Article:
Batley, R. (2007) On ordinal utility, cardinal utility, and random utility. Theory and Decision, Online. ISSN 1573-7187

https://doi.org/10.1007/s11238-007-9046-2

Published paper

Batley, Richard (2007) *On ordinal utility, cardinal utility, and random utility.* Theory and Decision, Online First, May 2007

On ordinal utility, cardinal utility, and random utility

Richard Batley

Institute for Transport Studies, University of Leeds, UK

rbatley@its.leeds.ac.uk

ABSTRACT

Though the Random Utility Model (RUM) was conceived entirely in terms of ordinal utility, the apparatus through which it is widely practised exhibits properties of cardinal utility. The adoption of cardinal utility as a working operation of ordinal is perfectly valid, provided interpretations drawn from that operation remain faithful to ordinal utility. The paper considers whether the latter requirement holds true for several measurements commonly derived from RUM. In particular it is found that measurements of consumer surplus change may depart from ordinal utility, and exploit the cardinality inherent in the practical apparatus.

Keywords: ordinal utility, cardinal utility, Random Utility Model, log sum, rule-of-a-half

JEL Classification: B41 Economic Methodology; D01 Microeconomic Behaviour, Underlying Principles

1. INTRODUCTION

The Random Utility Model (RUM) was conceived by Marschak (1960) and Block and Marschak (1960) as a probabilistic representation of the Neo-Classical theory of individual choice, with probability deriving from variability in the individual's preferences when faced with repeated choices from the same finite choice set. Subsequent to its theoretical conception, Marschak *et al.* (1963) introduced much of the apparatus by which RUM could be translated to practice; this exploited an analogy with psychophysical models of judgement and choice, in particular Fechner's (1859) model and the derivative models of Thurstone (1927) and Luce (1959). It was left to McFadden (1968, but unpublished until 1975) to complete the translation, which was achieved through a shift in the perspective of the presentation. McFadden reconstituted RUM from a model of an individual engaged in repeated choices, to one of the choices of a population of individuals. This was clearly motivated by pragmatism; the revised presentation was more amenable to the needs of practising economists, with interests more in markets than individuals, and data to match. McFadden's adjustment to the presentation - though somewhat innocuous in theoretical terms - opened the floodgates to practical RUM analysis. There began forty years of intensive application, culminating in McFadden's awarding of the 2000 Nobel Laureate in Economics, and with it the securing of RUM's place in the history of economic methodology.

Since the unwary might, quite understandably, draw confidence from such accolade, it would seem timely to illuminate a troublesome - and indeed potentially damning - property of RUM that persists. This may be seen as a problem of compatibility arising from the analogy between Neo-Classical theory and psychophysical models. More specifically, the metric of the latter carries cardinal properties, in contrast to the ordinality of utility within Neo-Classical theory. Several previous authors have alluded to this property (e.g. Goodwin and Hensher, 1978; Daly, 1978), and many have considered its empirical consequences - usually under the label of the 'scale factor problem' (e.g. Swait and Louviere, 1993). But one might ask: 'why are RUM practitioners so preoccupied with matters of utility scale, when utility is supposedly ordinal, and ordinal utility is robust to a reasonably general class of transformations?' The purpose of the present paper is to rationalise the answer to this question in economic theory.

The paper begins by rehearsing the derivation of RUM from the fundamental axioms of deterministic choice under certainty, thereby revealing the theoretical origins of RUM in terms of ordinal utility. The paper then proceeds to demonstrate that, in exploiting the vehicle of the Fechner (1859) model, the implementation of RUM implies the adoption of cardinal utility. This is not in itself a problem; cardinal utility can be used quite defensibly as a working representation of ordinal utility, provided any such representation does not actually depart from the ordinality on which it is founded. The paper considers whether practice is adherent to the latter requirement.

2. ORDINAL UTILITY AND DETERMINISTIC CHOICE

This section rehearses the theory of deterministic choice under certainty. Whilst the general theory will be well-known to many readers, the following attends to a particular interest in finite choice sets, and this may be less familiar to some. The purpose of the rehearsal is to establish a systematic basis for section 3, which seeks to illuminate the theoretical origins of RUM in terms of ordinal utility. The informed reader may however wish to omit this section, and move directly to section 3.

Before proceeding, and following RUM convention, deference is made to Lancaster's (1966) representation of goods - or in present parlance 'alternatives' - in terms of their constituent attributes. Formally, define an alternative to be a vector $\mathbf{x} = (x_1, ..., x_K)$, where x_k is the quantity of attribute k for $k = 1, ..., K$. In words, an alternative represents a point in a finite attribute space. Again following usual practice, define within this space a finite set T of alternatives that constitutes the 'choice set' of N alternatives available to an individual: $T = \{\mathbf{x}_1, ..., \mathbf{x}_N\}$.

The purpose of the theory of deterministic choice under certainty is to represent an individual's preferences in terms of a numerical utility function. Though there is reasonable agreement as to the means of achieving this, which is to impose an axiom system on the individual's preferences, several alternative presentations have been proposed. As is often the case, this variety has been motivated, in particular, by an interest in the possible generality of presentation. The approach employed below follows Block and Marschak (1960), although the reader is referred to Luce and Suppes (1965) for discussion of alternative presentations.

The theory rests fundamentally on two axioms[1], as follows:

Axiom of Completeness:

For every pair $\mathbf{x}_n, \mathbf{x}_m \in T$, either $\mathbf{x}_n \succeq \mathbf{x}_m$, or $\mathbf{x}_m \succeq \mathbf{x}_n$, or both $\mathbf{x}_n \succeq \mathbf{x}_m$ and $\mathbf{x}_m \succeq \mathbf{x}_n$.

Axiom of Transitivity:

For every triad $\mathbf{x}_n, \mathbf{x}_m, \mathbf{x}_l \in T$, if $\mathbf{x}_n \succeq \mathbf{x}_m$ and $\mathbf{x}_m \succeq \mathbf{x}_l$, then $\mathbf{x}_n \succeq \mathbf{x}_l$

These two axioms, when taken together, establish a complete (weak) preference ordering on T. Since T is finite, each alternative $\mathbf{x}_n \in T$ can be assigned a 'rank' on the basis of preference, where a rank is an integer r_n, $1 \leq r_n \leq N$, such that for every pair within the ranking:

$$r_n \leq r_m \text{ if } \mathbf{x}_n \succeq \mathbf{x}_m$$

Now define the vector $\mathbf{r} = (r_1, \ldots, r_N)$, which is an integer-valued function representing the complete ordering of T. Any strictly decreasing monotone

[1] Noting that relaxation of the assumption of a finite choice set necessitates an additional axiom imposing continuity of preference. Continuity is discussed further in section 7.1, albeit in a different regard.

function on $\mathbf{r}$ will induce a real-valued function U on T that preserves the preference ordering, specifically:

$$U_n \geq U_m \text{ iff } \mathbf{x}_n \succeq \mathbf{x}_m \qquad (1)$$

Hence we arrive at the 'ordinal' utility function U, wherein values are assigned to the function simply on the basis that the U-value of a preferable alternative will be greater than that of an inferior alternative. Otherwise the scale of U is arbitrary, and any transformation that preserves the ordering (i.e. the class of strictly monotonic transformations) is, in these terms, entirely equivalent. Thus (1) may be re-written equivalently:

$$\hat{U}_n \geq \hat{U}_m \text{ iff } \mathbf{x}_n \succeq \mathbf{x}_m \qquad (2)$$

where $\hat{U} = f(U)$, and f is a strict monotone of U

It is important to be clear that the above theory provides no basis for drawing inferences in terms of cardinal utility; i.e. in terms of differences in utility (or any ratio thereof) across alternatives. The ordinal utility function simply ensures that an individual prefers alternatives with greater utility to those with less; i.e. that the individual acts so as to maximise utility.

3. ORDINAL UTILITY AND PROBABILISTIC CHOICE

Let us now generalise the analysis to accommodate a probabilistic representation of choice, again under certainty. This representation is motivated by the conjecture, emanating principally from the discipline of psychology, that the choices of an individual may not necessarily be 'consistent'. Inconsistency is, in this sense, usually defined in terms of either (or both) of two phenomena; first, violations of the Axiom of Transitivity; and second, the observation that an individual, when faced with a repeated choice task, may not always make the same decisions. Having introduced the theory of probabilistic choice, attention now focuses on RUM, which constitutes one particular formalisation of the theory; interested readers may wish to refer to Luce and Suppes (1965), who summarise a fuller range of formalisations. With reference to the aforementioned phenomena, RUM is motivated by an aspiration to accommodate the latter, i.e. variations in individual preferences across repetitions; this will become apparent in the subsequent analysis. Following in particular the definition given in Block and Marschak (1960), RUM offers statement of the probability of choosing an alternative $\mathbf{x}_n$ from the set T, thus:

There is a random vector $\mathbf{U} = (U_1, ..., U_N)$, unique up to an increasing monotone transformation, such that:

$$P(\mathbf{x}_n | T) = \Pr(U_n \geq U_m) \text{ for all } \mathbf{x}_n \in T,\ \mathbf{x}_m \in T,\ m \neq n \qquad (3)$$

where $0 \leq P(\mathbf{x}_n | T) \leq 1$ and $\sum_{n=1}^{N} P(\mathbf{x}_n | T) = 1$

It is crucial to make clear that RUM, in common with the theory of deterministic choice in section 2, refers to the behaviour of a single individual. Note for example Marschak's (1960) assertion: *'Unless specified to the contrary, all concepts in this paper are associated with a single given person, the 'subject' of experiments'* (p219). It thus emerges that randomness in RUM derives from the repetition, in particular the facility for variability on these repetitions, of an individual's preference ordering. On any given repetition for a given individual, a complete preference ordering is established (and $\mathbf{U}$ defined) in the manner of the deterministic theory outlined earlier. RUM, however, permits variability in the preference ordering on successive repetitions. Block and Marschak (1960) make clear this proposition by re-couching the analysis in terms of rankings. Thus define R_n to be the set of rankings on T within which the alternative $\mathbf{x}_n$ is ranked first:

$$R_n = \{\mathbf{r} | r_n \leq r_m\} \text{ for all } \mathbf{x}_n \in T,\ \mathbf{x}_m \in T,\ m \neq n \qquad (4)$$

Remembering from section 2 that the first-ranked alternative can be represented as that maximising ordinal utility, define $P_n(\mathbf{r})$ to be the probability of $\mathbf{x}_n$ being first-ranked according to a given ordering:

$$P_n(\mathbf{r}) = \Pr\left(U_n^{\mathbf{r}} \geq U_m^{\mathbf{r}} \geq \ldots \geq U_N^{\mathbf{r}}\right) \geq 0 \tag{5}$$

Then summing across all orderings where $\mathbf{x}_n$ is ranked first, the probability statement (3) can be re-stated equivalently:

$$P(\mathbf{x}_n | T) = \sum_{\mathbf{r} \in R_n} P_n(\mathbf{r}) \tag{6}$$

Worked example of ordinal RUM

Since the above interpretation of RUM may be unfamiliar to some readers, let us offer illustration by means of worked example. Consider in particular an individual choosing an alternative from the set $T = \{\mathbf{x}_n, \mathbf{x}_m, \mathbf{x}_l\}$ on five repeated occasions, such that these five observations constitute the entire sample space. With reference to (4), the possible rankings under which each alternative can be chosen are as follows:

$$R_n = \{U_n \geq U_m \geq U_l ; U_n \geq U_l \geq U_m\}$$

$$R_m = \{U_m \geq U_n \geq U_l ; U_m \geq U_l \geq U_n\}$$

$$R_l = \{U_l \geq U_n \geq U_m ; U_l \geq U_m \geq U_n\}$$

Let us now introduce some data, by assuming that the utilities and choices actually arising on the five repetitions are as given in Table 1 (and remembering of course that these utilities are defined arbitrarily, save for their ordering).

Table 1: Worked example of ordinal RUM

Repetition	U_n	U_m	U_l	*Choice*
1	1	2	3	$\mathbf{x}_l$
2	4	2	3	$\mathbf{x}_n$
3	1	2	3	$\mathbf{x}_l$
4	4	2	3	$\mathbf{x}_n$
5	1	2	3	$\mathbf{x}_l$

Two initial observations may be drawn. First, $\mathbf{x}_m$ is dominated by $\mathbf{x}_l$ and therefore never chosen. Second, whilst the utilities of both $\mathbf{x}_m$ and $\mathbf{x}_l$ are constant across repetitions, the utility of $\mathbf{x}_n$ is variable; this engenders fluctuating preferences for $\mathbf{x}_n$ and $\mathbf{x}_l$. Now applying the data of Table 1, let us calculate, for each alternative, the probability of being first ranked by each possible ordering (5) and the probability of being chosen (6), thus:

$$\Pr(U_n \geq U_m \geq U_l) = 0,\ \Pr(U_n \geq U_l \geq U_m) = 2/5, \text{ hence } P(\mathbf{x}_n) = 2/5$$

$$\Pr(U_m \geq U_n \geq U_l) = 0,\ \Pr(U_m \geq U_l \geq U_n) = 0, \text{ hence } P(\mathbf{x}_m) = 0$$

$$\Pr(U_l \geq U_n \geq U_m) = 0,\ \Pr(U_l \geq U_m \geq U_n) = 3/5, \text{ hence } P(\mathbf{x}_l) = 3/5$$

The worked example serves to reiterate that RUM is defined entirely in terms of ordinal utility[2]. Indeed, this can be shown by applying the data of Table 1 to the order-preserving (but non-linear) transformation $\hat{U}_n = U_n^2$ for all $\mathbf{x}_n \in T$, and confirming that the choices remain unchanged (Table 2).

[2] Should the sample space be extended from five repetitions to 'arbitrarily many' repetitions, then the above interpretations of ordinal utility and choice probability would apply in exactly the same manner.

Table 2: Worked example of ordinal RUM, after transformation

Repetition	$\hat{U}_n$	$\hat{U}_m$	$\hat{U}_l$	*Choice*
1	1	4	9	$\mathbf{x}_l$
2	16	4	9	$\mathbf{x}_n$
3	1	4	9	$\mathbf{x}_l$
4	16	4	9	$\mathbf{x}_n$
5	1	4	9	$\mathbf{x}_l$

More generally, and analogous to (2), the probability of being first-ranked according to a given ordering (and therefore also the probability of choice) is robust to any strictly monotonic transformation f applied to utility. Hence (5) may be rewritten:

$$P_n(\mathbf{r}) = \Pr\left(f\left(U_n^{\mathbf{r}}\right) \geq f\left(U_m^{\mathbf{r}}\right) \geq \ldots \geq f\left(U_N^{\mathbf{r}}\right)\right) \geq 0$$

Before moving on, it might be remarked that whilst deterministic transitivity continues to hold in relation to the underlying preference orderings, the probability statement of RUM can itself be characterised in terms of a stochastic representation of transitivity. The relevant property is one of 'strong stochastic transitivity' (SST), which applies to various subsets of the choice set T as follows:

If $\min\left[P(\mathbf{x}_n|\mathbf{x}_n,\mathbf{x}_m), P(\mathbf{x}_m|\mathbf{x}_m,\mathbf{x}_l)\right] \geq \frac{1}{2}$, then the binary choice probabilities satisfy strong stochastic transitivity provided:

$$P(\mathbf{x}_n|\mathbf{x}_n,\mathbf{x}_l) \geq \max\left[P(\mathbf{x}_n|\mathbf{x}_n,\mathbf{x}_m), P(\mathbf{x}_m|\mathbf{x}_m,\mathbf{x}_l)\right]$$

Discussion of other properties will be avoided here, since these are dealt with comprehensively in previous contributions, notably those of Block and Marschak (1960) and Luce and Suppes (1965).

4. CARDINAL UTILITY AND PROBABILISTIC CHOICE

Though, as section 3 has demonstrated, the theoretical definition of RUM relies entirely on ordinal utility, such metrics are not particularly amenable to the usual parametric methods of econometric modelling. Any aspiration to implement RUM in practice provokes, therefore, an inclination towards adopting cardinal utility as a working representation of ordinal. Credit for translating RUM from theory to practice is often accorded to McFadden (1968), which was indeed path-breaking in demonstrating the potential of

RUM to inform public policy[3]. It is rarely acknowledged, however, that a fundamental tenet of the practical apparatus exploited by McFadden was conceived in the earlier work of Marschak *et al.* (1963), as follows.

Marschak *et al.* considered the relationship between the binary RUM and Fechner (also referred to as 'Strong Utility') models (Fechner, 1859), where they define the latter, thus:

> A binary probabilistic model is called a Fechner model if there exist constants $v_1,...,v_N$ and a non-decreasing real-valued function $\phi(q)$ defined for all real numbers q, such that for every binary offered set $S = \{\mathbf{x}_n, \mathbf{x}_m\} \subseteq T$,
>
> $$P(\mathbf{x}_n|S) = \phi(v_n - v_m)$$

As Marschak *et al.* note, if the usual conditions $P(\mathbf{x}_n|S) + P(\mathbf{x}_m|S) = 1$ apply, then $\phi(q)$ is restricted at a finite number of values. In particular, $\phi(q)$ can be specified as a continuous non-decreasing function such that:

$$\lim_{q \to -\infty} \phi(q) = 0,\ \lim_{q \to \infty} \phi(q) = 1, \text{ and } \phi(q) + \phi(-q) = 1 \text{ for all } (q)$$

[3] McFadden's (1968) contribution is discussed further in section 6.

Marschak *et al.* prove that every Fechner model such that $\phi(q)$ is the distribution function of the difference between two independent and identically-distributed (IID) random variables is binary RUM:

Define the random vector $\mathbf{U} = (U_1,...,U_N)$ by:

$$U_n = v_n + \varepsilon_n \quad \text{for all } \mathbf{x}_n \in T \tag{7}$$

where $\varepsilon_1,...,\varepsilon_N$ are independent and identically-distributed random variables such that $\phi(q)$ is the distribution function of $\varepsilon_n - \varepsilon_m$ for all $\mathbf{x}_m \in T$, $m \neq n$, and $v_1,...,v_N$ are the constants appearing in the Fechner model. Then for any $S = \{\mathbf{x}_n, \mathbf{x}_m\} \subseteq T$:

$$\begin{aligned} P(\mathbf{x}_n | S) &= \Pr\{U_n \geq U_m\} = \Pr\{v_n + \varepsilon_n \geq v_m + \varepsilon_m\} \\ &= \Pr\{v_n - v_m \geq \varepsilon_m - \varepsilon_n\} = \phi(v_n - v_m) \end{aligned} \tag{8}$$

A few comments are appropriate. First, the above analysis relates specifically to binary choice. This reflects the origins of the Fechner model in psychophysical discrimination, e.g. 'is stimulus *a* stronger or weaker than stimulus *b*?'. This is not unduly restrictive, however, since the model can be extended to multinomial choice simply by specifying $\phi(q)$ as a multivariate joint distribution pertaining to each and every pair in the choice set. Second,

$\varepsilon_1,...,\varepsilon_N$ must be defined independently of $v_1,...,v_N$; subsequent presentations of RUM such as Daly and Zachary (1978) are more explicit about this. Third, the IID assumption is not essential to the definition since any distribution function $\phi(q)$ will yield binary RUM; the IID assumption is however relevant to other properties of the model. A fourth, and associated, point is that U is characterised by two redundant degrees of freedom. The remainder of this section expands upon the latter point, whilst the penultimate point is developed further in section 7.1.

As is widely understood, the Fechner implementation of RUM carries two redundant degrees of freedom, as follows. First, a common constant may be added to the utility of each and every alternative without changing probability. This property, which is referred to as 'translational invariance', implies that (8) can be re-written equivalently:

$$P(\mathbf{x}_n|S) = \Pr\{U_n + K \geq U_m + K\} = \Pr\{v_n + \varepsilon_n + K \geq v_m + \varepsilon_m + K\} \quad (9)$$

$$= \Pr\{v_n - v_m \geq \varepsilon_m - \varepsilon_n\}$$

Second, the utility of each and every alternative may similarly be multiplied by a common positive factor without affecting probability; thus (8) can be again expressed equivalently:

$$P(\mathbf{x}_n|S) = \Pr\{\lambda U_n \geq \lambda U_m\} = \Pr\{\lambda v_n + \lambda \varepsilon_n \geq \lambda v_m + \lambda \varepsilon_m\} \qquad (10)$$

$$= \Pr\{v_n - v_m \geq \varepsilon_m - \varepsilon_n\}$$

where $\lambda > 0$

Moreover, the probability statement (8) is robust only to the utility transformations (9) and (10), where the latter constitute the class of increasing linear transformations. Contrast this with the probability statement (6), which is robust to the more general class of strictly monotonic transformations of utility. The important implication follows that *differences* in utility now matter; whilst increasing linear transformations would leave the quantity $(v_n - v_m)$ in (8) unchanged, any other transformation would not. In other words, if the Fechner model is adopted as a working representation of RUM then the above analysis shows that the probability statement now exhibits properties of cardinal utility.

5. ORDINALITY AND CARDINALITY THUS FAR

Let us summarise the story thus far, and consider the implications that follow. The theory of deterministic choice under certainty is explicit about the ordinal basis of utility (section 2). This basis is preserved in the translation from deterministic to probabilistic choice under certainty, and indeed to RUM (section 3). Such a representation of RUM is not particularly amenable to implementation, however, precluding conventional parametric methods of econometric modelling. The Fechner model is, by contrast, readily amenable to implementation, hence Marschak *et al.*'s (1963) interest in establishing analogy between the RUM and Fechner models (section 4). An implication of this analogy is that utility, though fundamentally ordinal, adopts working properties of cardinal utility. The latter does not in itself constitute a problem; provided utility is interpreted only in ordinal terms, it is perfectly reasonable to employ cardinal utility as a working operation of ordinal. Finally, it might be remarked that although a Fechner model with particular distributional assumptions is RUM, RUM is never a Fechner model. That the relation is unidirectional is perfectly intuitive; cardinal utility can always be interpreted in ordinal terms, but ordinal utility can never be interpreted as cardinal.

6. McFADDEN AND THE PRACTICE OF RUM

Let us now digress slightly by reconciling the above discussion with the presentation of RUM more commonly applied in practice. The alternative presentation can be attributed to Daniel McFadden, initially in his 1968 paper (published in 1975), and elaborated further in his many subsequent works. To understand the nub of his approach, however, a brief excerpt from McFadden (1976) suffices:

> *'… the assumption that a single subject will draw independent utility functions in repeated choice settings and then proceed to maximize them is formally equivalent to a model in which the experimenter draws individuals randomly from a population with differing, but fixed, utility functions, and offers each a single choice; the latter model is consistent with the classical postulates of economic rationality'* (McFadden, 1976 p365).

On this basis, (3) is re-written:

$$P(\mathbf{x}_{ni}|T) = \Pr\{U_{ni} \geq U_{mi}\} \qquad \text{for all } \mathbf{x}_{ni} \in T,\ \mathbf{x}_{mi} \in T,\ m \neq n \qquad (11)$$

where an additional subscript is introduced to denote individual i in the population. The probability statement (11) is thus distinct from (3) in that randomness derives from inter- (as opposed to intra-) individual variation in the preference ordering. McFadden's assertion is perhaps made with the pragmatism of the practising economist; given the usual dearth of data on *repeated* choices by an individual - he might well have asked - just how useful is the Marschak (1960) and Block and Marschak (1960) presentation?[4]

In common with Marschak *et al.* (1963), McFadden dissects U into two components, albeit with different interpretation. McFadden (1974) describes v_{ni} as '…non-stochastic and reflects the "representative" tastes of the population…' and ε_{ni} as '…stochastic and reflects the idiosyncracies of this individual in tastes…' (p108). His subsequent works show a gradual evolution in the interpretation of these terms; by the time of his retrospective of RUM in McFadden (2000), v_{ni} is referred to as '…systematic or expected utility', and ε_{ni} simply as 'unobserved factors' (p16). The latter interpretation would seem to imply that, within McFadden's RUM, ε_{ni} could be a catch-all for a number of possible sources of randomness[5]. A further, but related, distinction concerns the perspective of observation, since the latter interpretation is couched more in terms of the viewpoint of the analyst. That is, v_{ni} is deemed to be 'observable' to the analyst, and ε_{ni} 'unobservable' (or at least latent).

[4] Whilst the distinction between inter- and intra- individual variation is a pertinent one, it might be noted that the popular data collection technique of Stated Preference (SP) offers a facility for pursuing both interests. That is to say, SP data is typically collected from a sample of individuals, and each individual typically responds to several repetitions of a given choice task.

[5] Manski (1977) develops this proposition, postulating several specific sources of randomness.

The Fechnerian properties of RUM outlined in section 4 enter McFadden's analysis via the elegant vehicle of his Generalised Extreme Value (GEV) theory (McFadden, 1978). In paying due credit to McFadden, it is important to acknowledge that GEV offers considerable generality over Marschak *et al.* (1963); in particular, it considers choice sets larger than binary, and accommodates covariance between the random variables of different partitions of the choice set. Those observations notwithstanding, McFadden's RUM carries the same implication as Marschak *et al.*'s for the cardinality of utility.

7. THE PRACTICE OF RUM, AND ITS FAITHFULNESS TO ORDINAL UTILITY

The practice of RUM has evolved through the adoption of the Fechner model as a means of establishing a parametric relationship between the observed (ordinal) choices of the individual and 'observable' utility differences $(v_n - v_m)$. It remains to consider whether practice adheres to the requirement that utility be interpreted only in ordinal terms. We can develop this interest by applying order-preserving transformation to utility, and considering whether this results in any substantive change to the inferences deriving from the Fechner model. In particular, let us consider three measurements that are pertinent to practice: the marginal utility of an attribute, the marginal valuation of an attribute, and the change in consumer surplus arising from a change in the price of an alternative. For present purposes, it suffices to consider the choices of a single individual, though it might be noted that the reported implications readily transfer to McFadden's RUM.

7.1 Marginal utility of an attribute

In developing the first two measurements, those of marginal utility and marginal valuation, let us represent utility (7) as indirect, functional on the attributes of the alternatives and money budget, thus:

$$U_n = v_n(\mathbf{x}_n; y - p_n) + \varepsilon_n \qquad \text{for all } \mathbf{x}_n \in T \tag{12}$$

where y is money budget, and p_n is the price of alternative $\mathbf{x}_n$

The derivation of marginal utility depends of course on (12) being differentiable. This is not guaranteed by the axioms of completeness and transitivity alone, hence the need for an additional axiom, that of continuity (e.g. Debreu, 1954):

Axiom of Continuity

For any pair of alternatives $\mathbf{x}_n, \mathbf{x}_m \in T$ where $\mathbf{x}_n \succeq \mathbf{x}_m$; a neighbourhood $z(\mathbf{x}_n)$ exists such that for any alternative $\mathbf{x}_l \in z(\mathbf{x}_n)$, $\mathbf{x}_l \succeq \mathbf{x}_m$; and a neighbourhood $z(\mathbf{x}_m)$ exists such that for any bundle $\mathbf{x}_l{}' \in z(\mathbf{x}_m)$, $\mathbf{x}_n \succeq \mathbf{x}_l{}'$

Moreover, conventional practice is to specify the constant v_n of the indirect utility function (12) as a continuous and differentiable function g of the attributes $\mathbf{x}_n$ that embody the alternative, with the money budget entering additively[6]. This yields the so-called Additive Income RUM (or 'AIRUM') popularised by McFadden (1981), thus:

$$U_n = \alpha(y - p_n) + g(\mathbf{x}_n) + \varepsilon_n$$

Then explicating the redundant degrees of freedom (9) and (10), we arrive at the following:

$$U_n = \lambda[\alpha(y - p_n) + g(\mathbf{x}_n) + \varepsilon_n + K] \qquad (13)$$

[6] It might be noted that recent contributors have considered the possibility of non-linearity, for example McFadden (1999), Karlström (1998, 2001), Dagsvik (2001), Dagsvik and Karlström (2005).

Equation (13) reveals K to be confounded with y, such that choice probability is - in AIRUM at least - invariant to money budget. Now pursuing our interest in the marginal utility of an attribute, let us differentiate (13) with respect to an attribute $x_k \in \mathbf{x}$, thus:

$$\frac{\partial U_n}{\partial x_{kn}} = \frac{\partial U_n}{\partial g}\frac{\partial g}{\partial x_{kn}} = \lambda \frac{\partial g}{\partial x_{kn}} \quad \text{for all } \mathbf{x}_n \in T, x_{kn} \in \mathbf{x}_n \tag{14}$$

Equation (14) demonstrates that marginal utility derived from the Fechner representation of RUM is scaled in terms of the λ transformation of utility. Developing ideas further, let us apply an order-preserving and differentiable transformation f to the utility function (13) - remembering from (6) that probability in the ordinal representation of RUM would be robust to such transformation - and repeat the derivation of marginal utility:

$$\hat{U}_n = f\{\lambda[\alpha(y - p_n) + g(\mathbf{x}_n) + \varepsilon_n + K]\}$$

$$\frac{\partial \hat{U}_n}{\partial x_{kn}} = \frac{\partial \hat{U}_n}{\partial f}\frac{\partial f}{\partial g}\frac{\partial g}{\partial x_{kn}} = \frac{\partial \hat{U}_n}{\partial f}\lambda \frac{\partial g}{\partial x_{kn}} \tag{15}$$

Comparison of (14) and (15) reveals that the former is subject to the cardinal scale λ, whereas the latter is subject to both the same cardinal scale and the ordinal scale f.

As section 3 has demonstrated, the derivation of RUM relies entirely on ordinal utility, and thus establishes the basis for admitting order-preserving transformations of utility. RUM practice has tended to overlook the latter requirement, adopting for working purposes the marginal utility (14), which may be seen as a restricted case of (15). Indeed somewhat ironically, RUM practice has, whilst overlooking the f scale, invested considerable effort in managing the λ scale, usually under the banner of the 'scale factor problem' (e.g. Swait and Louviere, 1993). With reference to (10), the scale factor problem acknowledges the relation between the λ scale and the variance of the random variables that is engendered through the identity:

$$\text{var}[\lambda(\varepsilon_m - \varepsilon_n)] = \lambda^2 \text{ var}(\varepsilon_m - \varepsilon_n)$$

Rearranging:

$$\text{s.d.}[\lambda(\varepsilon_m - \varepsilon_n)] = \lambda \,\text{s.d.}(\varepsilon_m - \varepsilon_n),$$

$$\lambda = \frac{\text{s.d.}[\;]}{\text{s.d.}(\;)} \tag{16}$$

Whilst IID within any dataset, the random variables may not be identically distributed across datasets, and (16) therefore provokes the 'problem' that the utilities (and marginal utilities) deriving from those datasets may be of different λ scale. RUM practice has shown considerable support for this proposition, particularly when merging Revealed Preference (RP) data with Stated Preference (SP) data, where the latter invariably carry the lower variance. In seeking to resolve the scale factor problem, several contributors (e.g. Morikawa, 1989, 1994; Bradley and Daly, 1997) have proposed methods that combine one or more datasets within the same RUM, explicitly accommodating the different scale factor of each. These methods essentially amount to setting the cardinal scale of one dataset as the base, and the scales of other datasets as relative to the base (e.g. $\lambda^{SP} = \theta\lambda^{RP}$, where θ is the difference in scale between RP and SP, and one would expect *a priori* that $\theta > 1$). Whilst this serves to establish a unique λ scale across the various datasets, it should be noted importantly that λ does not however disappear. Thus utility continues to embody a cardinal scale pertaining to the λ transformation.

7.2. Marginal valuation of an attribute

It is well established in microeconomic theory that the ratio between the marginal utilities of two attributes represents the marginal rate of substitution between those attributes. If in particular the denominator of the ratio is the marginal utility of income net of price then the marginal rate of substitution can be further interpreted as the marginal valuation of the attribute represented in the numerator. To illustrate, and irrespective of whether the marginal utilities are of the form (14) or (15), the marginal valuation of an attribute $x_{kn} \in \mathbf{x}_n$ is given by:

$$Vo(x_{kn}) = \frac{\partial U_n}{\partial x_{kn}} \bigg/ \frac{\partial U_n}{\partial (y - p_n)} = \frac{\partial g}{\partial x_{kn}} \bigg/ \alpha \qquad \text{for all } \mathbf{x}_n \in T,\, x_{kn} \in \mathbf{x}_n$$

Moreover, marginal valuations are free of any scale, whether ordinal or cardinal, such that the distinction between ordinal utility and cardinal utility is in this context rendered immaterial.

7.3 Consumer surplus change

Having considered the marginal utility and marginal valuation of an attribute, a further practical interest is in forecasting the change in consumer surplus arising from a change in price. Note that such a relation exists only if the change in price induces a change in the probability of choice. Reconciling this with the ordinal basis of RUM, a change in attributes may impact upon an individual's preference orderings; if the preference orderings do indeed change, then a new set of choice probabilities could feasibly result. Two methods of measuring the change in consumer surplus are prevalent in RUM practice, referred to as the 'rule-of-half' and 'log sum' methods.

The rule-of-a-half method (Lane, Powell and Prestwood Smith, 1971; Neuburger, 1971) offers a first-order approximation to the change in consumer surplus deriving from a Marshallian demand function. Let us apply this to RUM, considering in particular an increase in the price of a specific alternative $\mathbf{x}_m \in T$, such that $p_m \rightarrow p_m + \Delta p_m$ where $\Delta p_m > 0$, holding all else constant. Now introducing the subscripts 0 and 1 to represent the states before and after the price increase, define Q_0 and Q_1 to be the number of repetitions of the choice task faced by the individual in the respective states[7], and $P_0(\mathbf{x}_m|T)$ and $P_1(\mathbf{x}_m|T)$ to be the associated choice probabilities[8]. Having equipped ourselves with the necessary notation, we can now write a statement - in accordance with the rule-of-a-half method - of the change in

[7] Or in other words, total units consumed in each state; i.e. in the example of Table 1, $Q = 5$.

[8] Note in passing that if $P_1(\mathbf{x}_m|T) \neq P_0(\mathbf{x}_m|T)$, such that the price increase Δp_m results in a change in the probability of choosing $\mathbf{x}_m \in T$, then the requirement from (3) that all probabilities sum to one implies the following: $\sum_{l \in T} P_1(\mathbf{x}_l|T) \neq \sum_{l \in T} P_0(\mathbf{x}_l|T)$, $l \neq m$. That is to say, any change in $P(\mathbf{x}_m|T)$ must be compensated by changes in one or more of the other probabilities.

consumer surplus ΔCS_m (in this instance a loss) arising from the price increase Δp_m, thus:

$$\Delta CS_m = \tfrac{1}{2}\left[Q_1 P_1\left(\mathbf{x}_m \middle| T\right) + Q_0 P_0\left(\mathbf{x}_m \middle| T\right)\right]\Delta p_m \quad \text{for all } \mathbf{x}_m \in T \qquad (17)$$

If the probabilities $P_0\left(\mathbf{x}_m \middle| T\right)$ and $P_1\left(\mathbf{x}_m \middle| T\right)$ are directly observable then, with reference to (6), they can be interpreted in ordinal terms. Accepting that, and noting that the quantities Q_0 and Q_1 simply represent the number of orderings that give rise to these probabilities, it becomes clear that (17) can be supported entirely by ordinal utility.

In many practical situations, however, the after state exists only hypothetically, and there may instead be interest in forecasting the effect of the price increase. The procedure of forecasting exploits, somewhat inevitably, the cardinal properties of the Fechner model. That is to say, the redundant degrees of freedom (9) and (10) become relevant, and the forecast of $P_1\left(\mathbf{x}_m \middle| T\right)$ will, with reference to (8), arise specifically from extrapolation of the quantity $\left(v_n - v_m\right)$ estimated on data for $P_0\left(\mathbf{x}_m \middle| T\right)$ in the before case. Moreover, the rule-of-a-half remains faithful to ordinal utility only if the probabilities before and after the price change are observed, and therefore reconcilable with (6), rather than forecasted by means of (8).

Now consider the second method of measuring the change in consumer surplus - the 'log sum' method. Although the origins of this method are evident in Williams (1977), Small and Rosen (1981) were first to offer a full and definitive treatment, with McFadden (1981) applying this treatment specifically to AIRUM. The log sum method captures, in the context of RUM, the Hicksian compensating variation of a price change, i.e. the minimum/maximum quantity of money that must be given to/taken from an individual in order leave him or her on the same indifference curve as before the price increase/reduction. If RUM carries no income effect, as would apply to AIRUM given the translational invariance of income, then the Hicksian compensating variation may be interpreted - entirely equivalently - as the change in consumer surplus pertaining to a Marshallian demand (i.e. analogous to the rule-of-a-half).

More formally, and with particular reference to the presentation of Dagsvik and Karlström (2005), the log sum method derives from an equality established between the maximal utilities that arise before and after a price change, wherein the utilities are of the form (12). To illustrate, let us consider the same price increase as before, that is an increase in the price of alternative $\mathbf{x}_m \in T$, such that $p_m \rightarrow p_m + \Delta p_m$ where $\Delta p_m > 0$, holding all else constant. In this case, the aforementioned equality may be written:

$$\max_{n=1,\ldots,N}\left[v_n\left(\mathbf{x}_n; y + c - \left(p_n + \delta_{mn}\Delta p_m\right)\right) + \varepsilon_n\right] = \max_{n=1,\ldots,N}\left[v_n\left(\mathbf{x}_n; y - p_n\right) + \varepsilon_n\right]$$

where c represents the compensating variation, p_n is price in the before state, and $\delta_{mn} = \begin{cases} 0 & m \neq n \\ 1 & m = n \end{cases}$

(18)

For the price increase Δp_m, any compensating variation must be given to the individual, such that $c \geq 0$. It is of course conceivable that the same alternative may not be utility-maximising in both states. Indeed, let us denote the utility-maximising alternatives in states 0 and 1 to be $\mathbf{x}_{n^0} \in T$ and $\mathbf{x}_{n^1} \in T$, where these alternatives may or may not be one and the same. Furthermore let the prices of these alternatives in the before state be denoted p_{n^0} and p_{n^1} respectively. Then re-stating the equality (18) specifically in terms of the Fechner model, once again explicating the redundant degrees of freedom:

$$\lambda\left\lfloor \alpha\left(y + c - \left(p_{n^1} + \delta_{mn^1}\Delta p_m\right)\right) + g\left(\mathbf{x}_{n^1}\right) + \varepsilon_{n^1} + K\right\rfloor$$
$$= \lambda\left\lfloor \alpha\left(y - p_{n^o}\right) + g\left(\mathbf{x}_{n^0}\right) + \varepsilon_{n^0} + K\right\rfloor$$

where $\delta_{mn^1} = \begin{cases} 0 & m \neq n^1 \\ 1 & m = n^1 \end{cases}$

(19)

It is conventionally assumed that the random variables ε_{n^0} and ε_{n^1} of the maximal utilities are distributed identically; with reference to the earlier discussion of section 7.1, this has the effect of imposing a common cardinal scale across the two states. This assumption of identical distributions is rather strong, perhaps, but let us accept it nonetheless, particularly as it facilitates considerable simplification of (19). Indeed, accepting this assumption we arrive at a reasonably succinct statement - according to the log sum method - of the consumer surplus loss ΔCS_m associated with the price increase Δp_m, thus:

$$\Delta CS_m = c = \frac{\lfloor g(\mathbf{x}_{n^0}) - \alpha p_{n^0} \rfloor - \lfloor g(\mathbf{x}_{n^1}) - \alpha(p_{n^1} + \delta_{mn^1} \Delta p_m) \rfloor}{\alpha} \quad \text{for all } \mathbf{x}_m \in T \tag{20}$$

It is important to acknowledge that whilst the subtraction - acknowledging that this is in itself a cardinal operation - between the maximal utilities in states 0 and 1 serves to remove K (together with money budget y) and λ, this does not mean that cardinality is absent from (20). Rather the cardinal scale has been standardised across the two states. We can demonstrate this property by applying an order-preserving but non-linear transformation h to the maximal utilities within (20), noting that h is a member of the class of order-preserving transformations f introduced earlier, but disjoint from the class of increasing linear transformations embodied by K and λ. Clearly, identification of the utility-maximising alternatives $\mathbf{x}_{n^0} \in T$ and $\mathbf{x}_{n^1} \in T$ is unaffected by the h transformation, since the transformation is order-preserving. By contrast, the difference between the maximal utilities in those states is not preserved under the same transformation. Hence with reference to (20), the following inequality applies:

$$\Delta CS_m = c \neq \frac{h\lfloor g(\mathbf{x}_{n^0}) - \alpha p_{n^0} \rfloor - h\lfloor g(\mathbf{x}_{n^1}) - \alpha(p_{n^1} + \delta_{mn^1}\Delta p_m)\rfloor}{\alpha}$$

Moreover, the compensating variation derived by means of the log sum method is robust only to cardinal transformations of utility, and not to other order-preserving transformations. Indeed, in order to interpret (20) even in ordinal terms (i.e. to say whether maximal utility in the after state is greater than or less than the maximal utility in the before state), it has been necessary to impose a common cardinal scale across the two states.

Finally, note that where the random variables $\varepsilon_1,...,\varepsilon_N$ of (8) are specifically IID Gumbel, the Fechner model adopts the logit form. Then adhering to the analysis outlined above, but taking expectations of the maximal utilities over repetitions and/or individuals[9], (20) may be re-written as follows, the form of which provokes the terminology 'log sum':

[9] It might in passing be remarked that the act of taking expectations would in itself seem to carry a strong inference of cardinality.

$$\Delta E(CS_m) = E(c) = \frac{\ln\left\{\sum_{n=1}^{N}\left[g(\mathbf{x}_n) - \alpha p_n\right]\right\} - \ln\left\{\sum_{n=1}^{N}\left[g(\mathbf{x}_n) - \alpha\left(p_n + \delta_{mn}\Delta p_m\right)\right]\right\}}{\alpha}$$

where p_n is price in the before state, and $\delta_{mn} = \begin{cases} 0 & m \neq n \\ 1 & m = n \end{cases}$

7.4 Prescriptive advice

Let us complete this discussion of RUM practice by synthesising its findings, and offering prescription to practitioners. Whilst marginal utility derived from the Fechner model (as with any marginal utility, it might be added) exhibits cardinal properties, it is usually applied only for purposes of inferring marginal valuation, and the latter is free of either cardinal or ordinal scale. The cardinality of utility within the Fechner model does however manifest when deriving measurement of consumer surplus change, whether implemented through the log sum method, or through the application of probability forecasts to the rule-of-a-half method. Hence, to the extent that practice exploits such methods, which would indeed seem considerable, it may stand accused of transgressing the theoretical tenets on which RUM is founded. There is however a convenient resolution: this is to apply the rule-of-a-half using the ordinal interpretation of RUM originally proposed by Marschak (1960) and Block and Marschak (1960).

8. SUMMARY AND CONCLUSION

The paper began by revisiting the theoretical origins of RUM, distinguishing the original presentation by Marschak (1960) and Block and Marschak (1960) from the popular McFadden (e.g. 1968, 1975) presentation. Consistent with Neo-Classical theory, the original presentation is couched at the individual level, and relies on an ordinal notion of utility. Randomness then derives from the repetition, in particular the facility for variability on these repetitions, of an individual's preference ordering. The paper proceeded to consider the implementation of RUM, which is achieved through a relation with Fechner's (1859) model of psychophysical discrimination. This served to illuminate the fundamental point of the paper; that the implementation of RUM yields a model that carries the properties of cardinal utility. Though this is not in itself a problem - cardinal utility can be used quite defensibly as a representation of ordinal utility - it is essential that the manner in which the representation is interpreted and applied does not depart from ordinality. The paper considered whether RUM practice is adherent to the latter requirement, particularly in relation to three measurements routinely derived from the Fechner model, namely the marginal utility of an attribute, the marginal valuation of an attribute, and the change in consumer surplus arising from a change in the price of an alternative.

As would seem inherent in the very notion of marginal utility, marginal utility derived from the Fechner model carries a cardinal scale. RUM practitioners have invested considerable effort in seeking to resolve the so-called 'scale factor problem', and this may be rationalised as an attempt to establish a unique cardinal scale across data pooled from different sources. If however the ratio of two marginal utilities is taken, thereby yielding the marginal rate of substitution (or in the particular case of the price attribute, marginal valuation), then the cardinal scale is removed and the resultant metric is defensible in ordinal terms. The cardinal properties of the Fechner model become rather more pertinent when measuring consumer surplus change, as follows. If the Fechner model is applied to the rule-of-a-half method for measuring consumer surplus change, then this would seem to be associated with the forecasting of choice probability by extrapolation, where the latter itself exploits cardinality. The log sum method for measuring consumer surplus change is defined entirely in terms of the Fechner model, and unambiguously departs from ordinal interpretation. RUM practice routinely exploits these two methods of measuring consumer surplus change, and may therefore stand accused of operating outside of the bounds for which theory offers legitimacy. There is however a convenient means of realigning theory and practice; this is to apply the rule-of-a-half method using the original presentation of RUM (Marschak, 1960; Block and Marschak, 1960). The original presentation is defined entirely in terms of ordinal utility, and therefore faithful to theory.

ACKNOWLEDGEMENTS

The author was supported by a Departmental Research Fellowship at the Institute for Transport Studies (ITS), University of Leeds, and by the Platform Grant GR/S46291/01 'Towards a unified theoretical framework for transport network and choice modelling', which is funded by the UK Engineering and Physical Sciences Research Council and also held at ITS.

The author acknowledges grateful thanks to several colleagues - Andrew Daly, Tony Fowkes, Nicolás Ibáñez, Jeremy Toner, Chris Nash, David Watling and Richard Connors - for comments on earlier versions of the manuscript. Any errors and misunderstandings that remain are of course the sole responsibility of the author.

REFERENCES

Block, H.D. and Marschak, J. (1960) Random orderings and stochastic theories of responses. In Marschak, J. (1974) *Economic Information, Decision and Prediction: Selected Essays* (Volume 1). D. Reidel, Dordrecht.

Bradley, M.A. and Daly, A.J. (1997) Estimation of logit choice models using mixed stated preference and revealed preference information. In Stopher, P.R. and Lee-Gosselin, M. (eds) *Understanding Travel Behaviour In An Era Of Change*. Pergamon, Oxford.

Dagsvik, J.K. (2001) Compensated Variation in Random Utility Models. Discussion Paper No. 299, Statistics Norway.

Dagsvik, J.K. and Karlström, A. (2005) Compensating Variation and Hicksian choice probabilities in Random Utility Models that are nonlinear in income. *The Review of Economic Studies*, 72 (250), pp57-76.

Daly, A.J. (1978) Issues in the estimation of journey attribute values. In Hensher, D.A. and Dalvi, Q. (eds) *Determinants of Travel Choice*. Saxon House, Farnborough.

Daly, A.J. and Zachary, S. (1978) Improved multiple choice models. In Hensher, D.A. and Dalvi, Q. (eds) *Determinants of Travel Choice*. Saxon House, Farnborough.

Debreu, G. (1954) Representation of a preference ordering by a numerical function. In Thrall, R.M., Coombs, C.H. and Davis, R.L. (eds) *Decision Processes*. Wiley, New York.

Fechner, G.T. (1859) *Elemente der Psychophysik*. Breitkopf and Härtel, Leipzig.

Goodwin, P.B. and Hensher, D.A. (1978) The transport determinants of travel choices: an overview. In Hensher, D.A. and Dalvi, Q. (eds) *Determinants of Travel Choice*. Saxon House, Farnborough.

Karlström, A. (1998) Hicksian welfare measures in a nonlinear random utility framework. Working paper, Royal Institute of Technology, Department of Infrastructure and Planning, Stockholm, presented at the 1st KFB Transport Economics Conference, Borlange, Sweden, June 1998.

Karlström, A.K. (2001) Welfare evaluations in a nonlinear Random Utility Models with income effects. In Hensher, D. (ed.) *Travel Behaviour Research: The Leading Edge*. Pergamon Press, Sydney.

Lancaster, K.J. (1966) A new approach to consumer theory. *Journal of Political Economy*, 74 (2), pp132-157.

Lane, R., Powell, T.J. and Prestwood Smith, P. (1971) *Analytical Transport Planning*. Duckworth, London.

Luce, R.D. (1959) *Individual Choice Behaviour: A Theoretical Analysis*. John Wiley, New York.

Luce, R.D. and Suppes, P. (1965) Preference, utility and subjective probability. In Luce, R.D., Bush, R.R. and Galanter, E. (eds) *Handbook of Mathematical Psychology* (Volume III). John Wiley and Sons, New York.

McFadden, D. (1968) The revealed preferences of a government bureaucracy. Department of Economics, University of California, Berkeley, California, unpublished.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behaviour. In Zarembka, P. (ed) *Frontiers in Econometrics*. Academic Press, New York.

McFadden, D. (1975) The revealed preferences of a government bureaucracy: theory. *The Bell Journal of Economics and Management Science*, 6 (2), pp401-416.

McFadden, D. (1976) Quantal choice analysis: a survey. *Annals of Economic and Social Measurement*, 5 (4), pp363-390.

McFadden, D. (1978) Modelling the choice of residential location. In Karlqvist, A., Lundqvist, L., Snickars, F. and Weibull, J. (eds) *Spatial Interaction Theory and Residential Location*. North-Holland, Amsterdam.

McFadden, D. (1981) Econometric models of probabilistic choice. In Manski, C. and McFadden, D. (eds) *Structural Analysis of Discrete Data: With Econometric Applications*. The MIT Press, Cambridge, Massachusetts.

McFadden, D. (1999) Computing willingness-to-pay in Random Utility Models. In Moore, J., Riezman, R. and Melvin, J. (eds.) *Trade, Theory and Econometrics: Essays in Honour of John S. Chipman*. Routledge, London.

McFadden, D. (2000) Disaggregate Behavioral Travel Demand's RUM side: A 30 year perspective. In Hensher, D. (ed) *Travel Behaviour Research: The Leading Edge*. Pergamon, Oxford.

Manski, C.F. (1977) The structure of random utility models. *Theory and Decision*, 8, pp229-254.

Marschak, J. (1960) Binary choice constraints and random utility indicators. In Marschak, J. (1974) *Economic Information, Decision and Prediction: Selected Essays* (Volume 1). D. Reidel, Dordrecht.

Marschak, J., Becker, G.M. and DeGroot, M.H. (1963) Stochastic models of choice behavior. In Marschak, J. (1974) *Economic Information, Decision and Prediction: Selected Essays* (Volume 1). D. Reidel, Dordrecht.

Morikawa, T. (1989) Incorportaing stated preference data in travel demand analysis. Dissertation for PhD, Department of Civil Engineering, Massachusetts Institute of Technology.

Morikawa, T. (1994) Correcting state dependence and serial correlation in the RP/SP combined estimation method. *Transportation*, 21, pp153-165.

Neuburger, H. (1971) User benefit in the evaluation of transport and land use plans. *Journal of Transport Economics and Policy*, 5 (1), pp52-75.

Small, K.A. and Rosen, H.S. (1981) Applied welfare economics with discrete choice models. *Econometrica*, 49 (1), pp105-130.

Swait, J. and Louviere, J. (1993) The role of the scale parameter in the estimation and use of multinomial logit models. *Journal of Marketing Research*, 30, pp305-314.

Thurstone, L. (1927) A law of comparative judgement. *Psychological Review*, 34, pp273-286.

Williams, H.C.W.L. (1977) On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A*, 9 (3), pp285-344.