

**MEASURING WILLINGNESS-TO-PAY
FOR TRANSPORTATION IMPROVEMENTS**

Daniel McFadden

Department of Economics, University of California, Berkeley

<http://elsa.berkeley.edu/~mcfadden>

July 1, 1997

ABSTRACT:

Consumers' Willingness-to-Pay (WTP) for transportation improvements can be estimated by analyzing travel choices in real or hypothetical markets. This paper examines the measurement and analysis problems that arise in forming WTP estimates and using them to guide public policy. Bounds are given for WTP deduced from random utility models. Issues in the use of stated preference data and the design of stated preference experiments are discussed.

ACKNOWLEDGEMENTS:

This research was supported by the E. Morris Cox Endowment at the University of California, and by National Science Foundation Grant SBR 9601149. Correspondence should be directed to the author at Department of Economics, University of California, Berkeley CA 94720-3880; e-mail mcfadden@econ.berkeley.edu.

KEYWORDS:

Choice Models, Travel Demand Models, Consumer Surplus, GEV Models, Random Utility Models, Willingness to pay, Contingent Valuation.

MEASURING WILLINGNESS-TO-PAY FOR TRANSPORTATION IMPROVEMENTS

Daniel McFadden

1. Introduction

A fundamental premise of economic science is that consumers have well-formed, stable preferences, and that their choices reflect these preferences. If this is true, then it is possible to deduce from choice behavior, or from carefully phrased direct questions about preferences, whether transportation improvements or other public policy initiatives are socially desirable. In practice, consumers make errors in perception and judgment that violate classical economic rationality. It is often still possible to use observed behavior to judge the social desirability of improvements, but additional problems arise for measurement and analysis. This paper discusses these problems for the estimation of willingness-to-pay (WTP) from choice observations in real or hypothetical markets, or from direct preference elicitation. Following common use, I will refer to revealed preference (RP) analysis of real market choice behavior and stated preference (SP) analysis of hypothetical market choice behavior. I will use the term preference scaling methods (PSM) from applied psychology to refer to various methods for measuring preferences, including SP analysis, scaling of perceptions, attitudes, and risk judgments, and direct elicitation of WTP. Included in PSM are overlapping techniques known in marketing as conjoint analysis (CA) and in resource economics as contingent valuation (CV).

Section 2 of this paper reviews the random utility model (RUM) that forms the conceptual linkage from preferences to choice, and discusses some issues that arise in applying this model to transportation behavior. Section 3 reviews the problem of computing WTP from RUM-consistent choice probabilities. Section 4 considers the conditions under which choice behavior in real or hypothetical markets will be consistent with a RUM that accurately represents underlying preferences, and examines the major issues in designing reliable PSM experiments.

2. Random Utility Models of Travel Behavior

2.1. Travel Demand Models. Travel demand models are used to forecast the impacts on the transportation system of changes in attributes of travel modes or destinations, and can also be used to estimate demands for visits to alternative sites. Travel demand models are particularly useful for analyzing the demand for quasi-public goods such as recreation that are not priced through user fees, but are indirectly priced through the time and dollar cost of travel; this is called the travel cost method (TCM) for measuring WTP. I will use the example of anglers' choice of fishing sites, and the value of restoring sites that are injured by pollution, to motivate the discussion in this paper. However, the issues that arise here are the same in more general applications such as determining WTP for new or improved shopping destinations, travel modes, or vehicles.

The problem of travel demand analysis and valuation of attribute changes is simple in concept and complex in practice. The demand for trips to a destination is influenced by the availability of travel and non-travel substitutes. When consumers are utility maximizers and improvements to a site can be expressed as an equivalent reduction in travel cost, a consumer's WTP is measured by the Hicksian consumers' surplus attached to the equivalent price change; see Hicks (1939), Diamond and McFadden (1974), Varian

(1984), and McFadden (1997). Mean WTP in the population of consumers, weighted if necessary to adjust for differences in private and social marginal utility of income, is a measure of social value.

The first complication is that different individuals facing the same menus of alternatives make different choices. Then, the theory of utility maximization and of equivalent and compensating variations that provide the foundation for consumer surplus measures must accommodate unobserved heterogeneity in tastes. It is now conventional to tackle this using the random utility model (RUM). The basic ideas of applying the RUM to transportation decisions and using it in cost-benefit analysis were introduced by McFadden (1973, 1974, 1981), Diamond and McFadden (1974), and Domencich and McFadden (1975).

A conventional RUM for recreational travel looks something like this:

$$(1) \quad \left[\begin{array}{c} \text{Utility} \\ \text{of a Site} \end{array} \right] = \left\{ \left[\begin{array}{c} \text{Disposable} \\ \text{Income} \end{array} \right] \& \left[\begin{array}{c} \text{Travel} \\ \text{Cost} \end{array} \right] \& \left[\begin{array}{c} \text{Value} \\ \text{of Time} \end{array} \right] \right\} \beta + \left[\begin{array}{c} \text{Hedonic} \\ \text{Attributes} \end{array} \right] \gamma + \epsilon,$$

where β and γ are parameters; ϵ is a disturbance that captures all unobserved heterogeneity; costs and income are deflated by a price index so that they are in real terms; and "sites" include both the no-travel option and alternative recreation destinations. The hedonic attributes of a site include components whose improvements we wish to value, such as fish abundance in the example of anglers' choice of site.

An important assumption is that the disturbance ϵ for an individual consumer is fixed, and is not altered by changes in attributes. Thus, randomness in individual tastes is absent, or suspended for the purpose of determining choice under counterfactual scenarios where the quality of some sites is improved. This is an assumption of *consumer sovereignty*, requiring that tastes are defined prior to market circumstances and invariant with respect to these circumstances. This assumption is a fundamental prerequisite for economic welfare calculations, and is also essential for identification of preferences from observations on choice behavior.

A consumer's WTP for a quality improvement in a site attribute, such as an increase in fish abundance when injury due to pollution is removed, is obtained as an equivalent variation -- the adjustment in income required to make the consumer exactly as well off after removal of the injury as she was when the injury was present. If a consumer visits an injured site despite the injury, she incurs its full impact, and her WTP for a change in the degree of injury is the ratio $(1/\beta)$ times the change. If a consumer visits an injured site only after the injury is eliminated, her WTP is less, since she was sheltered from the full impact of the injury by use of a substitute. The exact WTP calculation depends on the behavior of the disturbances ϵ ; in multinomial logit and nested multinomial logit models, these calculations are enshrined in what are called "log sum" formulas. Economically, this is a consumer surplus calculation, estimating the area behind the demand curve for the injured sites between their equivalent prices for the injured and non-injured states.

There are subtle modeling issues in the handling of hedonic site attributes. If attributes that are correlated with the degree of injury are omitted, then the coefficient γ spuriously attributes their impact on behavior to the injury. On the other hand, included attributes whose levels are influenced by the injury will capture part of the overall effect of the injury, and may not be handled appropriately when the WTP calculation is done.

2.2. Travel Times and Costs. Consider the term in the indirect utility function involving income, travel cost, and travel time. The level of the estimated coefficient β , which is inversely proportional to WTP,

depends on how travel costs are measured and how travel time is handled. Economic theory is clear: Travel costs should include all components that vary at the margin with recreational choices, and time spent traveling to and engaging in recreation should be valued in terms that reflect the individual's opportunity cost of time. Implementation of this principle introduces complications.

Individuals value time spent in different activities differently. In-vehicle travel time is not as onerous as working for most people, and they value it at a substantial discount relative to their wage. Another way of saying this is that travel time to a recreation site conveys some benefits that offset the opportunity cost of wages foregone. For time spent actually engaged in recreation, this offset may be complete. The structure of indirect utility is most easily seen by starting from a model of the direct utility associated with a fixed recreational choice,

$$(2) \quad \left[\begin{array}{c} \text{Direct utility} \\ \text{of a site} \end{array} \right] = c + G(h + (1-R)TT + (1-\delta)RT) + \left[\begin{array}{c} \text{Hedonic} \\ \text{Attributes} \end{array} \right] (c + g),$$

where c is consumption of goods other than recreation trips, h is hours of non-recreational leisure, TT is travel time, RT is on-site recreation time, and G is an increasing concave function. The coefficient R is less than one, reflecting the fact that recreational travel is less onerous than working, and the coefficient δ is near zero, reflecting the fact that recreational activities convey approximately the same benefits per hour as other forms of leisure. (If RT is the result of consumer optimization of the allocation of time between recreation and other leisure, then δ will be exactly zero.) The consumer will choose c and h to maximize this expression subject to the budget constraint $N + w'l = OPC + IVC + c$ and the time allocation constraint $H = l + h + TT + RT$, where N is non-wage income net of fixed costs, w is the wage rate, l is hours worked, OPC is out-of-pocket travel cost, IVC is the variable part of indirect vehicle cost (e.g., depreciation and maintenance due to "wear and tear"), and H is total time available. Let $z = h + (1-R)TT + (1-\delta)RT$. Using the budget and time constraints to solve for l and c in terms of z , the kernel of the consumer's problem is to choose z to maximize $-wz + G(z)$. This is accomplished by a function $z = g(w)$ that satisfies $w / G'(g(w))$, and correspondingly $l = H - R'TT - \delta'RT - g(w)$ and $c = N + wH - w'g(w) - OPC - IVC - R'w'TT - \delta'w'RT$. Substituting these in the direct utility function gives

$$(3) \quad \left[\begin{array}{c} \text{Utility} \\ \text{of a Site} \end{array} \right] = \left\{ Y - OPC - IVC - R'w'TT - \delta'w'RT \right\} + \left[\begin{array}{c} \text{Hedonic} \\ \text{Attributes} \end{array} \right] (c + g),$$

where $Y = G(g(w)) + N + wH - w'g(w)$ is "disposable income". This implies that one needs to distinguish travel time components. In particular, different travel time components such as in-vehicle time and walking time may have different coefficients, and on-site recreation time (RT) will have a much smaller coefficient than travel time.

Travel cost also causes difficulties. Economic theory says a marginal dollar is a marginal dollar, independent of the category in which it is spent. However, in reality consumers pay more attention to visible travel cost dollars, such as out-of-pocket gas and lodging expenses, than they do to less visible costs such as the variable components of maintenance and depreciation; see Hensher (1985), Manering (1983), Manering and Winston (1985), Small (1983) and Train (1986). This behavior appears to arise because people keep different categories of expenditures in different mental accounts; see Thaler (1990).

In the utility formula above, replace IVC by $\delta^\mu \text{IC}$, where IC is average indirect cost (including both fixed and variable components), μ is the fraction of average indirect costs that are variable, and δ is a behavioral parameter. For the conceptual experiment of calculating the equivalent variation in non-wage income N that keeps the consumer at the utility level with injury when the injury is removed, WTP should be calculated in terms of high-powered out-of-pocket dollars (e.g., multiplier $\$/\text{dollar}$) rather than low-powered indirect cost dollars (e.g., multiplier $\$/\text{dollar}^\mu$).

One would like to estimate the travel demand model as a function of OPC and IC to obtain consistent estimates of both δ and δ^μ . In practical travel demand analysis, OPC and IC are nearly collinear, as the detailed vehicle information necessary to measure relative variation in these variables is not collected. Then, it is necessary to use an external estimate of δ^μ ; popular choices are zero and one, while consensus estimates from the travel demand literature are roughly $\mu = 0.3$ and $\delta = 0.4$. At current price levels, OPC and IC in the U.S. are both about 14 cents per mile. If the estimates above are correct, then omitting IC from the analysis ($\delta^\mu = 0$) biases the WTP estimate downward by 12 percent, while combining OPC and IC into an overall average cost ($\delta^\mu = 1$) biases the WTP estimate upward by 100 percent.

2.3. Model Structure and Income Effects. The key restrictive features in the RUM for travel demand in (1) are that it is (i) additively separable into terms involving income, hedonic attributes, and the disturbance; (ii) linear in income; and (iii) linear in the hedonic attributes. An additional restrictive feature is introduced when the disturbances are specified to have distributions that do not depend on income or hedonic attributes. These restrictions may be relatively harmless for describing the empirical features of travel demand, as even poor specifications will often provide good within-sample fits. However, these restrictions can have a substantial impact on the correct formulas for WTP. To investigate these issues, consider a more general nonlinear specification of the direct utility function given a trip to a specific site, $U(c, h + R \text{TT}, \text{RT}, x, g)$, where c is consumption other than the recreation trip, h is leisure time, TT is recreation travel time, RT is time spent in the recreational activity, x is a vector of site attributes including those that are impacted by injury and those that are not, and g is a disturbance that is heterogeneous across individuals. In this specification, TT has a coefficient R reflecting the benefit of recreational travel relative to leisure. The consumer faces an income budget constraint $N + w'l = \text{OPC} + \text{IVC} + c$ and a time budget constraint $H = l + h + \text{TT} + \text{RT}$, as before, where N is non-wage income net of fixed costs, w is the wage rate, l is hours worked, OPC is out-of-pocket travel cost, IVC is the variable part of indirect vehicle cost, and H is total time available. The first-order conditions for maximization of direct utility subject to these constraints are

$$(4) \quad L_c U(c, h + R \text{TT}, \text{RT}, x, g) = \lambda,$$

$$L_h U(c, h + R \text{TT}, \text{RT}, x, g) = w \lambda,$$

and assuming RT can be optimized,

$$(5) \quad L_{\text{RT}} U(c, h + R \text{TT}, \text{RT}, x, g) = w \lambda.$$

The solution of these equations yields c , h , and RT as functions of TT, real non-wage income $y / N/w$, real

marginal travel cost $t = (\text{OPC} + \text{IVC})/w$, x , and g . Substituting these functions into the direct utility function yields an indirect utility function $U^\#(y - t, \text{TT}, x, g)$. Note that this function does not depend on RT , which is optimized out. This is not to say that duration of recreational activity does not influence utility level, but rather that this is already captured in the function $U^\#$. In practice, other constraints may prevent full optimization in RT . In econometric analysis of choice behavior, consistency will require modeling these constraints explicitly, rather than simply introducing the endogenous observed RT as an explanatory variable in the indirect utility function.¹

Assuming that utility is scaled so there are no problems in forming moments, one can always write

$$(6) \quad U^\#(y - t, \text{TT}, x, g) = V(y - t, \text{TT}, x) + O,$$

where $V(y - t, \text{TT}, x) = \mathbf{E}_{g, y, t, \text{TT}, x} U^\#(y - t, \text{TT}, x, g)$ and O is the residual. Then, the disturbance is additively separable, but with a distribution that will in general depend on y, t, TT, x . In the analysis below, I will often concentrate on the case of an additively separable disturbance whose distribution does *not* depend on y or x . This is quite restrictive, but essential if one is to isolate attribute changes from tastes, identify individual WTP, and give WTP an economic welfare content.

Consider the dependence of the indirect utility function on income, and the impact of this on demand behavior and WTP calculations. First, consider an indirect utility function that is additively separable and linear in income with a common coefficient across alternatives, and with a disturbance whose distribution does not depend on income,

$$(7) \quad U^\#(y - t, \text{TT}, x, g) = \alpha(y - t) + S(\text{TT}, x, g).$$

This is a case of the Deaton-Muelbauer (1980) *almost ideal demand system*, since y is income divided by a price index. An important, and restrictive, behavioral implication of this system is that income drops out of the comparison of the utilities of different alternatives. Then, income changes cannot change the probability of the event that a specific alternative has maximum utility. (Wage rate changes which alter the opportunity cost of time can change probabilities, and wage rates are empirically correlated with income, but transfers and non-wage income weaken the link.) In consumer surplus terms, this implies that there are no income effects, so that Hicksian and Marshallian surplus measures coincide. A related consequence of this linearity in income is that it is possible to aggregate preferences into a *social* or *representative* preference that generates the market demand functions using Roy's identity; see Chipman and Moore (1980, 1990).

Alternatives to the almost ideal demand system in which demand for trips or for particular sites *are* income-sensitive necessarily involve non-additive interaction between income and trip attributes in the indirect utility function. McFadden and Train (1976) discuss the direct utility structures required to obtain various patterns of income effects.

¹One method for handling RT consistently when time allocation between recreation and other leisure activities is incompletely optimized is to make trip duration a specific feature of choice alternatives, along with site, and model joint choice of site and duration.

There is no difficulty in principle in developing empirical RUM models containing non-linear and/or interactive income effects. For example, using the decomposition $U^\#(y - t, TT, x, g) = V(y - t, TT, x) + O$ and assuming that the disturbances have standard distributions that do not depend on variables of the problem, such as extreme value or generalized extreme value, leads to tractable choice probability functions. However, these income effects complicate considerably the calculation of WTP. Even additive non-linear income effects, or additive linear income effects with coefficients that differ by alternative, are sufficient to upset "log sum" formulas for WTP.

3. WTP from RUM-Consistent Choice Probabilities

3.1. Mean WTP. In this section, I discuss exact formulas and bounds for mean WTP when choices are RUM-consistent. Suppose consumers face $i = 1, \dots, J$ discrete transportation alternatives, interpreted to include the no-travel option as well as travel to alternative sites. The consumer has income y . If alternative i is chosen, there is a "generalized transportation cost" (t_i) and a vector of attributes (x_i). Included in x_i are any user fees that are collected. There may be other commodity prices in the background. If restricted to alternative i , the consumer will maximize preferences subject to this restriction and subject to the constraints imposed by the availability of time and income, as in Section 2. The result will be a level of satisfaction that can be characterized by an *indirect utility* function $U(y-t_i, x_i, g_i)$ that is homogeneous of degree zero and quasi-convex in income and prices (including those in the background), increasing in income, and non-increasing in prices. The form (1) commonly used for transportation behavior is a special case of this general RUM; see, for example, McFadden (1981). Consumers are heterogeneous in unobserved factors that influence their behavior (e.g., tastes). The effects of these unobserved factors are summarized by g_i . Let $\mathbf{t} = (t_1, \dots, t_j)$ and $\mathbf{x} = (x_1, \dots, x_j)$. Consumers choose i to maximize $U(y-t_i, x_i, g_i)$. The proportion $P_i(y, \mathbf{t}, \mathbf{x})$ of consumers choosing i is given by the probability of the event

$$(8) \quad U(y-t_i, x_i, g_i) > U(y-t_j, x_j, g_j) \text{ for } j \dots i.$$

Choice of recreational fishing trips is a typical application of this setup. The travel cost t_i will be transportation and living expenses for the trip, including opportunity cost of time; and x_i will include the observed site attributes that matter to anglers, such as site size measured by the number of accessible fishing spots, ease of access, fish abundance, facilities at the site, aesthetics, user fees, and marginal equipment costs.

Suppose a policy is adopted that raises the quality of some alternatives from *injured* to *restored* conditions. In the recreational fishing example, this might be a damage remediation program that removes pollutants and increases fish abundance. Use the notation $\mathbf{x}^1 = (x_1^1, \dots, x_j^1)$ for the attributes of sites in the *as is* case with injury, and the notation $\mathbf{x}^0 = (x_1^0, \dots, x_j^0)$ for the attributes of sites in the *but for* case with restoration. The *principle of equivalent variation* is that the *as is* income of each consumer is *lowered* in the *but for* case with restoration so that maximum utility *after* the quality improvement equals maximum utility *before* the quality improvement. Then, *mean WTP* is the *average* of these equivalent variation

income reductions in the target population.²

For *any* pair of alternatives j, m , define C_{jm} to be the *net* reduction in income combined with the quality improvement that equates the utility from alternative j with injury and the utility from alternative m without injury; i.e.,

$$(9) \quad U(y-t_j, x_j, \mathbb{1}, g_j) = U(y-t_m - C_{jm}, x_m, \mathbf{0}, g_m) .$$

Define $C = C(y, \mathbf{t}, \mathbf{x}, \mathbb{1}, \mathbf{x}_0, g)$ to be the income adjustment C that equates *maximum* utility at injured conditions to *maximum* utility at restored conditions:

$$(10) \quad \text{Max}_j U(y-t_j, x_j, \mathbb{1}, g_j) = \text{Max}_m U(y-t_m - C, x_m, \mathbf{0}, g_m) .$$

Then, C is the WTP of the individual for the quality improvement. If i is chosen before the improvement and k is chosen after the improvement, then $C = C_{ik}$ and these definitions imply $C_{ii} \neq C(y, \mathbf{t}, \mathbf{x}, \mathbb{1}, \mathbf{x}_0, g) \neq C_{kk}$. Thus, the income compensation actually required to keep a consumer's maximum utility constant is bracketed by the income compensation required to keep the utility of the *original* alternative constant and the income compensation required to keep the utility of the *final* alternative constant. The intuition for these bounds is that the ability of the consumer to substitute away from i under restored conditions may increase final utility, requiring additional income reduction to equate final utility to the initial utility level. Similarly, the ability of the consumer to substitute away from k under injury conditions may increase initial utility, requiring less income reduction to equate final utility to the initial utility level. If no alternative becomes less desirable as a result of the quality improvement, then $C_{ii} \geq 0$. If alternative k is *unaffected* by the quality improvement, then $C_{kk} = 0$. Together, these conditions imply that a consumer whose final choice k is *not* affected by the quality improvement must have $C(y, \mathbf{t}, \mathbf{x}, \mathbb{1}, \mathbf{x}_0, g) = 0$.

The distribution of the unobserved g 's in the target population will induce a distribution of income compensation levels $C(y, \mathbf{t}, \mathbf{x}, \mathbb{1}, \mathbf{x}_0, g)$. The mean of this distribution will be mean WTP for the quality improvement, $EC(y, \mathbf{t}, \mathbf{x}, \mathbb{1}, \mathbf{x}_0, g)$. The distribution of g 's also induces the choice probabilities $P_j(y, \mathbf{t}, \mathbf{x})$. Let $P_j \mathbb{1} = P_j(y, \mathbf{t}, \mathbf{x}, \mathbb{1})$ denote the injury case probability of alternative j , and $P_j \mathbf{0} = P_j(y, \mathbf{t}, \mathbf{x}, \mathbf{0})$ denote the restored case probability of this alternative when compensation is actually paid.³

²One could alternately define the *compensating variation* as the net increase in income in the case with injury that would make maximum utility before the quality improvement equal to maximum utility after the quality improvement. This is also called *willingness-to-accept* (WTA). When consumers are rational, WTP and WTA will differ only by a small income effect. In practice, when WTP and WTA are measured using stated preference data, WTA is usually much larger than WTP. This *endowment effect* may be due to consumer mistrust of hypothetical trade offers, or may arise from psychological asymmetries in the treatment of gains and losses; see Thaler (1990). The effect may also arise from strategic misrepresentation, since the consumer may see an opportunity to gain substantial rents by overstating WTA. For these reasons, valuation exercises concentrate almost entirely on WTP.

³When income effects are present, $P_j \mathbf{0}$ will differ by a small income effect from the restored case probabilities that prevail when compensation is *not* paid. When the probability of travel to each injured site has a positive income elasticity, the choice probability without compensation will be larger than the choice probability with compensation, and the right-hand bound in (11)

Evaluation of mean WTP $EC(y, \mathbf{t}, \mathbf{x}^1, \mathbf{x}^0, \mathbf{g})$ is often difficult, requiring multi- dimension numerical integration or simulation. However, bounds for mean WTP may suffice for many policy purposes. In most random utility model applications, the assumption is made that the \mathbf{g} 's enter the indirect utility function *additively*, so that $U(y-t_j, x_j, \mathbf{g}_j) = V(y-t_j, x_j) + \mathbf{g}_j$ and there is *no* interaction between income and the unobservable \mathbf{g} 's. In this case, *the income reduction C_{jj} for a fixed alternative j satisfies $V(y-t_j, x_j, \mathbf{1}) = V(y-t_j - C_{jj}, x_j, \mathbf{0})$ and does not depend on \mathbf{g} .* Then, one has the bounds

$$(11) \quad \sum_{j=1}^J P_j \mathbf{1}'' C_{jj} \leq EC(y, \mathbf{t}, \mathbf{x}^1, \mathbf{x}^0, \mathbf{g}) \leq \sum_{j=1}^J P_j \mathbf{0}'' C_{jj}.$$

If the choice probabilities depend on income, then the restored choice probabilities $P_j \mathbf{0}$ depend on the compensating income reduction, and will themselves have to be bounded to calculate the right-hand-side of (11). These bounds provide a convenient check on exact computations of population mean WTP. In applications where the quality improvement is small or induces little substitution among alternatives, an average of the bounds will often provide an adequate approximation to EC without further calculation.

The bound (11) is all that is needed to estimate mean WTP for small changes. An infinitesimal change from \mathbf{x}^1 to $\mathbf{x}^0 = \mathbf{x}^1 + \mathbf{dx}$ results in an infinitesimal change $P_j \mathbf{0} = P_j \mathbf{1} + \mathbf{P}_j$ in the choice probabilities. From the formula for the compensating income reduction for a fixed alternative j ,

$$(12) \quad C_{jj} = \sum_{j=1}^J \frac{MV(y \& t_j, x_j, \mathbf{1}) / Mx_j}{MV(y \& t_j, x_j, \mathbf{1}) / My} \mathbf{1}'' x_j + o(\mathbf{dx}).$$

Substituting this expression into the bounds above yields

$$(13) \quad EC(y, \mathbf{t}, \mathbf{x}^1, \mathbf{x}^1 + \mathbf{dx}, \mathbf{g}) \leq \sum_{j=1}^J P_j \mathbf{1}'' \frac{MV(y \& t_j, x_j, \mathbf{1}) / Mx_j}{MV(y \& t_j, x_j, \mathbf{1}) / My} \mathbf{1}'' x_j + o(\mathbf{dx}).$$

For small quality improvements, this formula will approximate closely the target population mean WTP, no matter how income influences choice. Also, note that mean WTP is linear in the fraction of the population visiting the injured sites; thus, mean WTP for small changes at unpopular sites must be very small.

3.2. Common Linear Income Effects. Consider the linear indirect utility function, $U(y-t_j, x_j, \mathbf{g}_j) = \mathbf{1}''(y-t_j) + (x_j + \mathbf{g}_j)$, with an income coefficient that is common across alternatives, and let

$$(14) \quad U^*(y, \mathbf{t}, \mathbf{x}) = \mathbf{E} \text{Max}_j U(y-t_j, x_j, \mathbf{g}_j)$$

denote expected maximum utility. McFadden (1981, 1997) establishes for this case that U^* is a representative or social utility function for this problem, with Roy's identity giving the choice probabilities,

$$(15) \quad P_i(y, \mathbf{t}, \mathbf{x}) = - \frac{MU^*(y, \mathbf{t}, \mathbf{x}) / Mt_i}{MU^*(y, \mathbf{t}, \mathbf{x}) / My},$$

holds for the uncompensated choice probability.

and mean WTP satisfying

$$(16) \quad U^*(y, \mathbf{t}, \mathbf{x} \parallel) = U^*(y - EC(y, \mathbf{t}, \mathbf{x} \parallel, \mathbf{x}_0, \mathbf{g}), \mathbf{t}, \mathbf{x}_0),$$

or

$$(17) \quad EC(y, \mathbf{t}, \mathbf{x} \parallel, \mathbf{x}_0, \mathbf{g}) = \frac{1}{\pi} \int \{E \text{Max}_i U(y \& t_i, x_i, 0, g_i) \& E \text{Max}_i U(y \& t_i, x_i \parallel, g_i)\} .$$

Thus, in the common linear income case, mean WTP can be calculated as the income adjustment needed to equate expected utility before and after the quality improvement. However, even this calculation can be difficult if expected utility does not have an analytic form.

In the common linear income case, the bound (11) reduces to

$$(18) \quad \int_{t_i=1}^J P_i \Pi''((x_i 0 - x_i \parallel) / \pi) \# EC(y, \mathbf{t}, \mathbf{x} \parallel, \mathbf{x}_0, \mathbf{g}) \# \int_{t_i=1}^J P_i 0''((x_i 0 - x_i \parallel) / \pi) ,$$

and the approximation (13) for small changes reduces to

$$(19) \quad EC(y, \mathbf{t}, \mathbf{x} \parallel, \mathbf{x} \parallel + d\mathbf{x}, \mathbf{g}) \approx \int_{t_i=1}^J P_j \Pi'' \left(\frac{\zeta}{\pi} \right) x_j .$$

These approximations will often provide a useful indication of the magnitude of mean WTP without full calculation of (17).

3.3. Nonlinear Income Effects. Next consider more general nonlinear dependence of indirect utility on income, $U(y-t_j, x_j, \mathbf{g}) = V(y-t_j, x_j \parallel) + \mathbf{g} = f(y-t_j) + (x_j \parallel + \mathbf{g})$, where the functions $f(y-t)$ are not all linear, or vary by alternative. The income compensation for a consumer who chooses i in the injury case and k in the restored case satisfies $C_{ik} = y - t_k - f_k^{-1}(f_i(y-t_i) - ((x_k 0 - x_i \parallel) + \mathbf{g} - \mathbf{g}))$. Unlike the linear in income case, C_{ik} cannot be written as a difference of utilities normalized by the marginal utility of income, so the formula for expected WTP cannot be expressed in terms of a difference in expected utilities in the injury and restored cases. In particular, because expectations are not preserved by nonlinear transformations, one cannot calculate $EC(y, \mathbf{t}, \mathbf{x} \parallel, \mathbf{x}_0, \mathbf{g})$ by solving for the income reduction that equates expected utility before and after the quality improvement. One can still define expected utility, but it will not be a social utility function that generates the market demand functions via Roy's identity, and the value C of an income reduction that keeps expected utility constant will not coincide with the WTP $EC(y, \mathbf{t}, \mathbf{x} \parallel, \mathbf{x}_0, \mathbf{g})$ derived by averaging the equivalent variation income reductions for each individual.

Calculation of $EC(y, \mathbf{t}, \mathbf{x} \parallel, \mathbf{x}_0, \mathbf{g})$ will in general require explicit consideration of the distribution of income compensations induced by the nonlinear transformation, with simultaneous or iterative solution of the random utility model for $C(y, \mathbf{t}, \mathbf{x} \parallel, \mathbf{x}_0, \mathbf{g})$. This calculation is in most cases analytically intractable; McFadden (1997) provides a simulation method that is asymptotically unbiased and can be made as accurate as is needed, and applies it to generalized extreme value (GEV) random utility models.

The general bounds (11) on mean WTP can be specialized to the case of indirect utility functions that are linear and additive in the disturbance and a (nonlinear) function of income. The income compensation C_{ij}

that equates the utility of alternative j before and after the quality improvement must then satisfy $(x_{j0} - x_{j1}) = f_j(y - t_j) - f_j(y - t_j - C_{jj})$. Using the theorem of the mean, $C_{jj} = ((x_{j0} - x_{j1}) / \mu_j)$, where μ_j is the marginal utility of income $df_j(y)/dy$ evaluated at some intermediate point between $y - t_j$ and $y - t_j - C_{jj}$. Substituting the expression for C_{jj} in (11), one obtains

$$(20) \quad \sum_{j=1}^J P_j \mu_j ((x_{j0} - x_{j1}) / \mu_j) \neq EC(y, t, p, x_0, g) \neq \sum_{j=1}^J P_j \mu_j ((x_{j0} - x_{j1}) / \mu_j) .$$

Then, substituting upper and lower bounds, respectively, on μ_j , one can use these inequalities to bound mean WTP when the dependence of the indirect utility function on income is nonlinear or varies by alternative. Again, for small quality improvements, the bounds are narrow, and establish that WTP in the target population is approximately equal to the sum over alternatives of the choice probabilities times the quality improvement for an alternative, divided by the marginal utility of income *for this alternative*.

3.4. WTP in Generalized Extreme Value Models. The family of *generalized extreme value* models, which includes multinomial logit, yields explicit forms for choice probabilities, and when combined with a common linear income effect, an explicit "log sum" form for mean WTP. Suppose indirect utility can be written as $U(y - t_i, x_i, g_i) = V(y - t_i, x_i) + g_i$, and assume that the g 's have a joint cumulative distribution function (CDF) that can be written as

$$(21) \quad F(g_1, \dots, g_J) = \exp(-H(e^{g_1}, \dots, e^{g_J})) ,$$

where $H(w_1, \dots, w_J)$ is a non-negative linear homogeneous function of $w \geq 0$, satisfying the properties that if any argument goes to $+\infty$, then H goes to $+\infty$; and the mixed partial derivatives of H exist, are continuous, and alternate in sign, with non-negative odd mixed derivatives. A function H with these properties will be termed a *GEV generating function*. McFadden (1978, Theorem 1, p. 80; 1997, Theorem 1), verifies that when the function H is a GEV generating function, (21) is a CDF whose one-dimensional marginals are extreme value distributions. Further the random utility model $u_i = v_i + g_i$ with g_i distributed $F(g)$ satisfies

$$(22) \quad E \max_i u_i = \log H(e^{v_1}, \dots, e^{v_J}) + 0.57721 ,$$

and the choice probabilities satisfy

$$(23) \quad P_i = e^{v_i} H_i(e^{v_1}, \dots, e^{v_J}) / H(e^{v_1}, \dots, e^{v_J}) .$$

The linear function $H = w_1 + \dots + w_J$ is a GEV generating function which yields the multinomial logit (MNL) model. The following result (McFadden, 1997, Lemma 2) can be applied recursively, starting from linear functions, to establish families of GEV generating functions that include nested multinomial logit models:

If sets A, B satisfy $A \cup B = \{1, \dots, J\}$, $H^A(w_A)$ and $H^B(w_B)$ are GEV generating functions in w_A and w_B , respectively, and if $s \in [0, 1]$, then $H(w) = H^A(w_A)^s + H^B(w_B)^s$ is a GEV generating function in (w_1, \dots, w_J) . Consider mean WTP when the unobserved g 's have a GEV distribution and the indirect utility function has a common linear income effect, so that $U(y - t_i, x_i, g_i) = v_i + g_i$, with $v_i = \mu(y - t_i) + \alpha(x_i)$. Let $v_{i1} = \mu(y - t_i) + \alpha(x_{i1})$ and $v_{i0} = \mu(y - t_i) + \alpha(x_{i0})$. Combine the formula (17) for mean WTP in this case with the formula (22)

for expected maximum utility in GEV models to obtain the "log sum" formula

$$(24) \quad EC(y, \mathbf{t}, \mathbf{x}, \mathbf{0}, g) = \frac{1}{\pi} \ln \left\{ \log H(e^{v_1}, \dots, e^{v_J}) + \log H(e^{v_1}, \dots, e^{v_J}) \right\} .$$

This is the formula first developed by Ben Akiva (1972), McFadden (1973), and Domencich and McFadden (1975) for the multinomial logit model, and by McFadden (1978, 1981) for the nested logit model. This formula is valid *only* when the indirect utility function has a common linear income effect and choice probabilities have the GEV form.

4. RUM-Consistent Choice Behavior

4.1. Consumer Rationality. The program outlined in Sections 2 and 3 for translating choice behavior into an acceptable measure of social value will fail if choices are not consistent with a RUM that accurately reflects preferences. This could happen because (i) consumers do not have the stable, rational preferences required to make benefit-cost analysis work; (ii) consumer information and market structure allow incentives and perceptions to confound the link from utility to choice; or (iii) true preferences are veiled by cognitive illusions or rule-driven behavior. These potential problems arise in both RP and SP analysis. Regarding the first possibility, cost-benefit analysis can tolerate a degree of irrationality, in the sense that if there is an acceptable way to aggregate values of consumers with different tastes into social value, then these same principles can be applied to intra-individual taste fluctuations. We have already seen this principle applied in deriving mean WTP from a RUM, where the randomness in utility was interpreted as interpersonal taste heterogeneity, but could equally well have been interpreted to include intra-personal randomness in tastes. However, beyond this, there is no accepted economic theory of social choice based on individual tastes. I will assume for the remainder of this paper that consumers have core preferences that are consistent with social valuation, and examine issues (ii) and (iii) in determining true preferences from RP or SP data.

4.2. Information and Incentives. Consumers on the demand side of a market can be described in terms of the information they have about the commodity, from previous experience and from the market itself, and their perceptions of uncertain product features. They can also be described in terms of their familiarity with market protocol, including trading rules, expected payoffs to strategic behavior, and the consequences of alternative decisions. The strongest case for classical rational behavior can be made when the consumer has complete information about the products in a market, has extensive experience and familiarity with market operations, and knows that her behavior cannot influence the price or quality of the product purchased. For example, one can argue that mode choices for commuters and site choices for avid anglers accurately reflect preferences for the alternatives. The more complex and heterogeneous the commodity, and the less experienced the consumer, the more difficult it is to convey and absorb the information required to understand what its consumption entails. Consumers deal with complex commodities by drawing heavily on past experience, search that provides information on market conditions (and on preferences), and perhaps a less careful analysis of the consequences of choices. When forced to make quick decisions involving unfamiliar complex commodities, consumers may fall back on more generalized experiences or ideologies, using simple screening rules. For example, some experiments in marketing suggest that when consumers must choose from a small set of multi-attribute commodities, they

weigh the trade-offs among the attributes, but when faced with large sets of commodities, they first screen on one or two attributes to reduce the choice set before beginning trade-off analysis. More generally, some psychological theories that suggest that individuals are more likely to adopt "black and white" ideologies to guide unfamiliar choices, and to weigh intermediate alternatives when making familiar ones.

An automobile is an example of a complex commodity where the information-gathering process influences choice. Potential buyers *actively* search for information on alternatives, through advertisements and motoring magazines, discussions with friends, and test drives. The depth and duration of search for information, and the order in which it is assembled, are all under the consumer's control. The weight given to different aspects of cars may depend on which are emphasized in various information sources, and on the sequence in which alternatives are screened out. The search environment involves *social interaction*, and in many channels is *visual* and *aural* rather than written. One might say that the consumer goes through a process of *constructing* preferences for autos. While a classical economist might bridle at the terminology, the presence of learning, and the dependence on history that learning implies, is neither inconsistent with rationality nor with benefit-cost analysis. However, when learning is important, preferences may appear to be context-dependent and unstable. Consequently, RP data may be accurate for simple commodities, but it may be difficult to find real markets for complex, infrequently purchased commodities in which learning has proceeded to a point where consumers' choices reflect core preferences. In principle, SP methods could do more than merely mimic a real market for a complex commodity; they could go beyond it in providing information and decision-making experience to the consumer. However, it is difficult to duplicate in the laboratory the time frame and social setting of complex consumer decisions, or the interaction between the degree to which the consumer actively initiates and controls search and the saliency of search results. Most SP experiments in practice provide abbreviated information to a passive consumer, little or no opportunity for questions or social interchange, and require quick response. I would not be surprised if consumers forced to purchase automobiles under the time and information conditions typical of these experiments behaved differently than consumers in the real auto market. Then, I believe we should not be surprised if stated preferences for innovative and complex goods appear unrealistic when held up to the yardstick of real market behavior.

An additional problem in linking preferences to behavior is strategic misrepresentation. Consumers may deliberately misstate preferences when they believe that they can influence product quality or price by their own behavior. For example, in a market such as the used car market where bid and ask prices are quoted and a transaction price is negotiated, the consumer has an incentive to understate value as a rent-seeking stratagem. The issue of strategic misrepresentation is important in SP experiments. Usually, consumers in these experiments have no *economic* incentives for misrepresentation, as it is clear that the market questions are hypothetical and will not entail purchase and consumption of the chosen commodity. In these circumstances, the consumer may find it easier to be truthful about preferences than to fabricate an alternative. On the other hand, there are no economic penalties for misrepresentation, so that factors that are insignificant when money changes hands may become important. For example, consumer preferences for less convenient environmentally safe ("green") commodities may be more positive in a hypothetical setting than in a real market, as the SP experiment makes it less costly to convince the observer (and oneself) that one is socially responsible.

4.3. Cognitive Illusions and Rule-Driven Behavior. Experiments from cognitive psychology that

address the effect of context on response suggest that preferences are *constructed* when the consumer faces a choice, and that this construction is highly sensitive to context. The *cognitive illusions* uncovered by these experiments appear to be the strongest when the consumer is presented with unfamiliar tasks, and seeks cues from context to shape an appropriate response. Table 1 lists some common cognitive illusions; many share the feature that attention is selective, and focuses on the most readily available information. Many economists argue that self-interest, habit, and discipline imposed by the market will limit the degree to which important market choices are clouded by cognitive errors. Clearly there are arenas, such as financial markets where consumers engage in frequent transactions, where this argument is quite persuasive. However, there is a broad sweep of market activity, where transactions are infrequent and inconsistencies in preferences hard to detect, and deviations from rationality are unlikely to be penalized by the market. Decisions in a hypothetical market regarding unfamiliar complex commodities will be particularly vulnerable to contamination by cognitive biases.

An example of a cognitive error that complicates RP and SP analysis is the phenomenon of mental accounting (Thaler, 1990), where consumers treat dollars differently depending on the category and direction of flow. I noted in Section 2 that consumers weigh out-of-pocket travel costs more heavily than indirect variable travel costs such as depreciation, and described in principle how acceptable calculation of WTP could proceed by identifying the relevant mental account. A problem in SP experiments is that presentation of cost information on complex commodities may lead to a different mental accounting than the analogous real market. Taking automobile choice as an example, consumers presented in an SP experiment with information on initial cost and operating cost may consider these equally salient and combine them in a single mental account. In real markets, there is a temporal and saliency mismatch between initial costs and operating costs, and consumers are likely to account for these components separately. This psychological phenomenon may explain the empirical finding that consumers appear to use irrationally high discount rates in trading off initial and operating costs of durables.

Cognitive illusions are likely to be particularly important in decisions that rely on recall of past experiences. When the consumer draws on memory and on information gathered from social interaction, availability, regression, and representativeness errors are likely to lead consumers to place too much weight on selected experiences, with inadequate allowance for statistical noise, and too little weight on background information. Thus, a few experiences with automobile brands may unjustifiably be taken as representative and permanent, and may then color perceptions. Core preferences are veiled, and measured preferences appear context- dependent and unstable.

Most market commodities are private goods, and do not carry any ethical baggage. Then, the decision to purchase at a price arguably establishes a value at least equal to this price. If, however, the consumer feels that she has an ethical entitlement or property right to the commodity, or that its purchase promotes or reduces "fairness", or induces externalities, then there is potentially a more complex relationship between purchase decisions and the values that should enter a social benefit calculation. An example of a private commodity with an ethical aspect is use of a seat belt in an automobile. Consumers may feel that freedom of choice on seat belt use is an entitlement, or may be concerned about external effects or social opprobrium. Or, the commodity may become a token for a larger ethical issue, such as the principle of self-responsibility for one's safety. The presence of ethical aspects to a commodity is also likely to make its valuation more dependent on context, since ethical issues tend to be cast in relation to past positions and

entitlements.

The use of "principles" or "rules" to control behavior can be explained by the conditioning of human behavior; see Anslie (1982) and Prelec (1990). Children acquire self-control, learning to delay gratification, until in normal adults some degree of abstinence and control becomes an end in itself. Processes of precommitment, particularly adoption of rules or principles, become an important part of behavior. Conduct in an abstract choice situation is likely to be determined by deciding which principles apply. There is no obvious reason for choices in different situations to be mutually consistent. Money is a prime example of a good where individuals develop what appear to be elaborate rules for how it is accounted for, how it can be used acceptably and unacceptably, and the process as well as the outcome of exchanges of money. "Neither a borrower nor a lender be" is a principle for conduct, not an instruction for rational behavior. Then, mental accounting may be governed not only by accessibility and saliency, but also by ethical rules. The consumer is most likely to rely on rules when there is a mismatch between cost and benefits in terms of time, saliency, or scale. Principles such as "I always use a seat belt" preempt a utilitarian calculation of the expected cost of one missed occasion to fasten your seat belt.

An important observation about rule-driven decision-making is that it does not deny the possibility of an underlying utilitarianism. It may be possible to rationalize a system of rules as a consequence of overarching preference maximization, in which the "sensible self" establishes precommitments that prevent the "indulgent self" from excesses that would have undesirable consequences later. For example, a fully rational individual will when entering a car compare the cost of the five seconds of time required to fasten her seat belt with the probability of avoiding various injuries, given the distance to be traveled and driving conditions. A "sensible self" might decide that this calculus is too difficult, or too prone to misjudgment, and conclude that the rule "always fasten your seat belt" is strategically optimal. (Most motivational psychologists would question the plausibility, or at least the relevance, of the idea of a rational "sensible self".) Even if there is a full utilitarian explanation for seat belt behavior, it may be one in which long term expectations and rule setting are paramount. Consequently, it might be quite misleading to infer WTP to avoid injury from behavior on a single trip. There is no reason for strategically optimal behavior to be tactically optimal when the purpose of strategy is to avoid strategically dangerous tactics.

4.4. Design of Stated Preference Experiments. The major issue in design of SP experiments is how to provide a context that elicits core preferences, and avoids distortions caused by cognitive illusions or strategic behavior. It is plausible that the more realistic a hypothetical market setting, the more likely that stated choice behavior will look like real choice behavior. Therefore, it is useful to point out the differences between real and hypothetical markets that are potential sources of difficulty in producing accurate forecasts from SP data. I conclude this discussion with brief reviews of several studies that have been more or less successful in using SP methods to measure values, and note the features that appear to be critical to success.

The most widespread practical development of SP methods has been in marketing, where a family of techniques called *conjoint analysis* (CA) have evolved for eliciting information on perceptions, attitudes, and preferences. The essential feature of conjoint analysis is that consumers are asked to make a sequence of choices from (or to rate or rank) sets of hypothetical alternatives. The sets and attribute levels of alternatives are typically specified using a statistical design that allows a value to be assigned to each feature of an alternative. Conjoint analysis procedures usually collect additional responses designed to measure

perceptions of commodities, and general attitudes ("values") that influence preferences. The foundation for scaling these features is weak, but they have been used with some success to individualize RUM and to detect clusters of consumer types. Traditionally, marketing applications of conjoint analysis data treated preferences as purely behavioral, and did not attempt to impose on them any assumptions of economic rationality. In recent years, it has become more common to process conjoint analysis data using models that assume the existence of economic preferences. However, the emphasis of these marketing methods is on the psychometric problem of scaling responses to achieve predictive accuracy, not on the objective of recovering preferences. McFadden (1986) surveys this literature and discusses its connection to classical consumer theory; other surveys are Green and Srinivasan (1978), Akaah and Korgaonkar (1983), and Leigh, Mackay, and Wittink (1984). There is also been a substantial literature in resource economics on CV methods, which can be viewed as a special case of CA. However, the multiple questions and experimental designs used in conjoint analysis to isolate aspects of complex commodities and reduce variance have so far not been much used in environmental economics.

Another strand of development of SP methods comes from the field of experimental game theory. This line of research has investigated a variety of market games, using players in a laboratory setting. Many of these games impose on the players the objective function of maximizing dollar payoff. The focus is then on the formation of expectations and strategies, and on the validity of alternative game solution concepts, but the experimental results also provide information on strategic misrepresentation, learning, and consistency of behavior. Some of the experiments use more complex games that are intended to mimic real markets, such as the market for natural gas, with dollar payoffs that are token shares of actual market scale. Experience with the effects of complexity, information, and learning in these experimental settings illuminates some of the phenomena observed in conjoint analysis. A particular issue that could be clarified using experimental games is the occurrence of gross misrepresentations of preferences when dollar payoffs are absent and perceptions of fairness or entitlements arise. The techniques of subject instruction and repeated play used in these experiments are also potentially useful in studying how learning leads to convergence to core preferences.

It is useful to review a few more or less successful constructed market studies, as judged by consistency with market behavior or by forecasting accuracy, and identify the features that seem to be associated with success. One can then ask whether these operating conditions are met in potential transportation applications.

1. Demand for Extended Area Telephone Service: Kridel (1990) reports on an experiment in which residents of Red Bluff, Texas were asked if they would subscribe to extended area telephone service, giving them the ability to call Dallas without incremental charges. The offering was made at a series of hypothetical monthly charges. This was a service that had been intensely discussed in a local election and petition campaign, and rulings of the Texas Public Service Commission made it clear that the service would be offered in the near future. The stated preference data were used to estimate a model of utility, which was then used to forecast the penetration of the new service and the additional calling volume it would generate. The model was used to establish the price of the new service in its subsequent offering, and the population was followed for one year to test the accuracy of the model. In this case, the model was accurate to within a few percent for both the penetration rate and the calling volume.

2. Demand for Goose-Hunting Licenses: Bishop and Heberlein (1979) conducted an experiment in which recipients of free early-season duck hunting licenses in Wisconsin were offered cash to sell their licenses, or alternately asked to state a hypothetical selling or buying price. The study also compared these values with those obtained by the "travel cost method". They found substantial discrepancies between mean willingness to buy and to sell in the hypothetical market, and between the hypothetical and cash markets, with the hypothetical mean selling price about 40 percent higher than the actual mean selling price, and the hypothetical mean buying price about one-third the actual mean selling price. The authors conclude that the reduced realism of the hypothetical offering leads to a significant bias. A companion paper by Bishop, Heberlein, and Kealy (1983) provides further evidence and reaches similar conclusions.

3. Demand for High-Speed Trains in Holland: Morikawa, Ben-Akiva, and McFadden (1989) analyze data in which the actual inter-city mode choices of Dutch travelers are observed, and in addition conjoint analysis data is collected on the demand for high-speed rail, with a variety of experimental treatments. In these experiments, the high-speed rail alternative was not familiar to most respondents, and the experimental description was quite stylized. The authors find that choice models fitted to the conjoint data are not consistent with observed choice behavior, but that after permitting the parameters for the hypothetical data to be shifted in location and scale, the real and stated preferences can be reconciled. The pattern of findings suggests that inexperienced consumers confronted with incomplete information on a commodity may make a biased imputation of unobserved attributes, and may make mistakes in weighing these attributes in comparison with observed attributes. In this application, real market data can be used to recalibrate the stated preferences so that they are consistent with revealed preferences, and add to predictive power. However, without the recalibration, the stated preferences would be substantially biased.

4. Demand for Telephone Services: Hausman and Ruud (1986) analyze ranking data from a conjoint analysis experiment, and find significant differences between the attribute tradeoffs in first ranks and subsequent ranks. They find that rescaling model parameters from subsequent rank data eliminates most, but not all, the discrepancies in preferences. Thus, apparently innocuous changes in the cognitive task induce behaviors that are inconsistent with maximization of stable classical preferences. The study suggests that SP results may be sensitive to framing of the SP task and elicitation format.

5. Demand for Wilderness Areas: Diamond and Hausman (1993) and McFadden (1994) analyze a series of experiments that examine WTP to protect wilderness areas from logging. The studies find statistically and economically significant *elicitation effects*, in which WTP from referendum questions is much higher than WTP from open-ended questions. They also find strong *embedding effects*, in which WTP fails to be sensitive to the scale of the resource being preserved. These studies conclude that current CV methods applied to problems of existence value are troubled by psychometric distortions that make them unreliable, and that there is some evidence of fundamental failure of the assumption of classical stable preferences. The primary problem is that one is dealing with an unfamiliar and extremely complex commodity for which ethical issues and rule-driven behavior are probably sufficient to obscure core preferences even if they exist.

The pattern that runs through these studies is that apparently minor deviations from standard market operating conditions, where the consumer is deciding on an exchange that will actually be executed, lead to substantial biases in behavior. In some cases, these biases are systematic, and can be corrected when there are market data available to be used for recalibration. An overall assessment is that constructed market methods provide information on preferences, but this information must be filtered to extract the preferences from surrounding noise. It is rare that the stated preferences measured by a constructed market method will coincide with true market preferences without recalibration. This suggests that for constructed market methods to have a reasonable probability of success, two reference operating conditions must be met:

1. The consumer must be fully informed on the attributes of the commodity, and either experienced in making decisions about it or trained in a manner that provides a context that resembles historical experience.
2. The exchange must parallel a real exchange sufficiently closely so that the consumer states accurately his or her intentions in the real exchange. Otherwise, there must be real market data against which stated preferences can be recalibrated.

Some transportation applications can meet these conditions, but some of the areas where SP methods would be most valuable, in dealing with new commodities and services with no close real market parallels, are also likely to be the most problematic. In a typical CA experiment, the commodities are presented in configurations that may be unfamiliar or implausible to many consumers. The consumer is a passive recipient of a linear, often brief, written or verbal description of the commodities, with limited opportunity for questions or discussion in a social setting. Response is required in an abbreviated time frame, possibly on unfamiliar scales such as semantic differentials on the importance of various attributes. As in the real purchase of an auto, learning and construction of preferences is likely to occur, but the time frame, poverty of information, and lack of economic incentives may be more likely to lead to protocols for this construction that are inappropriate; see Schkade and Payne (1993, 1994); Whittington et al (1992). One promising approach to making information-gathering in SP experiments more realistic is information acceleration, a multimedia interactive technique that gives the consumer active control of the search process, while permitting the analyst to monitor the search protocol; see Hauser *et al* (1993), Urban *et al* (1993, 1994, 1996).

5. Conclusions

The first part of this paper lays out the steps necessary to go from RUM- consistent core preferences to mean WTP, an acceptable measure of the benefits of a transportation improvement. The second part examines the question of when RP or SP experiments are likely to reveal core preferences, and the real or hypothetical market conditions and SP experimental designs which appear to be necessary to obtain reliable data on preferences.

REFERENCES

- Akaah, I.; P. Korgaonkar (1983) "An Empirical Comparison of the Predictive Validity of Self-Explicated, Huber-Hybrid, Traditional Conjoint, and Hybrid Conjoint Models", *Journal of Marketing Research*, 20, 187-197.
- Anslie, G. (1982) "Beyond Microeconomics: conflict among interests in a multiple self as a determinant of value," in J. Elster (ed) *The Multiple Self*. Cambridge: Cambridge University Press.
- Ben-Akiva, M. (1972) *The Structure of Travel Demand Models*, Ph.D. Thesis, MIT.
- Bishop, R.; T. Heberlein (1979) "Measuring the Values of Extra-Market Goods: Are Indirect Measures Biased?," *American Journal of Agricultural Economics*, 61, 926- 30.
- Bishop, R.; Heberlein, T.; Kealy, M. (1983) "Contingent Valuation of Environmental Assets: Comparisons with a Simulated Market," *Natural Resources Journal*; 23(3), 619-33.
- Chipman, J.; J. Moore (1980) "Compensating Variation, Consumer's Surplus, and Welfare", *American Economic Review*, 70, 933-949.
- Chipman, J.; J. Moore (1990) "Acceptable Indicators of Welfare Change", in J.
- Deaton, A.; J. Muellbauer (1980) *Economics and Consumer Behavior*, Cambridge University Press.
- Diamond, P.; Hausman, J. (1994) "Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives*; 8(4), 45-64.
- Diamond, P.; D. McFadden (1974) "Some uses of the expenditure function in public finance," *Journal of Public Economics*, 3-21.
- Domencich, T.; D. McFadden (1975) *Urban Travel Demand: A Behavioral Analysis*, North- Holland, Amsterdam.
- Green, P. ; Srinivasan, V. (1991) "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," *Journal of Marketing* 54, 3-19.
- Hauser, J.; Urban, G.; Weinberg, B. (1993) "How Consumers Allocate Their Time When Searching for Information," *Journal of Marketing Research* 30, 452-66.
- Hausman, J.; P. Ruud (1986) "Specifying and Testing Econometric Models for Rank- Ordered Data", *Journal of Econometrics*.
- Hensher, D. (1985) "An Econometric Model of Vehicle Use in the Household Sector," *Transportation Research* 19B (4), 303-314.
- Hicks, J. (1939) *Value and Capital*, Oxford Univ. Press.
- Kridel, D. (1990) "A Consumer Surplus Approach to Predicting Extended Area Service (EAS) Development and Stimulation Rates", *Information Economics and Policy*, 379- 390
- Leigh, T., D. Makay,; J. Summers (1984) "Reliability and Validity of Conjoint Analysis and Self-Explicated Weights: A Comparison", *Journal of Marketing Research*, 21, 456-62.
- Manering, F. (1983) "An Econometric Analysis of Vehicle Use in Multivehicle Households," *Transportation Research* 17A (3), 183-189.
- Manering, F.; C. Winston (1985) "A Dynamic Empirical Analysis of Household Vehicle Ownership and Utilization," Working Paper, Department of Transportation, MIT.
- McFadden, D. (1997) "Measuring Willingness-to-Pay in Discrete Choice Models," in J. Moore; R. Hartman (eds) *Essays in Honor of John Chipman*, forthcoming.
- McFadden, D. (1994) "Contingent valuation and social choice," *American Journal of Agricultural Economics*, 76, 689-708.
- McFadden, D. (1986) "The Choice Theory Approach to Market Research," *Marketing Science* 5, 275-297.
- McFadden, D. (1973) "Conditional logit analysis of qualitative choice behavior," in P. Zarembka (ed) *Frontiers in Econometrics*, Academic Press, New York.
- McFadden, D. (1974) "The measurement of urban travel demand," *Journal of Public Economics*, 303-328.
- McFadden, D.; K. Train (1976) "The goods/leisure tradeoff and disaggregate work trip mode choice models," *Transportation Research* 12, 349-353.
- McFadden, D. (1978) "Qualitative methods for analyzing travel behaviour of individuals," in D. Hensher; P. Stopher (eds) *Behavioural Travel Modelling*, 279- 318, Croom Helm, London.
- McFadden, D. (1981) "Econometric models of probabilistic choice," in Structural Analysis of Discrete Data with Econometric Applications, in C. Manski; D. McFadden, eds. MIT Press, Cambridge, MA.
- Morikawa, T., M. Ben-Akiva, D. McFadden (1990) "Incorporating Psychometric Data in Econometric Travel Demand Models", *Transportation Science*, under review.
- Prelec, D. (1990) "Values and principles: Some limitations on traditional economic analysis," in A. Etzioni; P. Lawrence (eds)

Perspectives on Socioeconomics. London: M.E. Sharpe.

Schkade, D.; Payne, J. (1994) "How People Respond to Contingent Valuation Questions: A Verbal Protocol Analysis of Willingness to Pay for an Environmental Regulation," *Journal of Environmental Economics and Management*; 26(1), 88-109.

Schkade, D.; Payne, J. (1993) "Where do the numbers come from? How people respond to contingent valuation questions," in J. Hausman, ed., *Contingent Valuation: A Critical Assessment*, North Holland: Amsterdam.

Small, K. (1983) "Bus Priority and Congestion on Urban Expressways," in Keener, T. ed *Research in Transportation Economics* 1, 27-74.

Thalor, R. (1990) "Savings, fungability, and mental accounts," *Journal of Economic Perspectives* 4, 193-205.

Train, K. (1986) *Qualitative Choice Analysis: Econometrics and an Application to Automobile Demand*. MIT Press.

Urban, G., Qualls, W., Bohiman, J. (1994) "Information Acceleration of High Tech Industrial Products: A Market Feasibility Test," Sloan School, MIT Working Paper.

Urban, G.; Weinberg, B., Hauser, J. (1996) "Premarket Forecasting of Really-New Products," *Journal of Marketing* 60, 47-60.

Urban, G.; Hauser, J.; Roberts, J. (1993) *Design and Marketing of New Products*. Englewood Cliffs: Prentice-Hall.

Whittington, D. et al. (1992) "Giving Respondents Time to Think in Contingent Valuation Studies: A Developing Country Application," *Journal of Environmental Economics and Management*; 22(3), 205-25.

Varian, H. (1984) Microeconomic Analysis (Second Ed.), Norton, New York.