earlier paper, I show that, under ideal conditions, competitive markets in private goods are effective in satisfying such demands. In Section 4, I propose an approach which treats cost-benefit analysis as an attempt to simulate these properties of ideal markets in situations of market failure. The central idea is that the valuations that should be used in cost-benefit analysis are those that individuals would reveal ‘at the moment of consumption’. This provides a general

principle for choosing among hypothetical payment mechanisms for eliciting valuations of non-market goods. In Section 5, I argue that this principle favours valuations derived from hedonic prices rather than from stated preference studies in cases in which both methods are practicable. I suggest that because hedonic-price valuations are elicited closer to the moment of consumption, they may be less susceptible to anomalies. In the final section, I draw some general conclusions.

## **1. The problem of preference incoherence**

Why is the assumption of preference coherence so important for normative economics? The most fundamental answer is that it plays a crucial role in the philosophical justification of the methods of normative economics. The standard justification works by treating a person's preferences as indicators of her 'welfare' or 'well-being'. This idea supports a presumption – a presumption that has been upheld by many generations of welfare economists – in favour of consumer sovereignty and against paternalism. If we were forced to conclude that people's preferences are typically unstable, context-dependent or internally inconsistent, it would be difficult to maintain the claim that preferences provide a measurement of well-being. That might seem to call into question the whole strategy of basing normative analysis on individuals' preferences.

The coherence assumption has an additional role when the principles of normative economics are put into practice in cost-benefit analysis (CBA). In a conventional CBA, the aim is to assess the impact of some public policy option on individuals, measured in terms of preference-satisfaction. This requires the analyst to *elicit* individuals' preferences over the outcomes of the policy and of a benchmark 'do-nothing' option. Notice that neither of these outcomes has yet been realised (the policy decision must precede its effects), that at least one of the two will not be realised at all (the policy cannot be both implemented and not implemented), and that individuals cannot make private choices between the two outcomes (the policy cannot be implemented for one person and not for another). So we cannot discover a person's preferences over these specific outcomes by reading them off from a choice problem in which they are revealed *directly*. Elicitation has to be indirect: we collect evidence from contexts other than the policy decision itself and then *infer* the preferences we need. Such inferences are possible only within *a model of preferences*. The model provides a common framework within which we can represent both the effects of the policy decision and

the environment in which preferences are elicited. It imposes a structure on the data we collect which corresponds with the structure of the CBA for which the data are needed. In almost all CBA, preference coherence is a fundamental part of the modelling strategy by which preferences are elicited and policy-relevant inferences are made. If the coherence assumption is not consistent with the data we collect, we are in trouble.

Where such inconsistencies occur, stated-preference methods of elicitation make them highly visible. Stated-preference surveys, like laboratory experiments, elicit individual responses under highly controlled conditions. This makes it very easy to compare two or more responses made by the same individual, at almost the same time, to questions which differ only in specific respects. If the same question elicits two different responses from the same person over a very short time interval, there is an obvious violation of the stability assumption. If two questions with the same logical content but different framings elicit different responses, and if this ‘framing effect’ has a consistent pattern, there is an obvious violation of the context-independence assumption. And if two questions are related in such a way that a certain combination of responses is inconsistent with some standard consistency axiom, but responses systematically reveal that pattern, there is an obvious violation of the assumption of internal consistency.

The controlled nature of stated-preference elicitation is, I think, one reason why stated-preference research was one of the first areas of applied economics to confront the problem of ‘anomalies’ (that is, systematic deviations between actual decision-making behaviour and the implications of standard assumptions about preferences). Because environmental policy is one of the areas in which CBA is most reliant on stated-preference methods, environmental evaluation became one of the first battlegrounds of the methodological war between rational-choice and behavioural economics. However, it would be wrong to think of anomalies merely as artefacts of stated-preference methods. It is now becoming clear that the psychological mechanisms that induce anomalies in stated-preference studies and in laboratory experiments have significant impacts in many other areas of economics, including finance and industrial organisation.<sup>1</sup>

I do not want to get diverted into a discussion of the evidence about the significance and robustness of anomalies.<sup>2</sup> I shall just record three familiar examples of anomalies that cause severe problems for stated-preference studies. These are: the disparity between willingness-to-accept (WTA) and willingness-to-pay (WTP) valuations; the apparently insufficient sensitivity of stated valuations to the scale or scope of the good being valued; and

the tendency for stated valuations to be influenced by apparently irrelevant ‘cues’, such as the starting point in a series of dichotomous valuation questions. No doubt some readers will want to cite their favourite pieces of evidence in support of a claim that some alleged anomaly can be eliminated by the use of the right elicitation method. Of course, there are important issues to be resolved about the robustness of anomalies. (I am one of the many economists who are currently trying to resolve these issues.) But that is not the point of the present paper.

This paper asks how normative analysis should be carried out *in the event that* individuals do not have coherent preferences. There is surely enough evidence of anomalies to make it prudent to start thinking about this question now. One can always hope that this precautionary work will turn out to be unnecessary.

## **2. Soft paternalism**

In the last few years, behavioural economists have begun to consider the normative implications of their empirical findings. Among the establishment of American behavioural economists, a consensus seems to be emerging around the position of *soft paternalism*. The essential idea is that, if individuals’ preferences are incoherent, the justification for using preference-satisfaction as the criterion for public decision-making lapses; it becomes legitimate for ‘planners’ to use their own judgements about individuals’ best interests in place of inferences about what those individuals prefer. The fact that this position is so widely supported suggests that it coheres with ideas about normative analysis that are deeply rooted in economics. Thus, it is useful to begin by considering the arguments for and against soft paternalism. The arguments against provide the starting point for an account of my rival proposal.

The ideas I now discuss have been presented in slightly different forms by two sets of co-authors. Cass Sunstein and Richard Thaler (2003a, 2003b) use ‘libertarian paternalism’ as the label for their favoured approach. A very similar approach has been proposed independently by Colin Camerer, Samuel Issacharoff, George Loewenstein, Ted O’Donoghue and Matthew Rabin (2003), under the banner ‘regulation for conservatives’. Each set of co-authors includes one American legal scholar (Sunstein and Issacharoff); the other five authors are leading American behavioural economists. These labels declare the authors’ intentions to propose a justification for paternalistic regulation of economic behaviour which (it is claimed) is compatible with ‘libertarian’ or ‘conservative’ principles. I shall use the term *paternalism*

in the same way that these authors do: an intervention is paternalistic if its intent is to induce individuals severally to act in ways that the intervening agent judges to be in those individuals' best interests, but which those individuals would not choose in the absence of the intervention (Sunstein and Thaler, 2003b, pp 1161-1162).<sup>3</sup>

The two sets of co-authors appeal to essentially the same evidence from behavioural economics to support the claim that traditional assumptions about individual rationality are untenable. They present similar lists of approved policy interventions. They justify these interventions on essentially the same grounds, namely that individuals are being steered away from erroneous or irrational decisions. They propose that policy analysis should use a form of CBA which does not presuppose that valuations of costs and benefits are always to be based on individual preferences (but, as I shall show later, provide almost no guidance as to how such a CBA might work). Both of these proposals appeal to the idea that, if an individual's preferences are incoherent, her choices can be influenced by interventions which, although paternalistic in intent, do not contravene the principles to which anti-paternalists are committed. The key thought here is that anti-paternalists believe that public policy decisions should respect, or be responsive to, individuals' preferences. But if an individual lacks coherent preferences, there is nothing to respect, nothing to be responsive to. This creates a vacuum within which paternalistic policies do not offend against anti-paternalist principles.

Sunstein and Thaler make this point particularly strongly, arguing that the findings of behavioural economics show that paternalism is 'inevitable', that the idea that there are 'viable alternatives to paternalism' is a 'misconception', and that the anti-paternalist position is 'incoherent', a 'nonstarter' (2003b, pp. 1164-65, 1182). In its essentials, their argument runs like this. Traditionally, economics has assumed that individuals have coherent preferences which are prior to the decision situations in which they are revealed. If this assumption were true, it would make sense to ask whether those preferences should be respected; the anti-paternalist would say that they should, while the paternalist would be willing to countenance policies which overruled them. But we have learned from behavioural economics that, in many cases, individuals do not come to decision problems with pre-existing preferences. Instead, preferences are formed only when individuals confront particular problems, and those preferences are sensitive to apparently arbitrary details of framing, such as which option is presented as the default. Thus, whoever is responsible for designing the presentation of options to individuals is able to affect the preferences that those individuals will reveal; there is no way of standing back and simply 'respecting' preferences.

This provides a space in which ‘planners’ can steer individuals’ choices (p. 1162). A planner, in Sunstein and Thaler’s terminology, is ‘anyone who must design plans for others, from human resource directors to bureaucrats to kings’ [p. 1190].<sup>4</sup> Once the planner recognises this fact, she cannot avoid a decision about the direction in which to steer the individual, and the only reasonable criterion for making this decision is the planner’s judgement about the individual’s best interests. It is in this sense that there are no viable alternatives to paternalism.

Both proposals include restrictions on the scope of legitimate paternalism which are intended to make them acceptable to libertarians or conservatives. Sunstein and Thaler’s official position is that interventions should be chosen to maximise welfare, subject to the constraint that individual freedom of choice is maintained:

The libertarian aspect of our strategies lies in the straightforward insistence that, in general, people should be free to opt out of specified arrangements if they choose to do so. Hence we do not aim to defend any approach that blocks individual choices. (2003b, p. 1161)

The idea seems to be that, in choosing how to *present* options to an individual, the planner should choose a frame which steers choices in the welfare-maximising direction, but ‘bad’ options should not be removed from the opportunity set. But Sunstein and Thaler sometimes seem not to respect this self-imposed constraint. In a section entitled ‘How much choice should be offered?’, they provide the following answer to their rhetorical question:

Libertarian paternalists want to promote freedom of choice, but they need not seek to provide bad options, and among the set of reasonable ones, they need not argue that more is necessarily better. Indeed that argument is quite implausible in many contexts. In the context of savings plans, would hundreds of thousands of options be helpful? Millions? ... [O]ne recent study finds that when 401(k) plans [i.e. US retirement savings plans] offer more choice, participants are slower to join, perhaps because they are overwhelmed by the number of choices and procrastinate. (2003b, pp. 1196-1197)

They go on to discuss how a libertarian paternalist might ‘decide how much (reasonable) choice to offer’. It should be said that, in the example of savings plans, Sunstein and Thaler are thinking about the menu of plans offered by a firm to its employees: the ‘planner’ is the human resources director of the firm. In this case, they are not advocating public policies to block individual choices; they are only advising firms to restrict the number of options they offer their employees. Nevertheless, the first sentence in the quoted passage seems to be intended as a general statement of principle. Even on the most sympathetic libertarian reading, it is a less than wholehearted endorsement of freedom of choice.

Camerer et al propose a different constraint on the scope of paternalism. We are told that all the authors share ‘trepidations’ about supporting paternalistic regulations. However, it turns out that these concerns are not about the principle of imposing restrictions on individual choice, but only about the risk of miscalculation. The concerns are that ‘paternalistic policies may impose undue burdens on those people who are behaving rationally’ and that, because behavioural economics is at an early stage of development, mistakes may be made; thus, ‘caution’ is appropriate ‘at this stage’ (p. 1214). This caution is expressed in Camerer et al’s proposed criterion of *asymmetric paternalism*. This is defined in relation to a model in which there are two types of consumer, the ‘fully rational’ and the ‘boundedly rational’. Paternalistic policies confer benefits (denoted  $B$ ) on boundedly rational consumers by ‘counteract[ing] mistakes’, but may impose costs ( $C$ ) on fully rational consumers by restricting their opportunities to make optimal decisions. Camerer et al define a policy to be asymmetrically paternalistic ‘if it creates large benefits for those people who are boundedly rational ( $B$  is large) while imposing little or no harm on those who are fully rational ( $C$  is small)’ (p. 1219). In other words, paternalistic interventions are favoured only when the expected benefit/cost ratio is high. However, there is little here to reassure anti-paternalists. If benefits and costs are measured by reference to the planner’s conception of welfare, paternalistic proposals are likely to show high benefit/cost ratios. The whole point of paternalistic restrictions is to steer people towards actions which (it is claimed) they would have chosen anyway, had they been fully rational. Thus, no costs are imposed on people whom the planner deems to be fully rational. Conversely, anyone who claims to be harmed by a paternalistic policy is revealing the bounds on his rationality.

Both sets of co-authors propose that paternalistic interventions should be appraised by CBA. In Camerer et al’s account, paternalism is justified as a means of correcting ‘errors’ in decision-making (p. 1218). To say that a class of decisions are in error is to say that they ‘do not accurately reflect the benefits [the decision-makers] derive’ from them. It is a ‘crucial assumption’ that the bounds on rationality which generate such errors ‘are empirical questions subject to systematic analysis, and thus cost-benefit calculations can be made’ (pp. 1221-1222). The implication is that Camerer et al favour a form of CBA in which benefits are not measured in terms of revealed preferences; but they do not say how, even in principle, benefits *are* to be measured.

At the level of philosophical principle, Sunstein and Thaler are a little more forthcoming. With respect to any given individual, the objective of policy is to promote that

person's 'best interests'. Individuals are treated as not acting in their own best interests if their decisions are ones 'they would change if they had complete information, unlimited cognitive abilities, and no lack of willpower' (2003b, p. 1162); in such cases, the planner should try to steer them towards the choices they would have made under those conditions. In the language of moral philosophy, this is a criterion of *informed desire*. When applying such criteria, however, it is notoriously difficult to distinguish between, on the one hand, supposedly empirical claims about what a particular person would choose under specified counterfactual conditions and, on the other, normative claims about what every person *ought* to choose. It is all too easy to suppose that one's own beliefs about what is valuable in human life are supported by rational arguments, and hence that any truly rational person would share those beliefs. We might hope that Sunstein and Thaler would limit the discretion of their planner by prescribing some methodology for identifying informed desires.

Sunstein and Thaler's recommendation is that, where possible, planners should use 'a form of cost-benefit analysis'. But they provide no guidance about the methodology to be used. They say that their proposed analysis would differ from conventional CBA in *not* always using WTP measures. This is as one would expect: the whole argument is premised on the supposition that individuals do not have well-defined preferences. But all they say about the methods that *should* be used are the following two anodyne statements: '[The CBA should be] one in which a serious attempt is made to measure the costs and benefits of outcomes' (p. 1166) and 'The goal of a cost-benefit study would be to measure the full ramifications of any design choice' (p. 1190). I conjecture that, at least at the time they wrote this, Sunstein and Thaler had been unable to define a CBA methodology that they felt able to defend.

One possible interpretation is that the advocates of soft paternalism have in mind a form of CBA which uses *experienced utility* – that is, ex post hedonic experience – as the standard of cost and benefit.<sup>5</sup> This approach has been proposed by Daniel Kahneman, Peter Wakker and Rakesh Sarin (1997) and by Kahneman (2000). If hedonic experience can be measured on a one-dimensional scale (as these authors argue may be possible), this approach could be used to firm up the definitions of 'benefit' and 'informed desire' that soft paternalism needs. However, it is not at all self-evident that everything that people value in their lives can be reduced to hedonic experience.<sup>6</sup> A non-paternalistic methodology is not entitled to *assume* that individuals, if fully informed, would want to maximise experienced utility.

To sum up, the logic of the soft paternalist position allows CBA to be decoupled from the criterion of preference-satisfaction whenever individuals lack coherent preferences. In these situations, it licenses public decision-makers or their policy advisers to use their own judgements about individuals' best interests rather than trying to infer, from choices or survey responses, what those individuals in fact want. But no methodology for arriving at or justifying such judgements has yet been proposed.

Traditionally, CBA has been seen as a way of making public decision-making responsive to individuals' preferences in much the same way that competitive markets are (or have been thought to be). Many economists – I for one – have valued CBA as a barrier against paternalistic motivations in policy-making. The soft paternalist approach runs the risk of throwing away these features of CBA. Of course, if Sunstein and Thaler are right, we have no choice in the matter: there are no viable alternatives to paternalism. But *are* they right? Or can there be a non-paternalistic form of CBA which does not require us to assume that preferences are coherent?

### **3. How the market deals with incoherent preferences**

To try to convey the intuition behind my rival proposal, I begin by looking at a stylised example which Sunstein and Thaler use to illustrate libertarian paternalism. This is a story about the cafeteria at some unspecified 'organisation'. Customers proceed along a line, choosing food items from a display, until they reach the checkout. The cafeteria director notices that given items are more likely to be chosen if they are placed earlier in the line. On the basis of current medical knowledge, she judges that it would be in the best interests of most customers to eat fewer sweet desserts and more fruit. Which should she display first, the fruits or the desserts? Sunstein and Thaler claim that an anti-paternalist would advocate the rule that 'give[s] customers what [the director] thinks they would choose on their own'. But this rule is meaningful only if 'what the customer would choose' can be defined independently of the director's choice and, by hypothesis, this is not the case: consumers' preferences are context-dependent (they depend on how the items are displayed). Hence, we are told, the only reasonable rule is that the director should do what she judges to be best for consumer well-being – that is, put the fruits first. This story is intended to illustrate the claim

that there are no viable alternatives to paternalism and that anti-paternalism is incoherent (2003b, p. 1164).

One of the most surprising features of Sunstein and Thaler's treatment of this example is that they never consider the possibility that the cafeteria director might choose the display that maximises profit. Although they do not specify whether their cafeteria is operated for profit, they are using the example to illustrate a general argument in favour of paternalistic interventions; so it is surely natural to ask how the director's problem would be solved in a market. Sunstein and Thaler recognise that the director will be constrained by 'market pressures', but then offer the following reasons for not pursuing the implications of this idea:

But some of the time, market success will come not from tracking people's ex ante preferences, but from providing goods and services that turn out, in practice, to promote their welfare, all things considered. Consumers might be surprised by what they end up liking; indeed, their preferences might change as a result of consumption. And in some cases, the discipline imposed by market pressures will nonetheless allow the director a great deal of room to maneuver, because people's preferences are not well-formed across the relevant domains. (p. 1165)

The first argument here, that firms can be successful without tracking ex ante preferences, fails to recognise a significant property of markets. It is true that successful firms do not simply find out what consumers currently *think they want*, and then supply that; they anticipate what consumers *will come to want*, starting up the process of satisfying demands before consumers experience them. Market rewards come from supplying what, in the event, consumers are willing to pay for. But a profit-seeking firm has no need to make its own judgement about what is good for the consumer. As the business maxim puts it: the customer is always right. I take this maxim to mean that the customer is right *at the moment at which she buys*; a firm succeeds by predicting what she will want to buy when that moment arrives.

I cannot understand the second argument, that market pressures are less when consumers' preferences are ill-formed. The pressure of the market is not to satisfy well-formed preferences but to make profit. A firm's room to manoeuvre is the degree to which it can pursue non-profit-maximising strategies without being taken over or going bankrupt. The extent of this freedom depends (negatively) on the competitiveness of the markets in which the firm operates, not on whether its customers' preferences are coherent or incoherent.

Without adding more detail to Sunstein and Thaler's story, it is not possible to say whether a profit-maximising cafeteria would display the desserts before the fruit or vice versa.

But the problem of determining which display would lead to greater profit is well-defined, even if customers' preferences are affected by the display. The tendency for customers to favour the earlier items in the line is just one of the many regularities in their behaviour that needs to be taken into account in solving the problem. We might expect a competent and experienced retailer, working under the pressures of the market, to be able to find the profit-maximising display. Let us suppose that this display puts the desserts first. If Sunstein and Thaler's claims about the inevitability of paternalism are correct, this market solution must be deemed to be not viable, a non-starter. But on what grounds?

Since what is at issue are the general principles by which paternalism might be defended or criticised, we need to consider those properties of market solutions *in general* that are exemplified by the profit-seeking director's decision to display the desserts first. We immediately run up against the problem that standard economic analyses of the workings of markets assume that individuals have coherent preferences. The centrepieces of conventional welfare economics are the two 'fundamental theorems' which establish connections between competitive equilibrium and Pareto efficiency. Traditionally, anti-paternalist economists have appealed to these theorems as showing that competitive markets are efficient in satisfying individuals' preferences, while regulations which restrict consumers' choices introduce inefficiencies into the system. But the Pareto-efficiency criterion presupposes coherent preferences. Thus, when analysing cases in which preferences are incoherent, we cannot use the fundamental theorems to justify the allocation of resources through markets. This thought may lie behind Sunstein and Thaler's rejection of the idea of a market solution to the cafeteria problem.

Before dismissing the market solution as a non-starter, however, we should ask whether there are other ways of characterising the outcomes of market processes which do not assume preference coherence but which might still provide a basis for a normative defence of the market. I have explored one such line of enquiry in a recent paper (Sugden, 2004). The analysis centres on an idea with a long history in economics: that the market respects consumer sovereignty. Although theoretical economists have often treated this claim as merely a rhetorically-charged reference to the fundamental theorems, the idea of being sovereign seems to imply much more than that one's preferences will be respected insofar as they are coherent. If you are the sovereign in the traditional monarchical sense, the supreme source of authority is your *will*. If your will is arbitrary, capricious or unstable, that may be unfortunate for your subjects, and perhaps even for you; but its authority is not thereby

invalidated. Analogously, one might think, an economic system which respects consumer sovereignty should be understood as one which responds to consumers' demands without asking whether those demands issue from coherent preferences.

In my 2004 paper, I analyse a model of an exchange economy in which preferences need not be coherent. In this model, there are two types of agent – *consumers* and *traders*. Each consumer is endowed with some bundle of private goods, one of which goods plays the role of money. Trading takes place over an interval of time, the 'trading period'. At the start of the period, traders make public offers to buy and sell non-money goods at prices (expressed in units of money) specified by themselves, subject to specified constraints on total quantities traded. These offers are fixed for the duration of the trading period. Over the period, consumers have the continuous opportunity to take up any available offers, subject to their budget constraints. The only assumption about the rationality of consumers is that, at any given moment, each individual is *price-sensitive*. That is, if a given good is offered for sale at two different prices, the consumer does not buy at the higher price; similarly, if there are offers to buy at two different prices, he does not sell at the lower price. From one moment to another, no consistency of any kind is assumed. A consumer may even choose to buy a good at a high price at the same time as selling it at a low price. Traders are assumed to be highly rational in the pursuit of profit, which can be achieved by arbitrage – that is, by buying cheap and selling dear. Goods that have been bought or sold during the trading period are delivered only at the end of the period. This allows traders to conduct arbitrage across time, but within the trading period. There are many traders in competition with one another. An equilibrium state of the model is defined as a configuration of offers by traders such that each trader's array of offers is profit-maximising for her, given the other traders' offers and given the behaviour of consumers.

I show that, in equilibrium, each good has a single price at which each consumer is always free to buy and sell as much as he chooses, subject only to his budget constraint; all markets clear (that is, traders do not end up with positive or negative holdings of any good); and each trader makes zero profit. In these respects, an equilibrium state of the model *looks like* a Walrasian competitive equilibrium, in which the traders collectively have performed the role of the Walrasian auctioneer. But we cannot say that this market is efficient in satisfying consumers' preferences, since coherent preferences may not exist.

What we can say, however, is that the following *opportunity criterion* is satisfied. Let  $x = (x_1, \dots, x_n)$  denote an *allocation*, that is, an  $n$ -tuple of consumption bundles, one bundle for

each of the  $n$  consumers. Let  $x^*$  be the allocation that is in fact realised through trade. This allocation satisfies the opportunity criterion if, for every other feasible allocation  $x'$ , we can find some consumer  $i$  such that  $x'_i \neq x_i^*$  and such that  $i$  had the opportunity to reach  $x'_i$  by accepting some combination of the offers made to him by traders (but in fact chose to do something else). In other words, any proposal to substitute some feasible  $x'$  for the actual allocation  $x^*$  involves overruling the actual choice of some consumer  $i$ , giving him something he chose not to have in place of what he chose to have. We might say that the proposal requires  $i$ 's opportunities to be constrained. If the opportunity criterion is satisfied, every proposal for feasible change requires at least one consumer's opportunities to be constrained in this way. In this sense, the opportunity criterion is a test of efficiency in the provision of opportunity. In my 2004 paper, I show that this criterion is satisfied in the equilibrium state of the model. I shall call this result the *opportunity theorem*.

Here is another way of understanding the opportunity criterion. Suppose that, under some resource-allocation regime, the actual outcome is  $x^*$ . Consider any other feasible outcome  $x'$ . Let  $I$  be the set of consumers  $i$  for whom  $x'_i \neq x_i^*$ . Thus, a movement from  $x^*$  to  $x'$  involves some set of exchanges of private consumption goods among the members of  $I$ , with no change to the bundles held by other consumers. We can ask: Why is the outcome is  $x^*$  and not  $x'$ ? If the opportunity criterion is satisfied, the following is a valid answer: Because at least one member of  $I$  chose not to undertake his part of the set of exchanges that would have been required for a move from  $x^*$  to  $x'$ . In other words, the members of  $I$  cannot collectively object that they have been denied the opportunity to make this move by mutual consent. The essential idea behind the opportunity theorem is that, in equilibrium, no opportunities for mutually agreed exchanges remain unrealised.

This idea allows the following intuitive explanation of the opportunity theorem. Because of the assumption of price sensitivity, it is possible to define a measure of the *surplus* created for a consumer in any given transaction carried out between him and a trader. For example, suppose that, at some moment, a consumer buys one (indivisible) unit of a particular good at a market price  $p^M$ . It follows from the assumption of price sensitivity that there must be some reservation price  $p \geq p^M$  such that the consumer would have bought the good if the price had been  $p$  or less, but would not have bought if the price had been greater than  $p$ . Then the consumer's surplus on the transaction is  $p - p^M$ . Now consider a pair of transactions by which, through the intermediation of a trader, one consumer sells one unit of a good at one

moment and another consumer buys it (possibly at a different moment). Suppose consumer  $i$  sells at a price of  $q^M$ , having been willing to sell at any price greater than or equal to  $q$ , and consumer  $j$  buys at  $p^M$ , having been willing to buy at any price less than or equal to  $p$ ; obviously  $q \leq q^M$  and  $p \geq p^M$ . Then the transaction creates a surplus of  $q^M - q$  for  $i$ , a surplus of  $p - p^M$  for  $j$ , and a surplus or profit of  $p^M - q^M$  for the trader. Summing these three components, the total surplus created is  $p - q$ . Notice that this value is independent of the actual prices  $p^M$  and  $q^M$ ; it is a property of consumers' attitudes to a net trade *between themselves*, considered without reference to the intermediation of the trader. Clearly, if  $p > q$ , there are values of  $p^M$  and  $q^M$  at which the trader's profit is positive. Thus, whenever a putative exchange between consumers would generate positive surplus, there is an incentive for traders to intermediate it. In equilibrium, however, traders' profits are zero, having been driven down by competition. The foregoing argument can be extended to any set of putative exchanges among any set of consumers. For any such set of exchanges, a measure of total surplus can be defined. If this measure is positive, there is an opportunity for traders to make profits by intermediation; but competition forces profits to zero. Thus, all opportunities for mutually agreed exchanges between consumers are realised.

This property of my model is the analogue of the first fundamental theorem of welfare economics – that in competitive equilibrium, no opportunities for Pareto improvements remain unrealised. But there is a crucial difference between the two results: the fundamental theorem is expressed in terms of individuals' assumedly coherent preferences, while the opportunity theorem is expressed in terms only of price-sensitive behaviour.

This difference can be illustrated by considering a case in which individuals' preferences are systematically context-dependent. Daniel Kahneman proposes the psychological hypothesis that many individuals are subject to a *focusing illusion*: in tackling judgement and decision tasks, they give disproportionate weight to those features that are brought to their attention by the nature or framing of the task.<sup>7</sup> I conjecture that this effect is one of the causes of a form of context-dependence that I have noticed in my own preferences, and which I suspect is quite common. Immediately after I have visited a tourist site, I find that I have a desire to own goods whose value depends on my interest in, or emotional attachment to, features of that site. Thus, having visited an art gallery, I feel a desire for cards and posters of the art works I have seen; having visited a nature reserve, I feel a desire for books which provide information about the plants that grow there. Before the visit, I had no such desires. A few hours after the visit, my desire to own these goods has evaporated. If my

conjecture is right, the stimuli of the visit have induced a transient increase in the subjective significance I attach to objects that are associated with thoughts about the site. In this respect, my preferences are (depending on how we choose to describe the focusing illusion) unstable or context-dependent. Nevertheless, for a short period, there is an opportunity for surplus-creating trade between me and the suppliers of the goods I temporarily desire. In my experience, the incentives of the market ensure that these opportunities are in fact realised: there are market agents who anticipate the transient desires of tourists like me by setting up gift shops in just the places we are most likely to be at the time we feel the desires. The point of this example is that the market does not respond to my preferences only insofar as they are coherent; it responds to my desires as and when I experience them, and insofar as they are then backed up by willingness to pay.

I now return to the case of Sunstein and Thaler's cafeteria. The customers' preferences between dishes are context-dependent. In my version of the story, the director puts the desserts before the fruits because this display maximises profit; customers who might otherwise have bought fruit buy desserts. Is this an example of consumer sovereignty at work? I submit that it is. The director is seeking to create and appropriate surplus in essentially the same way as are the traders of my abstract model, or the gift-shop owners in the example of the tourist site. Presumably, the director's reason for putting the dessert first is that consumers are willing to pay more for 'dessert early in the line' than for 'dessert late in the line', to an extent that outweighs their lower willingness to pay for 'fruit late' rather than 'fruit early'. Collectively, then, consumers are willing to pay to have the dessert displayed first, and the market responds by providing this display.

It might be objected that the cafeteria customers do not consciously desire this display; in responding to their differential willingness to pay for different displays, the director is merely taking advantage of them. But it is often the case in markets that consumers are not conscious of subtle features of their own desires, while entrepreneurs recognise willingness to pay and respond to it. Think of how consumers can be willing to pay a premium for products which are attractively styled, even if they cannot articulate what makes one product more attractive to them than another. The tendency of the market is to provide people with the things that they are willing to pay for. That is consumer sovereignty.

Sunstein and Thaler might accept this conclusion but ask: What is so good about consumer sovereignty? In their example, consumer sovereignty fails to promote the best interests of consumers, as judged by the planner. If we presuppose that normative judgments

are made from the viewpoint of a planner, and that planners should be concerned with individuals' best interests, the case against consumer sovereignty has been made. But should those principles be presupposed?

The idea that normative judgements about social arrangements should be made from the viewpoint of a 'planner', 'public official' or 'social observer' is deeply embedded in economics; it reflects the discipline's roots in nineteenth-century utilitarianism. But there is another possible perspective, that of *contractarianism*.<sup>8</sup> In the contractarian perspective, there are as many viewpoints as there are members of society. Each person makes his or her own judgements about what is good for her. The role of normative assessment is not to arrive at a unified conception of what is good for society, but to find fair terms on which these separate individuals can reach agreement about how their society should be arranged. When normative analysis is understood in this way, paternalism in its purest sense is simply off-limits: no one has the standing to overrule individuals' judgements about their own interests. It remains possible that an individual might *appoint* an agent to act on his behalf, steering his choices according to principles that the individual himself endorses. Conceivably, the members of society might collectively license some public agency to impose 'paternalistic' regulations on their own choices. But if regulation of this kind is to have a contractarian justification, each individual must judge the regulation to be in his own interests. In the case of the cafeteria, it is not enough that *the director* believes that it is in the interests of the customers to be steered away from the desserts; the customers must believe this themselves (or, at least, be willing to delegate such dietary judgements to the director). Thus, justified paternalism is not made inevitable by the mere fact that individuals' preferences are context-dependent.

Still, we need to consider whether an individual who recognises that his preferences are context-dependent will see consumer sovereignty, rather than regulation by a planner, as in his best interests. I cannot prove that this is *necessarily* the case. An individual may be aware of properties of his psychology which, in particular circumstances, lead him to act contrary to what he would normally judge to be his interests; and he may be so concerned about this that he chooses to be regulated. (In terms of Sunstein and Thaler's example: a person who wants to eat a healthy diet but knows that his resolution is liable to be undermined by the sight of sugary food might choose a regulatory regime which reduces the visibility of desserts and increases that of fruits.) I make the more modest claim that it is *possible* – that it

*makes sense* – to see consumer sovereignty as in one’s interests, even in situations in which one lacks coherent preferences.

As I have explained, a regime of consumer sovereignty tends to provide each of us, at each moment, with those things that we then want and are willing and able to pay for. The question is whether each of us, reflecting about what matters to us, can conceive of this as in his or her interests. Intuitively, an affirmative answer is surely natural: what could be more obvious than that a person has an interest in being able to get what he wants, when he wants it? This idea, I suggest, captures an important part of the visceral appeal of consumer sovereignty as a rhetorical device.

Thinking more philosophically, we can distinguish between a *person* as a continuing entity through time, and a succession of that person’s *selves*; at each moment in time, a particular self is in control of the person’s actions. We might say that a regime of consumer sovereignty tends to provide each self with what *it* wants and is willing and able to pay for. Is this in the interests of the continuing person? Thinking reflectively about myself as a continuing person, I can say that consumer sovereignty is in my interests if I can *identify with* each of the selves which, at their respective moments, are doing the wanting and the paying. That is, I need to treat the transient wants of each of my possible future selves as *my* wants – even if those selves have conflicting wants. In other papers, I have tried to show that this way of thinking about one’s identity is philosophically coherent (Sugden, 2004, 2006, 2007).

#### **4. Cost-benefit analysis as market simulation**

I have argued that competitive markets are effective in meeting individuals’ demands for *private* goods, and that this claim need not presuppose the existence of coherent preferences. But this argument does not extend to *public* goods. In the context of the opportunity theorem, the problem is that a mutually beneficial multilateral contract for the supply of a public good cannot be decomposed into a set of independent bilateral contracts between individual consumers and traders. For this reason, we cannot rely on competition between profit-seeking arbitrageurs to satisfy demands for public goods.

In conventional welfare economics, public goods are recognised as giving rise to market failure. Individuals are assumed to have coherent preferences over bundles of private-good and public-good consumption, and so the concept of Pareto-efficiency in the provision of public goods is well-defined. Efficiency can be achieved if decisions about the provision

of public goods are made according to the Kaldor-Hicks compensation test (or ‘potential Pareto improvement criterion’). Since markets cannot be relied on to meet this criterion in relation to public goods, there is a role for CBA. But what if the assumption of preference coherence does not hold?

As I explained in Section 2, soft paternalists have argued for a form of CBA in which the measurement of costs and benefits can be decoupled from individuals’ preferences, as revealed in actual choices. Very little has been said about how such a CBA might be carried out, but the implication seems to be that, where preferences are incoherent, planners’ judgements about individuals’ best interests can be substituted. For anyone who is attracted by the idea of consumer sovereignty, this approach will be unappealing. But is there a feasible alternative? Is there a form of CBA that is compatible with the principle of consumer sovereignty and that does not require preferences to be coherent?

In trying to develop such a form of CBA, my starting point is the familiar idea that the purpose of CBA is to *simulate* the workings of competitive markets in situations of market failure. A typical statement of this idea can be found in the *Green Book* produced by HM Treasury, the definitive source of guidance on methods of appraisal for projects funded by central government in the UK. The section entitled ‘Valuing non-market impacts’ states:

Where market values are not available for an identified cost or benefit, there are a number of approaches to attributing a value for inclusion in an appraisal .... The preferred method of valuation is to simulate the market by estimating the ‘willingness to pay’ (WTP) or ‘willingness to accept’ (WTA) [for] a project’s outputs or outcomes.<sup>9</sup>

Implicit in the concept of market simulation is the idea that, under suitably idealised conditions, the outcomes produced by markets satisfy some normative criterion. Market simulation then uses that same criterion to make policy recommendations in situations in which the idealised conditions do not hold. In such situations, the normative criterion is applied by asking a counterfactual question of the form: Had there been an ideal market, what outcomes would have come about?

In the passage from the *Green Book*, the idea is that, in general, ‘market values’ should be used to measure costs and benefits. In this context, the normatively relevant feature of market values is that they reveal individuals’ WTP or WTA, at the margin, for the relevant goods. When goods are not traded on markets, it may still be possible to estimate individuals’ WTP or WTA. By using these estimates in place of the missing values, we simulate the

market. More generally, it can be shown that, under conditions of perfect competition and with the assumption of preference coherence, a project gives a positive profit to the firm which undertakes it, if and only if the project satisfies the Kaldor-Hicks compensation test.<sup>10</sup> The compensation test is the conventional criterion in CBA. Thus, it can be argued, the use of conventional CBA to determine decisions about the provision of public goods simulates market decision-making about private goods.

One way of defending the market-simulation approach is to identify the criterion that market allocations satisfy, and then to endorse that criterion directly as a normative principle. Thus, for example, one might appeal to arguments from welfare economics in favour of the compensation test as a measure of economic efficiency, and argue that economic efficiency is an appropriate objective for at least certain activities of government. An alternative, indirect defence starts from the observation that, in political debate in most developed countries, there is a general presumption in favour of the market as the mechanism for determining the allocation of resources with respect to private goods. If this presumption is accepted, consistency requires that the normative criterion used in decision-making about public goods should be satisfied by market allocations in normal cases involving private goods.

The problem we now have to face is that the traditional argument for CBA as market simulation depends on the assumption of preference coherence. If that assumption does not hold, we need to re-specify the normative criterion that is satisfied by market allocations under ideal conditions; and then we need to develop a form of CBA which uses that criterion. The preceding sections of this paper have responded to the first part of the problem by defining a condition that ideal market allocations satisfy: the condition that all opportunities for mutually agreed composite transactions between consumers are realised. The second part of the problem is to find a form of CBA which allows that condition to be satisfied in the provision of public goods.

In thinking about this problem, it is useful to begin by examining the properties of an ‘opportunity for a mutually agreed composite transaction’ in the context of markets for private goods. Such an opportunity is a feasible combination of transactions which generates positive surplus. The total surplus generated in any combination of putative transactions is measured in terms of the WTP of buyers at the moments at which they buy (that is, for each buyer, the highest price at which he would buy) and the WTA of sellers at the moments at which they sell (that is, for each seller, the lowest price at which she would sell). Given the assumption of price sensitivity, these WTP and WTA values are well-defined, whether or not

individuals' preferences are coherent. Thus, surplus is well-defined too. So what I have called an 'opportunity for a mutually agreed composite transaction' can usefully be called a *surplus-creating opportunity*.

It is important to recognise that the total surplus generated by a composite transaction is a sum of surpluses generated in distinct transactions, carried out by different people and perhaps at different times. Each individual may be aware only of the transaction to which he is a party, and only at the moment at which he makes it. The surplus that accrues to an individual depends on his WTP or WTA for that specific transaction at that particular moment. If the composite transaction is extended over time, some individuals may have completed their component transactions before other individuals are conscious of any desire to undertake theirs. Take the example of the tourist who, on leaving the art gallery, feels a transient desire to buy a souvenir. He may have had no prior expectation of having this desire. Even so, the gift shop stocks the goods that the tourist is now willing to pay for. Why? Because the shop-owner *anticipated* the tourist's demand. She was trading with the wholesale suppliers of these goods before the tourist knew he was going to want them.

As this example illustrates, the realisation of surplus-creating opportunities can require entrepreneurship and arbitrage. In the context of private goods, it is because the market facilitates and gives incentives for entrepreneurship and arbitrage that it is so effective at generating surplus. But the surplus itself is defined from the viewpoint of each individual *as a final consumer*. Thus, in the example, the value of the souvenir is measured by the tourist's WTP at the moment of purchase. Whether that WTP reflects a stable preference or a transient whim is of no significance for the measurement of surplus. This privileging of the viewpoint of the consumer is intrinsic to the concept of consumer sovereignty.

Now consider how the concept of a surplus-creating opportunity can be extended to public goods. Using the case of private goods as a template, the obvious way to proceed is to interpret a public good as a common source of benefits which accrue separately to each member of a group of individuals. Because the good is public, there is *non-rivalry* among the beneficiaries – that is, whatever quantity is consumed by one of them must be consumed by all. But this does not prevent us from thinking of the benefits as separable into benefits enjoyed by individuals. For example, the transmissions of a public radio station have the property of non-rivalry, but each person who tunes in can be thought of as privately enjoying her own benefits of entertainment and information. We now need to imagine a hypothetical payment mechanism by which, counterfactually, individuals can separately choose whether or

not to buy specific benefits in return for specific payments.<sup>11</sup> For any specific mechanism of this kind, given the assumption of price sensitivity, it is principle possible to define the surplus that an individual enjoys by virtue of receiving a particular benefit free of charge: this surplus is the maximum amount that the individual would be willing to pay, *within the mechanism*, for the relevant benefit. Thus, if the total cost of supplying the public good is known, we can say whether or not the policy of supplying it is a surplus-creating opportunity. Notice that there is no need to assume that preferences are coherent. What *is* necessary is to define a specific (but hypothetical) payment mechanism. I shall argue that the principle of market simulation points us towards particular kinds of hypothetical payment mechanism.

A market-simulating approach to CBA aims to identify and realise surplus-creating opportunities in situations in which the market cannot be relied on to do this. As I have emphasised, in the situations in which the market *can* be relied on, the identification of surplus-creating opportunities requires entrepreneurship and arbitrage. Surplus is defined from the viewpoint of each individual as consumer, but there has to be a mechanism for integrating these viewpoints. If CBA is to be market-simulating, the analysis has to take the viewpoint of a special kind of arbitrageur – one who is seeking mutually-beneficial multilateral composite contracts, but without being able to appropriate surplus for herself. We might say that, in place of the ‘social planner’ of traditional welfare economics and libertarian paternalism, the market-simulation approach substitutes a ‘social arbitrageur’. A social arbitrageur seeks to *anticipate* individuals’ future wants, but not to make judgements about what is good for those individuals, or about what they would want if they were more rational than they really are. Crucially, the wants that are being anticipated are wants that, when the time comes, individuals will be willing to pay to satisfy. They are not *ex post* measurements of hedonic experience, or expressions of regret about decisions that were made in the past.<sup>12</sup>

This concept of social arbitrage is already implicit in many of the practices of conventional CBA. For example, consider the methods used in the CBA of transport projects. For concreteness, consider a policy proposal to introduce congestion pricing on urban roads. The standard CBA approach would typically use a computerised theoretical model to predict changes in traffic flows over future periods, with and without road pricing. This model would include parameters to represent road-users’ subjective trade-offs between money outlays and travelling time; those parameters would have been estimated from past observations of individuals’ transport decisions. The model’s predictions would then be interpreted as points

on demand functions, and the effects of road pricing on road users would be measured as changes in consumers' surplus. Notice that this methodology uses the *analyst's* predictions about the aggregated demands of future road users. It does not ask potential road users *now* (that is, at the time the CBA is carried out) to report their monetary valuations, positive or negative, of the entire road-pricing option relative to the do-nothing alternative. Why not?

One reason is that the analysis requires only aggregate predictions about demand, and behaviour is much more predictable in the aggregate than at the level of the individual. To use individuals' own predictions about their future travel plans would introduce unnecessary noise into the calculations. Another reason is that individuals' predictions of their own responses to unfamiliar stimuli (such as new systems of road pricing) may be less reliable than forecasts derived from well-specified models. A third reason, and perhaps the strongest of all, is that individuals' predictions about the effects of a policy on a complex system (such as a road network) are likely to be particularly unreliable. It is because such effects are so difficult to predict that traffic engineers and transport economists use computer models rather than informed guesses. Significantly, arguments such as these presuppose that CBA is an attempt to value the *actual* costs and benefits of policy options, not the costs and benefits that individual citizens currently expect will flow from those options. The standard of valuation for benefits is the actual willingness-to-pay of the beneficiaries, at the moment of consumption. In other words, this kind of CBA is social arbitrage.

This conclusion can be restated in terms of hypothetical payment mechanisms. In defining the net surplus created by the road-pricing policy, the standard methodology assumes a hypothetical payment mechanism by which benefits are paid for at the moment of consumption. I have argued that the conventional CBA methodology does not investigate the valuations that individual citizens place on the policy as a whole, at the moment of the analysis. Clearly, this *could* be investigated – for example, by a suitably designed contingent valuation survey. To do so would be to define a different concept of surplus, based on a different hypothetical payment mechanism. My claim is that the principle of market simulation favours hypothetical payment mechanisms *which operate at the moment of consumption*.

## **5. How the market-simulation approach copes with anomalies: an example**

I shall now suggest that some of the problems caused by anomalies may be less severe when surplus is measured at the moment of consumption, rather than in other ways. The ideas underlying this suggestion are illustrated by the following example. I shall present the argument in a relatively informal way; a more formal treatment is given in the Appendix.

The example concerns the evaluation of some environmental bad, as it affects individuals in the role of occupiers of property. I take the case of noise nuisance. For clarity, I use the following model. There is a fixed set of houses, each with its own location in some urban area. Houses are identical in all respects except their exposure to noise. The level of noise varies across space; its level at any house can be measured on a one-dimensional scale, common to all houses. Each house is rented by an individual occupier from a property-owning firm; there are many such firms. There is some random process which generates turnover in the population: current occupiers sometimes move away and are replaced by new arrivals. Houses are rented in a competitive market, but occupiers who move from house to house may incur transaction costs. Individuals are identical to one another and are price-sensitive. For the moment, the only other assumption I make about individuals' preferences is that, other things being equal, lower levels of noise are preferred to higher levels. I do not assume preference coherence. A policy intervention – say, the construction of a new road – is predicted to impact on noise levels across the area. The CBA problem is to assign a money value (positive or negative) to this effect.

One standard method of tackling this problem is by *hedonic pricing*. The minimal assumptions I have made about preferences are sufficient to imply that, at any given time, the prices at which houses are rented are decreasing in noise exposure. If the market is sufficiently large and if there are no external shocks, we should expect the relationship between rental price and noise to be reasonably stable. If the new road is built, that will have the effect of a shock; but we should expect that, after some time, the market will settle down, again with a stable relationship between price and noise. (If the road has had a significant effect on overall noise levels, the new relationship may be different from the old one. If, for example, there is a significant reduction in the number of houses at low noise levels, the premium for low-noise property may increase.) In principle, a well-constructed forecasting model, calibrated on evidence from similar housing markets, would generate reliable estimates of the price/noise relationships for the policy option and for the do-nothing alternative. Such a forecasting model would typically be *direction-neutral*: the predicted price/noise relationships for the two scenarios would not depend on which of them was given

the ‘do-nothing’ label. Using these relationships, the effect of the policy on property-owners can be measured by the net loss of rental income, while the effect on occupiers can be measured as a change in consumers’ surplus. The sum of these effects measures the net benefit of the change in noise levels. (If the price/noise relationship remains constant, there is no change in consumers’ surplus, and so the net effect is simply the change in the rental value of property.)

The hedonic pricing method produces a long-run valuation of the effects of the road – that is, it values the effects after allowing for consumers’ responses to changes in market prices. We might expect it to understate the negative effects of the road in the period in which adjustment is taking place because, until adjustment is complete, individuals will not be living in optimal locations. (For example, some individuals who are willing to pay much more than the market premium for quietness may initially be subjected to high levels of noise; over time, such individuals will move to low-noise locations, or leave the market and be replaced by new entrants who optimise at the new prices.) An alternative and equally standard method, *contingent valuation*, is capable of discovering non-marginal valuations.

Suppose that, before the policy decision is made, we interview a random sample of occupiers. We inform each individual about the predicted effect of the road project on noise levels at her house, and then ask her to state her valuation of this effect in money units. The sum of these valuations, scaled up by the ratio of population to sample (and adjusted to allow for any differences in response rates between occupiers of different types of property) is a measure of the net effect of the road, calculated on the assumption that individuals do not change location in response to price changes. In principle, these valuation data, combined with predictions about changes in house prices and information about the rate at which people move house, could be used to evaluate the long-run effect of the road.

How might these analyses be affected by anomalies? The question asked in the contingent valuation survey can be framed in different ways. If the effect of the road on a particular respondent is an increase in noise, the question might be framed in terms of WTA (‘What is the smallest reduction in rent which would compensate you for this additional noise?’). Or it might be framed in terms of WTP to reverse the policy proposal (‘Imagine that this increase in noise has already taken place. What is the largest increase in rent that you would be willing to pay in order to reduce noise to its current level?’) Notice that the difference between these two frames is one of *direction*. The WTA question postulates the noise level of the do-nothing alternative as the status quo, and asks the respondent to think

about a change to the new-road option. The WTP question postulates the noise level of the new-road option as the status quo, and asks the respondent to think about a change to the do-nothing option. Similarly, if the road would reduce noise for the respondent, the valuation question can be framed either in terms of WTP ('What is the largest increase in rent that you would be willing to pay for this reduction in noise?') or WTA ('Imagine that this reduction in noise has already taken place. What is the smallest reduction in rent that would compensate you for a return to current noise levels?').

If the noise effects of the road are relatively small, the conventional theory of (reference-independent) preferences implies that any differences between the valuations elicited by WTA and WTP framings should be tiny.<sup>13</sup> However, there is overwhelming evidence that, in contingent valuation surveys and in corresponding experimental designs, respondents' reported valuations are strongly influenced by reference points.<sup>14</sup> In practice, the WTA frame tends to elicit much higher reported valuations than the WTP frame. Thus, the contingent valuation approach is not direction-neutral. Viewed in relation to standard economic theory, this is an anomaly. The best available theoretical explanation for this anomaly is that individuals' preferences are *reference-dependent* – that is, an individual's preferences between given consumption bundles differ according to the individual's perceived reference point.<sup>15</sup>

Comparing the hedonic pricing and contingent valuation methods, we have arrived at an apparently puzzling conclusion. The hedonic pricing method yields a direction-neutral valuation of the noise effects of the road, while the contingent valuation method does not. But the specification of the model seems to be the same in both cases. Clearly, the (assumed) discrepancy between WTA and WTP valuations in the contingent valuation study implies that individuals' preferences do not satisfy standard coherence conditions; but, in analysing the workings of the hedonic pricing method, I did not assume coherence. Still, the hedonic pricing method has washed out the direction-specific effects which appeared in the contingent valuation method. How can this be?

The (assumed) direction-neutrality of the hedonic pricing method stems from the assumption that the price/noise relationship in the housing market depends only on *current* causal factors, such as the relative numbers of houses at different noise levels, the characteristics of the population of potential occupiers, and so on; it is assumed to be unaffected by whether noise levels in the recent past were higher or lower than they are now. Is it credible to maintain this assumption while accepting the evidence of WTA/WTP

disparities? I think so, for the following reason. Both methods aim to elicit individuals' valuations of *flows* of benefits (positive or negative) which occur over time. But the contingent valuation method elicits these valuations *at a single point in time*, namely the moment at which the respondent answers the relevant survey question. In contrast, the hedonic pricing method elicits valuations *over time*: valuations are inferred from the transactions that individuals make (or choose not to make) as the market continues to operate. Thus, the hedonic pricing method takes account of the effects of endogenous changes in reference points over time, while the contingent valuation method does not (unless the respondent herself anticipates them).

It is a striking feature of the experimental evidence that individuals' reference points are easily displaced. For a subject who comes into an experimental laboratory and is offered the opportunity to buy some good, say chocolates, the reference point is the pre-experiment state in which she had no chocolates. But if the same subject is given the chocolates at the start of the experiment, and then a few minutes later is offered the opportunity to sell them, her reference point has already shifted to include the chocolates (Kahneman et al, 1990; Bateman et al, 1997). The implication is that reference points adapt readily to changes in a person's current circumstances. When subjects are given repeated opportunities to buy and sell a good, disparities between WTA and WTP tend to be eroded (Loomes et al, 2003). Similarly, in cross-section comparisons of participants in sportscard markets, people with more market experience tend to show smaller WTA/WTP disparities (List, 2003). It seems that the salience of one's current holding of a good as a reference point dissipates as one gains experience of trade.

It would be wrong to conclude that the reference-dependence of preferences never impinges on market transactions. For example, an individual-level analysis of behaviour in housing markets has shown that, when a person moves from one US city to another, her housing decisions in the new city are influenced by prices in the old one. (Someone who has just moved from a high-price city to a low-price city is likely to choose a better-quality house than an otherwise similar person who has lived for a long time in the low-price city.) This effect disappears after the immigrant has become acclimatised to the prices of the new city (Simonsohn and Loewenstein, 2006). For my purposes, however, the crucial point is that reference points adapt to current circumstances, and this is confirmed by Simonsohn and Loewenstein's findings. Thus, in a market equilibrium, we should expect that most individuals' reference points approximately coincide with their actual consumption patterns.

It is for this reason that we should expect cross-section price/noise relationships to be direction-neutral. When we observe a housing market in equilibrium, each individual's chosen price/noise combination can be interpreted as a decision to stay at that combination rather than switch to another, evaluated in relation to *that combination* as reference point. Thus, the cross-section price/noise relationship is not defined in relation to exogenously given reference points. Rather, it is defined in a way that makes reference points endogenous.

In equilibrium, the market premium for quietness measures *both* WTP for quietness by a marginal occupier of a quiet house *and* WTA for noise by a marginal occupier of a noisy house. To say that an individual is 'marginal' is to say that a small change in the premium would induce her to make a different choice: if the premium increased slightly, the marginal occupier (or would-be occupier) of the quiet house would instead choose a noisy one, while if it decreased slightly, the marginal occupier of the noisy house would instead choose a quiet one. Notice that the individuals who are most likely to be marginal are those who, either because they are insensitive to perceptions of gains and loss or because their reference points are not well-defined, have preferences which exhibit relatively little loss aversion. For example, at any time, there will be a proportion of new entrants to the housing market, for whom neither level of noise is the status quo. For this reason, new entrants may be particularly likely to be marginal.

It must be admitted that this analysis raises the theoretical possibility of multiple equilibria. If the agents in a market have reference-dependent preferences, and if reference-dependence takes the usual form of aversion to movements away from reference points, the overall effect is to reduce the volume of trade (relative to what would occur if preferences were reference-independent). It is conceivable that the market could be so thin as to make more than one price/noise relationship compatible with equilibrium. But notice that market prices are determined by the behaviour of active traders. Provided that, at any given time, there is a sufficient number of market participants whose preferences are not reference-dependent, the behaviour of these agents will fix the equilibrium.

If this analysis is correct, the hedonic pricing method measures surplus in terms of *long-run* demand relationships – that is, demand relationships that are defined for endogenous reference points – while the contingent valuation method measures surplus in terms of *short-run* demand relationships. Thus, provided that reference points adjust sufficiently quickly to changes in an individual's circumstances, the hedonic pricing method will generate reasonably accurate measures of the surplus that accrues to consumers at the moment of

consumption. In contrast, the contingent valuation method measures surplus at the moment that the relevant survey questions are answered. I conclude that the hedonic pricing approach is more consistent with the principle of market simulation.

In this example, the hedonic pricing approach is less susceptible to anomalies caused by reference-dependence. Notice, however, that I am not claiming that the hedonic pricing approach is superior *because* it is less susceptible to anomalies. Nor am I claiming that the anomalies revealed by the contingent valuation approach are the result of error or irrationality, and not properties of ‘real’ preferences. My normative claim is that, in order to embody a commitment to consumer sovereignty, CBA should value benefits in terms of what the beneficiaries are willing to pay for them at the moment of consumption. This principle favours particular kinds of hypothetical payment mechanisms as means of measuring surplus. In particular, it favours hedonic pricing over contingent valuation. If, as I have suggested, these favoured mechanisms are less prone to anomalies, that is all to the good. But in a certain sense, it is just good luck.

## **6. Conclusion**

If the only conclusion to be drawn from the analysis in Section 5 was that hedonic pricing should be preferred to contingent valuation when both methods are feasible, that would not be particularly helpful. In the policy-making community, there is already a general consensus that revealed preference methods are more reliable than stated preference ones, even if the precise interpretation of ‘reliable’ is unclear.<sup>16</sup> Stated preference methods are generally used only in situations in which revealed preference methods are impractical. But I would prefer to interpret my analysis both more generally and more conjecturally, as suggesting a way of thinking about how to do CBA in the face of preference incoherence, and as identifying an agenda for further research.

The principle of market simulation tells us to look for measurements of surplus that are made as close as possible to the moment of consumption. The first step towards implementing this proposal is simply to recognise the significance of the distinction between valuations that are revealed at the moment of consumption and valuations that are elicited *ex ante*. This paper has largely been concerned with that step. Having recognised this distinction, we can investigate the causal mechanisms which generate differences between the two kinds of measurement. In the process, we may find that some anomalies are found in *ex*

ante valuations but are absent from, or much less significant in, valuations that are revealed at the moment of consumption. I have conjectured that this may be true of anomalies associated with reference-dependence. By studying cases in which both kinds of valuation can be elicited, we may discover general methods for estimating moment-of-consumption valuations from data on ex ante valuations. For example, investigations of the effects of market experience on WTA/WTP discrepancies show that, as experience accumulates, there is a tendency for WTP to increase and for WTA to fall, and that latter effect is stronger than the former (Loomes et al, 2003; List, 2005). This suggests that moment-of-consumption valuations might be predicted by a weighted average of ex ante WTP and ex ante WTA, with WTP having the larger weight.

At the level of principle, I have tried to show that the principle of consumer sovereignty can be understood in a way that does not presuppose that preferences are coherent, and which can be implemented in CBA. Someone who reads the evidence from behavioural economics as showing that individuals often lack coherent preferences can continue to use many of the standard methods of CBA, and can remain a robust opponent of paternalism.

### **Appendix: a simple model of a market in which consumers' preferences are reference-dependent**

This model is a simplified representation of the case discussed informally in Section 5. Consider a market in which there are  $N$  houses;  $N$  is taken to be large, so the law of large numbers can be used in analysing the model. Of these houses,  $qN$  have the attribute *quiet*; the others are *noisy*. In all other respects, houses are identical. There is a succession of time periods, with a rental market in each period. In each period, there is a market-clearing price for each type of house. The analysis is concerned only with the *premium* for quietness – that is, the amount by which the rental price of quiet houses exceeds that of noisy ones. In each period, there are  $N$  consumers, each of whom rents a house of one of the two types. After each period,  $rN$  consumers, drawn at random from those currently in the market, leave the market and are replaced by new ones, drawn at random from a large population (obviously,  $0 < r < 1$ ). Consumers are *entrants* in the period in which they join the market; in subsequent

periods, as long as they remain in the market, they are *experienced*. To avoid complications that are orthogonal to the main line of argument, I postulate that transaction costs are zero.

Reference-dependence is represented in the following way. Each consumer  $i$  has a *base valuation* of quietness,  $v_i$ . This is interpreted as a reference-independent valuation, relevant when neither type of house can be treated as the reference point. In the population of potential consumers, base valuations have a rectangular distribution with a minimum at zero and a maximum at  $z$ . Entrants' choices between house types are determined by their base valuations: an entrant chooses a quiet house if and only if his base valuation is greater than the current market premium. The choices of an experienced consumer in any given period is influenced by the type of house he rented in the previous period. For any given experienced consumer  $i$ , his valuation of quietness is  $\alpha v_i$  if he rented a noisy house in the previous period and  $\beta v_i$  if he previously rented a quiet house, with  $\beta > 1 > \alpha > 0$ . We can think of  $\alpha v_i$  as a WTP valuation of quietness (the consumer is willing to pay this amount to upgrade from noisy to quiet) and  $\beta v_i$  as a WTA valuation of quietness (he is willing to accept this amount as compensation for downgrading from quiet to noisy).

In each period  $t$ , the supply of the quietness attribute is  $qN$ . In general, the demand for quietness depends on how, in period  $t - 1$ , consumers with different base valuations were distributed between the two types of houses. However, we can define *long-run equilibrium* as a price and a corresponding pattern of consumption such that, if they occur in any period  $t - 1$ , they will be reproduced in period  $t$ . This equilibrium is straightforward to characterise. In long-run equilibrium, in each period, quietness is purchased by those  $qN$  consumers who have the highest base valuations. That is, consumers with base valuations greater than  $z(1 - q)$  buy quietness, while those with lower base valuations do not. This requires that the market premium for quietness is  $z(1 - q)$ . To understand the nature of this equilibrium, suppose that all this is true of some period  $t - 1$ . Then in period  $t$ , there are  $(1 - r)(1 - q)N$  experienced consumers who did not buy quietness in period  $t - 1$ , and so have WTP valuations of quietness in period  $t$ ; these valuations have a rectangular distribution over the range from zero to  $\alpha z(1 - q)$ . There are  $(1 - r)qN$  experienced consumers who bought quietness in period  $t - 1$ , and so have WTA valuations of quietness in period  $t$ ; these valuations have a rectangular distribution over the range from  $\beta z(1 - q)$  to  $\beta z$ . And there are  $rN$  entrants whose base valuations have a rectangular distribution over the range from 0 to  $z$ . The valuations of these three classes of consumer fix the (short-run) demand function for quietness. The supply is  $qN$ . The only market-clearing premium is  $z(1 - q)$ . Facing this price, all experienced

consumers choose to repeat the choices they made in period  $t$ , while entrants distribute themselves between quiet and noisy according to whether their base valuations are greater than or less than  $z(1 - q)$ . Thus, the property that quietness is purchased by those consumers with the highest base valuations is conserved.

Long-run equilibrium is illustrated in Figure 1. The broken line  $D_1$  is the short-run demand function of entrants. The broken line  $D_2$  is the short-run demand function for experienced consumers, defined in relation to reference points set by consumption in the previous period. The kinks in  $D_2$  at the quantity  $(1 - r)qN$  occur because WTA valuations are relevant at quantities lower than this, while WTP valuations are relevant at higher quantities.  $D_T$ , the horizontal sum of  $D_1$  and  $D_2$ , is the total short-run demand function for the market.  $D_L$ , the *long-run demand function*, shows what total market demand would be if all  $N$  consumers acted on their base valuations. Notice that long-run demand is reference-independent. The long-run equilibrium price,  $z(1 - q)$ , is determined by the intersection of  $D_L$  with the supply function, i.e. with the vertical line at  $qN$ . In long-run equilibrium, the short-run demand function  $D_T$  also passes through this intersection. At the long-run equilibrium price, entrants buy the quantity  $rqN$  while experienced consumers buy  $(1 - r)qN$ .

Now suppose we start from some period  $t$  in which the supply of quietness is  $qN$  and the market is in long-run equilibrium with the price  $p' = z(1 - q')$ . Then some policy is proposed which, if implemented, will impose noise on some previously quiet houses, reducing the supply of quietness to  $q''N$ . This case is illustrated in Figure 2. As before,  $D_L$  is the long-run demand function.  $D_T'$  is the short-run demand function corresponding with the initial equilibrium. The long-run effect of the policy is to increase the price of quietness to  $p'' = z(1 - q'')$  and to shift the short-run demand function to  $D_T''$ . On the assumption that consumers do not anticipate changes in their reference points, the adjustment to the new equilibrium occurs in the following way. In period  $t + 1$  (the first period after the implementation of the policy), the relevant demand function is  $D_T'$ , and so the price increases to a level above  $p''$ . (Intuitively: since there has to be a reduction in the consumption of quietness, the WTA valuations of experienced consumers who enjoyed quietness in period  $t$  have become relevant for the determination of the market price.) Because the price in period  $t + 1$  is greater than  $p''$ , the proportion of entrants who buy quietness is less than  $q''$ , while the opposite is true for experienced consumers. (Intuitively: entrants have reference-independent valuations, which tend to be outbid by the WTA valuations of experienced consumers.) But

this implies that, among those consumers who leave the market at the end of period  $t + 1$ , the proportion who are vacating quiet houses is greater than  $q''$ , while among those who enter the market in period  $t + 2$ , the proportion who would buy quietness at the price prevailing in period  $t + 1$  is less than  $q''$ . If the price in period  $t + 2$  were to be the same as in period  $t + 1$ , quietness would be in excess supply; so the price must fall. This process continues until the price falls to the long-run equilibrium value of  $p''$ .

What is the net change in surplus as a result of the policy? *In the short run*, the loss of surplus is given by the area of the vertical strip below the initial short-run demand function  $D_T'$  and bounded by the quantities  $q''N$  and  $qN$ . This is the measure of loss that is generated by the hedonic pricing method, since that method elicits individuals' valuations using the initial equilibrium as the reference point. Notice that this measure is not direction-neutral. Instead, it takes account of the reference-dependent WTA valuations of quietness held by experienced consumers who, having enjoyed quietness in the previous period, are reluctant to give it up. Were the policy to be reversed (starting from a long-run equilibrium in which the supply of quietness is  $q''N$ ), the gain in short-run surplus would be the area below  $D_T''$ , which takes account of the relatively low WTP valuations held by experienced consumers who, *not* having enjoyed quietness in the previous period, are now reluctant to buy it.

*In the long run*, the loss of surplus per period is equal to the area of the vertical strip below  $D_L$ . This is the measure of loss that is generated by the hedonic pricing method. This measure is direction-neutral: were the policy to be reversed, there would be an equal and opposite gain of long-run surplus. If reference points adjust sufficiently quickly to changes in consumption, long-run surplus provides an appropriate measure of the per-period effects of the policy. Conversely, if reference points adjust very slowly, or if the discount rate is very high, short-run surplus becomes a more relevant measure.

## Notes

1. For surveys of 'behavioural finance' and 'behavioural industrial organisation', see Barberis and Thaler (2003) and Ellison (2007).
2. Sugden (1999) reviews the evidence of anomalies in relation to stated-preference studies.

3. By ‘severally’, I mean that the regulation acts on each individual independently; regulations that are intended to deal with externalities or to enforce contributions to public goods are not paternalistic.
4. Through personal communication, I understand that Sunstein and Thaler now prefer the term ‘choice architect’ rather than ‘planner’. They do not want this role to be identified with the traditional ‘social planner’ of welfare economics. Their idea, as I understand it, is that in a mature market economy, the task of ‘designing plans for others’ is part of the job description of many people, working in both the public and private sectors. The human resource director of a private firm is one of their paradigm examples.
5. Camerer et al point to evidence that ‘not all “risk preferences” that manifest themselves in choice behavior seem to be fully rational in the sense of maximizing *experienced* welfare’ and interpret this as a potential warrant for paternalism (p. 1217, note 21). When discussing loss aversion, they say that, in many contexts, ‘the degree of loss aversion exhibited in people’s choices seems inconsistent with their actual experiences of gains and losses’, and offer this as another justification for paternalism (p. 1218). These passages suggest that Camerer et al are defining ‘error’ as failure to maximise experienced utility. In private communication, Thaler has suggested that experienced utility might provide a criterion of informed desire.
6. For more on this, see Loewenstein and Ubel (2006). Extreme mountaineering, as discussed by Loewenstein (1999), provides a case in point: this is an activity that is highly valued by participants, but the evidence (particularly from self-reports) suggests that the balance of hedonic experience is strongly negative. Closer to home for most readers, I suspect that the pursuit of academic ambition often has similar characteristics on a smaller scale.
7. This idea is explained in Kahneman and Sugden (2005); the idea itself is Kahneman’s.
8. I am referring to contractarianism of the kind espoused by Buchanan (1975). Buchanan is strongly influenced by the much earlier work of Wicksell (1896/ 1958).
9. *Green Book*, Annex 2. The *Green Book* can be accessed online at <http://greenbook.treasury.gov.uk>.
10. When pecuniary externalities are taken into account, it is not necessarily the case that no one loses as a result of a privately-profitable project; all one can show is that, hypothetically, the gainers could fully compensate the losers without becoming net losers themselves.

11. A classic example is the mechanism proposed by Erik Lindahl (1919/ 1958). In this mechanism, each individual  $i$  is assigned a given share  $s_i$  of the costs of supplying a public good; these shares sum to 1 across all individuals. Each  $i$  then chooses the quantity  $q_i$  that he wants to be supplied, given this cost-sharing rule. There is a *Lindahl equilibrium* if the cost shares  $s_i$  are such that all individuals choose the same quantity.
12. In the market, arbitrageurs are rewarded for correct anticipations of consumers' demands, and penalised for incorrect ones. It might be possible to design mechanisms of accountability for public decision-makers which, at least imperfectly, simulate this incentive property of markets. But this issue is beyond the scope of the present paper.
13. For a formal analysis, and an explanation of what 'tiny' means, see Sugden (1999).
14. Horowitz and McConnell (2003) review a very large number of studies which have elicited comparable WTA and WTP valuations. Using the theoretical analysis of Sugden (1999), they show that observed WTA/WTP ratios are much too large to be explained by income effects.
15. The first comprehensive theory of reference-dependent preferences over consumption bundles is due to Tversky and Kahneman (1991). Later refinements of this theory include those of Munro and Sugden (2003) and Kőszegi and Rabin (2007). An alternative approach, based on the assumption that individuals are uncertain about their own preferences, is proposed by Loomes et al (2007).
16. The Treasury *Green Book* (Annex 2: see note 6) gives the following recommendation: 'As a general rule, revealed preference methods are fairly reliable, and should be used where the relevant information can be inferred. However, they cannot estimate the value placed on an asset by people who make no direct use of it. In these circumstances, stated preference methods may be helpful.' List (2005) reports that US policy-makers are more willing to trust valuations if they have been by revealed preference rather than stated preference methods.

## References

- Barberis, Nicholas and Richard Thaler (2003). A survey of behavioral finance. In George Constantinides, M. Harris and R.M. Stulz (eds), *Handbook of the Economics of Finance*, pp. 1053-1128, Elsevier.

- Bateman, Ian, Alistair Munro, Bruce Rhodes, Chris Starmer and Robert Sugden (1997). A test of the theory of reference-dependent preferences. *Quarterly Journal of Economics* 112: 479-505.
- Buchanan, James (1975). *The Limits of Liberty*. University of Chicago Press.
- Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue and Matthew Rabin (2003). Regulation for conservatives: behavioral economics and the case for 'asymmetric paternalism'. *University of Pennsylvania Law Review* 151: 1211-1254.
- Ellison, Glenn (2007). Bounded rationality in industrial organisation. In Richard Blundell, Whitney Newey and Torsten Persson (eds), *Advances in Economics and Econometrics*; Cambridge University Press, forthcoming.
- Kahneman, Daniel (2000). Experienced utility and objective happiness: a moment-based approach. In Amos Tversky and Daniel Kahneman (eds), *Choices, Values, and Frames*. Cambridge: Cambridge University Press.
- Kahneman, Daniel, Jack Knetsch and Richard Thaler (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy* 98: 1325-1348.
- Kahneman, Daniel and Robert Sugden (2005). Experienced utility as a standard of policy evaluation. *Environmental and Resource Economics* 32: 161-181.
- Kahneman, Daniel, Peter Wakker and Rakesh Sarin (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics* 112: 375-405.
- Kőszegi, Botond and Matthew Rabin (2007). A model of reference-dependent preferences. *Quarterly Journal of Economics*, forthcoming.
- Horowitz, John and Kenneth McConnell (2003). Willingness to accept, willingness to pay and the income effect. *Journal of Economic Behavior and Organisation* 51: 537-545.
- Lindahl, Erik (1919/1958). Just taxation: a positive solution. In Richard Musgrave and Alan Peacock (eds), *Classics in the Theory of Public Finance*, London: Macmillan. Original publication in German.
- List, John (2003). Does market experience eliminate market anomalies? *Quarterly Journal of Economics* 118 (1): 41-71.

- List, John (2005). Scientific numerology, preference anomalies, and environmental policy-making. *Environmental and Resource Economics* 32: 25-53.
- Loewenstein, George (1999). Because it is there: the challenge of mountaineering ... for utility theory. *Kyklos* 52: 315-344.
- Loewenstein, George and Peter Ubel (2006). Hedonic adaptation and the role of decision and experience utility in public policy. Paper prepared for conference 'Happiness and Public Economics', 22-23 September 2006, London.
- Loomes, Graham, Shepley Orr and Robert Sugden (2007). Preference uncertainty and status quo effects in consumer choice. Unpublished paper.
- Loomes, Graham, Chris Starmer and Robert Sugden (2003). Do anomalies disappear in repeated markets? *Economic Journal* 113: C153-C166.
- Munro, Alistair and Robert Sugden (2003). On the theory of reference-dependent preferences. *Journal of Economic Behavior and Organization* 50: 407-428.
- Simonsohn, Uri and George Loewenstein (2006). Mistake #37: the effect of previously encountered house prices on current housing decisions. *Economic Journal* 116: 175-199.
- Sugden, Robert (1999). Alternatives to the neoclassical theory of choice. In Ian Bateman and Ken Willis (eds), *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EC and Developing Countries*, pp. 152-180, Oxford University Press.
- Sugden, Robert (2004). The opportunity criterion: consumer sovereignty without the assumption of coherent preferences. *American Economic Review* 94: 1014-1033.
- Sugden, Robert (2006). Taking unconsidered preferences seriously. *Philosophy*, Supplement 59: 209-232.
- Sugden, Robert (2007). The value of opportunities over time when preferences are unstable. *Social Choice and Welfare*, forthcoming.
- Sunstein, Cass and Richard Thaler (2003a). Libertarian paternalism. *American Economic Review, Papers and Proceedings* 93: 175-179.
- Sunstein, Cass and Richard Thaler (2003b). Libertarian paternalism is not an oxymoron. *University of Chicago Law Review*, 70: 1159-1202.

Wicksell, Knut (1896/ 1958). A new principle of just taxation. In Richard Musgrave and Alan Peacock (eds), *Classics in the Theory of Public Finance*, London: Macmillan. Original publication in German.

**Figure 1: Long-run equilibrium**

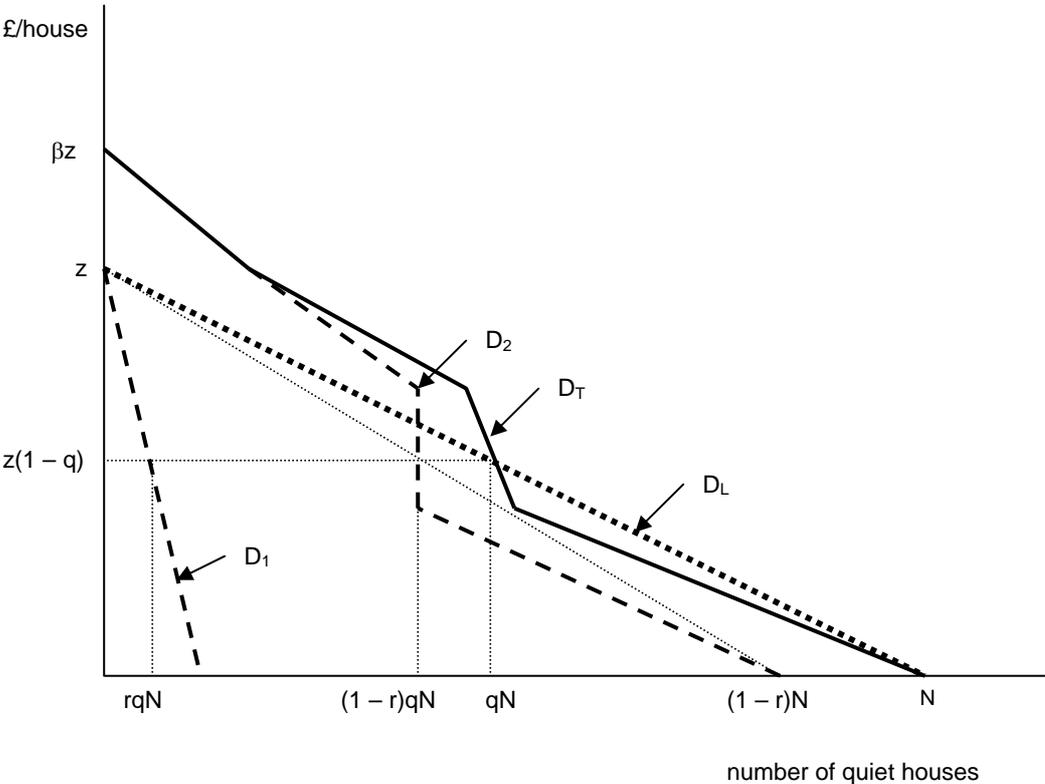


Figure 2: Measurements of surplus for a change in noise levels

