SOME PROBABILISTIC MODELS
OF SIMPLE CHOICE AND RANKING

I. THE REVERSIBLE RANKING MODEL

II. CONJOINT UTILITY MODELS

by

Anthony Alfred John Marley

A DISSERTATION

IN

PSYCHOLOGY

Presented to the Faculty of the Graduate School of Arts and Sciences of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy.

1965

R. Duncan Luce

Supervisor of Dissertation

Henry Gleitman

Graduate Group Chairman

Acknowledgements

Preparation of this dissertation was supported in part by National Science Foundation Grant NSF GB-1462 to the University of Pennsylvania. Also, I acknowledge the support of University of Pennsylvania Fellowships during 1964 and 1965. I wish to thank Mrs. Sally Kraska for typing the manuscript.

I am particularly indebted to Dr. R. Duncan Luce for his generous help and guidance with the various drafts of this thesis.

M. A. A. J.

1966 g. University of Pennsylvania

Index

Table of Contents

List of Figures

Bibliography

Aczel, J. A remark on functional dependence. J. math. Psychol., 1965, 2, 125-127.

Block, H.D. and Marschak, J. Random orderings and stochastic theories of responses. In Olkin et al. (Eds.) Contributions to probability and statistics. Stanford: Stanford Univer. Press, 1960. Pp. 97-132.

Davidson, D., Suppes, P. and Siegel, S. Decision-making: an experimental approach. Stanford: Stanford Univer. Press, 1957.

Debreu, G. Stochastic choice and cardinal utility. Econometrica, 1958, 26, 440-444.

Luce, R.D. On the possible psychophysical laws. Psychol. Rev., 1959, 66, 81-95 (a).

Luce, R.D. Individual choice behavior. New York: Wiley, 1959 (b).

Luce, R.D. Response latencies and probabilities. In K.J. Arrow, S. Karlin and P. Suppes (Eds.) Mathematical methods in the social sciences. Stanford: Stanford Univer. Press, 1960.

Luce, R.D. A generalization of a theorem of dimensional analysis. J. math. Psychol., 1964, 1, 278-284.

Luce, R.D. and Tukey, J.W. Simultaneous conjoint measurement: a new type of fundamental measurement. J. math. Psychol., 1964, 1, 1-27.

Luce, R.D. and Suppes, P. Preference, utility and subjective probability. In R.D. Luce, R.R. Bush and E. Galanter (Eds.) Handbook of mathematical psychology, III. New York: Wiley, 1965.

Marley, A.A.J. The relation between the discard and regularity conditions for choice probabilities. J. math. Psychol. 1965, 2. In press.

McLaughlin, D.H. and Luce, R.D. Stochastic transitivity and cancellation of preferences between bitter-sweet solutions.

Psychon. Sci., 1965, 2, 89-90.

Pendergrass, R.N. Ranking in triple comparisons (Ph.D. dissertation). Library, Virginia Polytechnic Institute, Blacksburg, Va., 1958.

Pendergrass, R.N. and Bradley, R.A. Ranking in triple comparisons. In Olkin et al. (Eds.). Contributions to probability and statistics. Stanford: Stanford Univer. Press, 1960, Pp. 331-351.

Sedov, L.I. Similarity and dimensional methods in mechanics. New York: Academic Press, 1959.

1. Introduction

"Of the major areas into which experimental psychology has been traditionally partitioned, motivation is the least well understood and systematized. This is true whether we consider theory, experimental paradigms or experimental results.... Moreover, of the various notions usually considered to be primarily motivational, preference is the only one that mathematical psychologists have attempted to analyze with any care: there are almost no satisfactory formal theories concerning, for example, drive and incentive, and those that exist are best discussed as aspects of learning. So this chapter on mathematical theories of motivation is limited to a study of preference and the closely related constructs of utility and subjective probability." (Luce and Suppes, 1965, p. 252)

In this thesis we develop certain aspects of the theories of preference discussed in the above mentioned chapter.

Luce and Suppes (p. 256-257) classify asymptotic theories of preference according to three binary distinctions, two of which are relevant to the present study. If we suppose that the responses that a subject makes to stimulus presentations are governed by probability mechanisms then a theory is <u>algebraic</u> if it requires that these probabilities be either 0, 1/2, or 1 and <u>probabilistic</u> otherwise. We confine our attention to probabilistic theories. A <u>simple choice experiment</u> is one in which a subject is asked to select among several outcomes and a <u>ranking experiment</u> is one in which he is asked to rank order them. We consider theories for both classes of experiments.

When decisions are governed by a probabilistic process which generates probabilities other than 0, 1/2 and 1 the connections between simple choice and ranking are not apparent. Nonetheless, we expect regular relations to exist between behavior in these two kinds of experiments and Luce and Suppes (Pp. 351-358) expound certain ideas on this topic. In the first half of the thesis we show that their assumptions have many interesting implications beyond those that they discuss.

In the second half of the thesis we suppose that a subject is making binary comparisons between elements of a set $A = X \times X^*$ where each element of A is called a <u>two component object</u>. For $a,b \in A$, $p(a,b)$ denotes the probability that a is chosen over b. If the ordering $\geq$ on A defined by

$$a \geq b \text{ if and only if } p(a,b) \geq 1/2$$

is such that $(A, \geq)$ satisfies the axioms of Luce and Tukey (1964) then the equivalent axioms for the binary probabilities yield a representation theorem of the form: there exist real valued functions f, g with domains X, X^*, respectively, such that for $x,y \in X$, $x^*,y^* \in X^*$,

$$p[(x,x^*),(y,y^*)] \geq 1/2 \text{ if and only if}$$

$$f(x) + g(x^*) \geq f(y) + g(y^*) ,$$

where f,g are interval scales with a common unit.

This example is given by Luce and Suppes (p. 333) and is the full extent of their discussion of probabilistic theories for binary choices between two component objects. The second half of the thesis contains a deeper, though still incomplete, discussion of such theories.

2. Probabilistic Preference Theories

From Luce and Suppes' (1965, Sec. 5, 6, Pp. 331-358) summary and extension of results about models of preference, it is apparent that most theorists have been concerned with the particular paradigm in which a subject has to choose the most preferred element from the available set. A much smaller body of work concerns experiments in which the subject has to rank order the set according to some criterion or has to choose the least preferred element from the available set. In particular, the strict utility model [Luce and Suppes (1965, p. 336)] has been applied almost exclusively to choice of the most preferred element, the major exceptions being Luce (1959, Pp. 56-58), Block and Marschak (1960, p. 111) and Luce and Suppes (1965, Pp. 356-358), in each of which it is also applied to choice of the least preferred element. Block and Marschak (1960, p. 111) show, however, that such constraints on the choice behavior are incompatible with certain reasonable conditions on ranking behavior unless all of the choice probabilities are equal. There are at least two ways of circumventing this undesirable theorem. The first is to assume that the choice behavior is in agreement with the strict utility model but that the ranking hypotheses are incorrect; and the second is to assume that the ranking hypotheses are correct, but that the choice behavior is not in agreement with the strict utility model. In this thesis we study the second alternative.

2.1 Notation.

A = a set of elements.

For $x \in X \subseteq A$,

$P_X(x)$ = the probability of choosing x as the most preferred element in X ,

$P_X = \{P_X(x): x \in X\}$

$P^*_X(x)$ = the probability of choosing x as the least preferred element in X

$P^*_X = \{P^*_X(x): x \in X\}$

$R(X)$ = the set of rank orders of X

ρ = an arbitrary element of $R(X)$.

Provided X has at least two elements, we may denote $\rho = \rho_1\rho_2\cdots\rho_{n-1}\rho_n \in R(X)$ by $\rho = \rho_1\sigma$ where $\sigma = \rho_2\rho_3\cdots\rho_{n-1}\rho_n \in R(X-\{\rho_1\})$ and provided X has at least three elements we may write $\rho = \rho_1\mu\rho_2$ where $\mu = \rho_2\rho_3\cdots\rho_{n-2}\rho_{n-1} \in R(X-\{\rho_1\rho_n\})$.

$R(x;X) = \{\rho|\rho \in R(X)$ and $\rho_1 = x\}$, i.e. $R(x;X)$ is the set of rank orders of X in which x has rank 1.

ρ^* = the inverse ranking of ρ , i.e. $\rho^*_1 = \rho_n$, $\rho^*_2 = \rho_{n-1}$, $\ldots\rho^*_i = \rho_{n-i+1},\ldots\rho^*_n = \rho_1$.

$p(\rho)$ = the probability of obtaining $\rho \in R(X)$ when the subject is ranking from the most preferred element to the least preferred element, in which case ρ_1 is the most preferred element.

$p^*(\rho)$ = the probability of obtaining $\rho \in R(X)$ when the subject is ranking from the least preferred element to the most preferred element, in which case ρ_1 is the least preferred element.

For the notation to be complete, $p(\rho)$ and $p^*(\rho)$ should include reference to the set X of which ρ is a rank order; however this omission will not lead to confusion and so we use the simpler notation. We use the convention that $P_{\{x\}}(x) = 1 = P^*_{\{x\}}(x)$ for each $x \in A$ and write $p(x,y)$ for $P_{\{x,y\}}(x)$, $p^*(x,y)$ for $P^*_{\{x,y\}}(x)$.

2.2 Reversible ranking models.

We are interested in how the <u>ranking probabilities</u> $p(\rho)$, $p^*(\rho)$ relate to the <u>preference probabilities</u> P_X and to the <u>aversion probabilities</u> P^*_X. Luce (1959, Pp. 69-70), Block and Marschak (1960, p. 109) and Luce and Suppes (1965, Sec. 6, Pp. 351-358) have considered the following possibility. A person might rank order a set of alternatives by first selecting the most preferred outcome and giving this rank one; he then selects the best outcome from the remaining set and ranks it second and so on until the set is exhausted. On the other hand it is equally possible that he may select the least satisfactory outcome and rank it last, then the least satisfactory outcome from the remaining set and rank it next to last and so on. A possible formulation of the first suggestion is that for any $\rho \in R(X)$,

$$p(\rho) = P_X(\rho_1)P_{X-\{\rho_1\}}(\rho_2)\cdots p(\rho_{n-1},\rho_n) \tag{1}$$

and of the second suggestion that

$$p^*(\rho) = P^*_X(\rho_1)P^*_{X-\{\rho_1\}}(\rho_2)\cdots p^*(\rho_{n-1},\rho_n) \ . \tag{2}$$

It is possible that the probability of obtaining a particular ranking ρ when ranking from most preferred to least preferred equals the probability of obtaining this ranking when ranking from least preferred to most preferred, i.e. for each $\rho \in R(X)$,

$$p(\rho) = p^*(\rho^*) \ . \tag{3}$$

These conditions are formalized as:

<u>Definition 1</u>. <u>A</u> reversible ranking model <u>is a set of preference, aversion and non-zero ranking probabilities for the subsets of a set</u> A <u>which satisfy Eqs</u>. 1-3 <u>for all</u> $X \subseteq A$.

A special case of this model was studied by Pendergrass (1958) and by Pendergrass and Bradley (1960). These authors restricted their attention to three element sets and also assumed that the binary choice probabilities satisfy the strict binary utility model.

We can now state precisely the result of Block and Marschak (1960, p. 111) mentioned earlier. They proved that the set of preference and the set of aversion probabilities of a reversible ranking model both satisfy the strict utility model only if the preference and aversion probabilities associated with any particular presentation set are all equal. However, Pendergrass (1958) presents an interesting reversible ranking model in which all choices need not be equally likely and thus it is the joint assumption of a reversible ranking model and the strict utility model which is entirely too strong.

2.3 Concordant choice models.

We are also interested in how the preference probabilities P_X are related to the aversion probabilities P^*_X. Suppose that a subject has to select both the most preferred and the least preferred element in a set X and he does this sequentially. Then it is possible that the probability of choosing x as best and $y \neq x$ as worst is independent of whether the subject first chooses $x \in X$ as best and then $y \in X-\{x\}$ as worst or first chooses $y \in X$ as worst and then $x \in X-\{y\}$ as best. This condition is formalized as:

<u>Definition 2</u>. <u>A</u> concordant choice model <u>is a set of preference and aversion probabilities for the subsets of a set</u> A <u>such that for</u> $x,y \in X \subseteq A$,

i) $p(x,y) \neq 0$,

and ii) $P_X(x)P^*_{X-\{x\}}(y) = P^*_X(y)P_{X-\{y\}}(x)$.

We note that a concordant choice model for a set $A = \{x,y\}$ is any set of binary preference and aversion probabilities such that $p(x,y) = p^*(y,x) \neq 0$.

2.4 Outline of results.

We first show that a set of preference and aversion probabilities for the subsets of a set A satisfies a concordant choice model if and only if the preference and aversion probabilities for any subset $X \subseteq A$ each have a particular representation in terms of the binary preference probabilities $p(x,y)$, $x,y \in X$. These representations are more complicated than those of the strict utility model and are equivalent to the latter if and only if the preference and aversion probabilities associated with any particular presentation set are all equal; we shall see later that this result is strictly stronger than Block and Marschak's result (1960, p. 111).

We then show that any reversible ranking model is a concordant choice model and use this result to prove that a set of preference, aversion and ranking probabilities for the subsets of a set A satisfies a reversible ranking model if and only if the preference, aversion and ranking probabilities for any subset $X \subseteq A$ each have a particular representation in terms of the binary preference probabilities $p(x,y)$, $x,y \in X$.

Finally we discuss the relation of concordant choice models to certain other conditions which have appeared in the mathematical literature on preference. The reader is referred to Fig. 1 for a summary of the results.

3. Concordant Choice Models

<u>Lemma 1</u>. <u>Any concordant choice model is such that all the preference and aversion probabilities are non-zero</u>.

Proof. We proceed by induction, using part i) of Def. 2 as the first step. We may suppose that A has at least three elements and that all the preference and aversion probabilities are non-zero for sets $X \subseteq A$ with less than n elements, $n \geq 2$. Let $Y \subseteq A$ be an n element set and suppose that $P_Y(w) = 0$ for some $w \in Y$. Then by part ii) of Def. 2, for any $y \in Y-\{w\}$,

$$0 = P_Y(w) P_{Y-\{w\}}(y)$$
$$= P^*_Y(y) P_{Y-\{y\}}(w) .$$

By the induction hypothesis, $P_{Y-\{y\}}(w) \neq 0$ for any $y \in Y-\{w\}$ and hence $P^*_Y(y) = 0$ for all $y \in Y-\{w\}$. Reversing the argument, we see that $P_Y(x) = 0$ for all $x \in Y$ which is impossible.

3.1 Concordant choice and strict utility models.

We next prove that the assumption that both the preference and the aversion probabilities for the subsets of a set A satisfy the choice axiom is in general incompatible with the assumption that they satisfy a concordant choice model. The following definition and lemma are basic to any study of the choice axiom.

<u>Definition 3</u>. <u>A set of preference probabilities for the subsets of a set</u> A <u>satisfy the</u> choice axiom <u>provided that for all</u> $x \in Y$ $X \subseteq A$, $P_Y(x) = P_X(x|Y)$.

<u>Lemma 2</u>. <u>If the preference probabilities for the subsets of a finite set</u> A <u>satisfy the choice axiom, then there exists</u>

<u>a ratio scale</u> v <u>on</u> A <u>such that for any</u> $P_Y(x) \neq 0, 1$,

$$P_Y(x) = \frac{v(x)}{\sum_{y \in Y} v(y)} .$$

This is Theorem 31 of Sec. 5.2 of Luce and Suppes (1965). v is called a <u>strict utility scale</u>. Similar statements of Def. 1 and Lemma 2 may be made in terms of the aversion probabilities.

<u>Theorem 1</u>. <u>If the sets of preference and aversion probabilities for the subsets of a finite set</u> A <u>with at least three elements both satisfy the choice axiom and a concordant choice model, then the strict utility scales</u> v <u>and</u> v^* <u>are constant functions</u>.

Proof. By Lemmas 1 and 2 there exist ratio scales v and v^* such that for $x \in X \subseteq A$,

$$P_X(x) = v(x) \Big/ \sum_{y \in X} v(y) > 0 \tag{4}$$

and

$$P^*_X(x) = v^*(x) \Big/ \sum_{y \in X} v^*(y) > 0 . \tag{5}$$

And by Def. 2,

$$P_A(x) P^*_{A-\{x\}}(y) = P^*_A(y) P_{A-\{y\}}(x) .$$

Substituting Eqs. 4 and 5 into this equation, we obtain that

$$\frac{v(x)}{\sum_{z \in A} v(z)} \; \frac{v^*(y)}{\sum_{w \in A-\{x\}} v^*(w)} = \frac{v^*(y)}{\sum_{z \in A} v^*(z)} \; \frac{v(x)}{\sum_{w \in A-\{y\}} v(w)} > 0 .$$

So,

$$\sum_{z\in A} v(z) \sum_{w\in A-\{x\}} v^*(w) = \sum_{z\in A} v^*(z) \sum_{w\in A-\{y\}} v(w) .$$

It follows immediately that

$$\sum_{z\in A} v(z) \sum_{w\in A} v^*(w) - \sum_{z\in A} v(z)\ v^*(x) =$$

$$\sum_{z\in A} v^*(z) \sum_{w\in A} v(w) - \sum_{z\in A} v^*(z) v(y) ,$$

and so

$$v^*(x) \sum_{z\in A} v(z) = v(y) \sum_{z\in A} v^*(z) . \tag{6}$$

Because A has at least three elements there exists an $r\in A$, $r \neq x,y$ and replacing y by r in the above argument we obtain that

$$v^*(x) \sum_{z\in A} v(z) = v(r) \sum_{z\in A} v^*(z) . \tag{7}$$

Equating Eqs. 6 and 7 and dividing by $\sum_{z\in A} v^*(z)$ yields $v(r) = v(y)$. A similar argument shows $v^*(r) = v^*(y)$.

QED.

3.2 Representation theorem.

Given $x \in X \subseteq A$ let

$$f(x,X) = \prod_{y\in X-\{x\}} p(x,y) ,$$

$$f^*(x,X) = \prod_{y\in X-\{x\}} p^*(x,y)$$

and for $\rho \in R(X)$, define

$$F(\rho,X) = \prod_{i=1}^{n-1} f(\rho_i, X - \bigcup_{j<i} \rho_j)$$

and

$$F^*(\rho,X) = \prod_{i=1}^{n-1} f^*(\rho_i, X - \bigcup_{j<i} \rho_j)$$

Lemma 3. *If for all* $x,y \in A$, $p(x,y) = p^*(y,x)$ *then for each* $\rho \in R(X)$, $X \subseteq A$, $F(\rho,X) = F^*(\rho^*,X)$.

Proof. For $X = \{x,y\}$ and $\rho = xy$,

$$\begin{aligned} F(\rho,X) &= f(x,X) \\ &= p(x,y) \\ &= p^*(y,x) \\ &= f^*(y,X) \\ &= F^*(\rho^*,X) \end{aligned}$$

and so the lemma is true for two element sets. For larger A we proceed by induction. Let $X \subseteq A$ be an n element set and for each $\rho \in R(X)$ write $\rho = \rho_1\sigma$ where $\sigma = \rho_2\rho_3 \cdots \rho_{n-1}\rho_n \in R(X-\{\rho_1\})$. Then

$$F(\rho,X) = \prod_{i=1}^{n-1} f(\rho_i, X - \bigcup_{j<i} \rho_j)$$

$$= f(\rho_1,X) \prod_{i=2}^{n-1} f(\rho_i, X - \bigcup_{j<i} \rho_j)$$

$$= f(\rho_1,X) \prod_{i=2}^{n-1} f(\rho_i, X - \{\rho_1\} - \bigcup_{1<j<i} \rho_j)$$

$$= f(\rho_1,X)F(\sigma,X-\{\rho_1\})$$

$$= f(\rho_1,X)F^*(\sigma^*,X-\{\rho_1\}) \quad . \text{ (by induction hypothesis)}$$

Also

$$F^*(\rho^*,X) = \prod_{i=1}^{n-1} f^*(\rho_i^*,X-\bigcup_{j<i}\rho_j)$$

$$= \prod_{i=1}^{n-2} f^*(\rho_i^*,X-\bigcup_{j<i}\rho_j)\ f^*(\rho_{n-1}^*,\{\rho_{n-1}^*,\rho_1\})$$

$$= \prod_{i=1}^{n-2}[f^*(\rho_i^*,X-\{\rho_1\}-\bigcup_{j<i}\rho_j^*)p^*(\rho_i^*,\rho_1)]\ p^*(\rho^*_{n-1},\rho_1)$$

$$= \prod_{i=1}^{n-2} f^*(\rho_i^*,X-\{\rho_1\}-\bigcup_{j<i}\rho_j^*)\prod_{i=1}^{n-1} p(\rho_1,\rho_i^*)$$

$$= F^*(\sigma^*,X-\{\rho_1\})f(\rho_1,X)$$

and thus

$$F(\rho,X) = F^*(\rho^*,X) \qquad \text{QED.}$$

<u>Theorem 2</u>. <u>A set of preference and aversion probabilities for the subsets of a set</u> A <u>satisfies a concordant choice model if and only if for each</u> $x,y \in X \subseteq A$,

i) $p(x,y) = p^*(y,x) \neq 0$,

$$\text{ii) } P_X(x) = \sum_{\rho\in R(x;X)} F(\rho,X) \Big/ \sum_{\sigma\in R(X)} F(\sigma,X) \qquad (8)$$

<u>and</u>

$$\text{iii) } P_X^*(x) = \sum_{\rho\in R(x;X)} F^*(\rho,X) \Big/ \sum_{\sigma\in R(X)} F^*(\sigma,X) \ . \qquad (9)$$

Proof. As pointed out after Def. 2, if $\{P_X, P^*_X : X \subseteq A\}$ satisfies a concordant choice model then for $x,y \in A$, $p(x,y) = p^*(y,x) \neq 0$. Also for any set $X = \{x,y\} \subseteq A$,

$$\sum_{\rho \in R(x;X)} F(\rho,X) \Big/ \sum_{\sigma \in R(X)} F(\sigma,X) = \frac{p(x,y)}{p(x,y)+p(y,x)} = p(x,y) = P_X(x)$$

and Eq. 8 holds. It can be shown similarly that Eq. 9 holds for two element sets and we proceed by induction. If $X \subseteq A$ is an n element set then Lemma 3 implies that

$$\sum_{\rho \in R(X)} F(\rho,X) = \sum_{\rho \in R(X)} F^*(\rho^*,X) = \sum_{\rho \in R(X)} F^*(\rho,X) \quad . \tag{10}$$

and by Lemma 1 we may write part ii) of Def. 2 in the form that for $x,y \in X$.

$$\frac{P_X(x)}{P^*_X(y)} = \frac{P_{X-\{y\}}(x)}{P^*_{X-\{x\}}(y)} \quad .$$

But the induction hypothesis implies that

$$\frac{P_{X-\{y\}}(x)}{P^*_{X-\{x\}}(y)} = \frac{\sum_{\rho \in R(x;X-\{y\})} F(\rho,X-\{y\})}{\sum_{\rho \in R(X-\{y\})} F(\rho,X-\{y\})} \;\; \frac{\sum_{\rho \in R(X-\{x\})} F^*(\rho,X-\{x\})}{\sum_{\rho \in R(y;X-\{x\})} F^*(\rho,X-\{x\})}$$

$$= \frac{\sum_{\rho \in R(X-\{x\})} F^*(\rho,X-\{x\})}{\sum_{\rho \in R(X-\{y\})} F(\rho,X-\{y\})} \;\; \frac{\sum_{\rho \in R(x;X-\{y\})} F(\rho,X-\{y\})}{\sum_{\rho \in R(y;X-\{x\})} F^*(\rho,X-\{x\})}$$

(by rearranging terms)

$$= \frac{\displaystyle\sum_{\rho\in R(X-\{x\})} F(\rho,X-\{x\})}{\displaystyle\sum_{\rho\in R(X-\{y\})} F^*(\rho,X-\{y\})} \; \frac{\displaystyle\sum_{\sigma\in R(X-\{x,y\})} f(x,X-\{y\})F(\sigma,X-\{y,x\})}{\displaystyle\sum_{\sigma\in R(X-\{x,y\})} f^*(y,X-\{x\})F^*(\sigma,X-\{y,x\})}$$

(using Eq. 10 and the definition of F)

$$= \frac{\displaystyle\sum_{\rho\in R(X-\{x\})} F(\rho,X-\{x\})}{\displaystyle\sum_{\rho\in R(X-\{y\})} F^*(\rho,X-\{y\})} \; \frac{f(x,X-\{y\})\displaystyle\sum_{\sigma\in R(X-\{x,y\})} F(\sigma,X-\{x,y\})}{f^*(y,X-\{x\})\displaystyle\sum_{\sigma\in R(X-\{x,y\})} F^*(\sigma,X-\{x,y\})}$$

$$= \frac{\displaystyle\sum_{\rho\in R(X-\{x\})} f(x,X-\{y\})F(\rho,X-\{x\})}{\displaystyle\sum_{\rho\in R(X-\{y\})} f^*(y,X-\{x\})F^*(\rho,X-\{y\})}$$

(using Eq. 10 and collecting terms)

$$= \frac{p(x,y)\displaystyle\sum_{\rho\in R(X-\{x\})} f(x,X-\{y\})F(\rho,X-\{x\})}{p^*(y,x)\displaystyle\sum_{\rho\in R(X-\{y\})} f^*(y,X-\{x\})F^*(\rho,X-\{y\})}$$

[because $p(x,y) = p^*(y,x) \neq 0$]

$$= \frac{\displaystyle\sum_{\rho\in R(X-\{x\})} f(x,X)F(\rho,X-\{x\})}{\displaystyle\sum_{\rho\in R(X-\{y\})} f^*(y,X)F^*(\rho,X-\{y\})}$$

(from the definition of f)

$$= \frac{\sum_{\sigma \in R(x;X)} F(\sigma,X)}{\sum_{\sigma \in R(y;X)} F^*(\sigma,X)} ,$$

and thus

$$\frac{P_X(x)}{P^*_X(y)} = \frac{\sum_{\rho \in R(x;X)} F(\rho,X)}{\sum_{\rho \in R(y;X)} F^*(\rho,X)} .$$

Inverting both sides of this equation, summing over $y \in X$ and using Eq. 10 we have for each $x \in X$,

$$P_X(x) = \sum_{\rho \in R(x;X)} F(\rho,X) \Big/ \sum_{\rho \in R(x)} F^*(\rho,X)$$

$$= \sum_{\rho \in R(x;X)} F(\rho,X) \Big/ \sum_{\rho \in R(X)} F(\rho,X)$$

and similarly summing over $x \in X$ and using Eq. 10 we obtain that for each $y \in X$,

$$P^*_X(y) = \sum_{\rho \in R(y:X)} F^*(\rho,X) \Big/ \sum_{\rho \in R(X)} F^*(\rho,X) .$$

Now assume that conditions i), ii) and iii) hold. Then clearly for all $x \in X \subseteq A$, $P_X(x) \neq 0$ and $P^*_X(x) \neq 0$. These conditions with Lemma 3 imply that

$$\frac{P_X(x)}{P^*_X(y)} = \frac{\sum_{\rho\in R(x;X)} F(\rho,X)}{\sum_{\rho\in R(y;X)} F^*(\rho,X)}$$

and

$$\frac{P_{X-\{y\}}(x)}{P^*_{X-\{x\}}(y)} = \frac{\sum_{\rho\in R(x;X-\{y\})} F(\rho,X-\{x\})}{\sum_{\rho\in R(X-\{y\})} F(\rho,X-\{y\})} \quad \frac{\sum_{\rho\in R(X-\{x\})} F^*(\rho,X-\{x\})}{\sum_{\rho\in R(y;X-\{x\})} F^*(\rho,X-\{x\})} .$$

Lemma 3 implies that Eq. 10 holds and an examination of the first part of the present theorem shows that this equation is sufficient to prove that the right-hand sides of the above two equations are equal and hence

$$\frac{P_X(x)}{P^*_X(y)} = \frac{P_{X-\{y\}}(x)}{P^*_{X-\{x\}}(y)}$$

i.e. $P_X(x)P^*_{X-\{x\}}(y) = P^*_X(y)P_{X-\{y\}}(x)$ QED.

3.3 Regularity and concordant choice models.

A condition on non-binary choice probabilities which has certain rather surprising implications [see Marley (1965) and Chap. 5 of this thesis] is <u>regularity</u> which imposes the condition that adding a new alternative to a choice set never increases the probability of choosing an old alternative. We formalize this condition and then consider its relation to concordant choice models.

<u>Definition 4.</u> <u>A set of preference probabilities for the subsets of a set</u> A <u>is</u> regular <u>if for all</u> x, X <u>and</u> Y <u>such that</u> $x \in X \subseteq Y \subseteq A$ <u>then</u> $P_X(x) \geq P_Y(x)$. <u>A set of aversion</u>

<u>probabilities for the subsets of a set</u> A <u>is</u> regular <u>if for all</u> x, X <u>and</u> Y <u>such that</u> $x \in X \subseteq Y \subseteq A$ <u>then</u> $P^*_X(x) \geq P^*_Y(x)$

It is clear that $\{P_X : X \subseteq A\}$ is regular if and only if for all $x,y \in Y \subseteq A$, $P_{Y-\{y\}}(x) \geq P_Y(x)$ and similarly $\{P^*_X : X \subseteq A\}$ is regular if and only if for all $x,y \in Y \subseteq A$, $P^*_{Y-\{y\}}(x) \geq P^*_Y(x)$. We use the conditions in the latter form to prove the next theorem.

<u>Theorem 3</u>. <u>If a set of preference and aversion probabilities for the subsets of a set</u> A <u>satisfy a concordant choice model then the preference probabilities are regular if and only if the aversion probabilities are regular</u>.

Proof. If $\{P_X : X \subseteq A\}$ is regular then for arbitrary $x,y \in Y \subseteq A$,

$$P_{Y-\{y\}}(x) \geq P_Y(x) \tag{11}$$

and by part ii) of Def. 2,

$$P_Y(x)P^*_{Y-\{x\}}(y) = P^*_Y(y)P_{Y-\{y\}}(x) . \tag{12}$$

Lemma 1 implies that all the preference and aversion probabilities are non-zero and thus Eqs. 11 and 12 imply that $P^*_{Y-\{x\}}(y) \geq P^*_Y(y)$ which shows that $\{P^*_X : X \subseteq A\}$ is regular. The proof of the theorem is completed by reversing the argument.

QED.

<u>Theorem 4</u>. <u>Sets of preference and aversion probabilities that satisfy a concordant choice model need not be regular. Moreover, regular preference and aversion probabilities for the subsets of a set</u> A <u>need not satisfy a concordant choice model</u>.

Proof. Let $A = \{x,y,z\}$ and assume that $\{P_X, P_X^*: X \subseteq A\}$ satisfies a concordant choice model. Theorem 2 shows that

$$P_A(x) = \frac{p(x,y)p(x,z)}{p(x,y)p(x,z)+p(z,y)p(z,x)+p(y,z)p(y,x)} .$$

If $p(x,y) = 1/2$, $p(x,z) = p(z,y) = 3/4$ then the above equation implies that $P_A(x) = 6/11 > 1/2 = p(x,y)$. Hence $\{P_X: X \subseteq A\}$ is not regular and Theorem 3 implies that $\{P_X^*: X \subseteq A\}$ is not regular.

Conversely, if both $\{P_X: X \subseteq A\}$ and $\{P_X^*: X \subseteq A\}$ satisfy the choice axiom with all the choice probabilities non-zero, then Lemma 2 and the corresponding Lemma for $\{P_X^*: X \subseteq A\}$ imply that both $\{P_X: X \subseteq A\}$ and $\{P_X^*: X \subsetneq A\}$ are regular. However, if for some $x,y \in A$, $P_A(x) \neq P_A(y)$ then the strict utility scale v is not a constant function and Theorem 1 implies that $\{P_X, P_X^*: X \subsetneq A\}$ does not satisfy a concordant choice model. QED.

4. Reversible Ranking Models

Theorem 5. A set of preference, aversion, and ranking probabilities for the subsets of a set A satisfy a reversible ranking model if and only if the preference and discard probabilities satisfy a concordant choice model and for each $\sigma \in R(X)$, $X \subseteq A$,

$$p(\sigma) = F(\sigma,X) \Big/ \sum_{\rho \in R(X)} F(\rho,X) \tag{13}$$

and

$$p^*(\sigma) = F^*(\sigma,X) \Big/ \sum_{\rho \in R(X)} F^*(\rho,X) \ . \tag{14}$$

Proof. It is immediate from the definition of a reversible ranking model that for $x,y \in A$, $p(x,y) = p^*(y,x) \neq 0$ and thus if A is a two element set the probabilities satisfy a concordant choice model. If A has more than two elements then for arbitrary $x,y \in A$ let Y be any set such that $\{x,y\} \subset Y \subseteq A$. Choose an arbitrary $\sigma \in R(Y-\{x,y\})$ and let $\rho = x\sigma y$. Then $\rho \in R(Y)$ and it follows from the definition of a reversible ranking model that

$$\begin{aligned}
P_Y(x)P^*_{Y-\{x\}}(y)p^*(\sigma^*) &= P_Y(x)p^*(y\sigma^*) \\
&= P_Y(x)p(\sigma y) \\
&= p(x\sigma y) \\
&= p^*(y\sigma^* x) \\
&= P^*_Y(y)p^*(\sigma^* x) \\
&= P^*_Y(y)p(x\sigma) \\
&= P^*_Y(y)P_{Y-\{y\}}(x)p(\sigma)
\end{aligned}$$

$$= P^*_Y(y)P_{Y-\{y\}}(x)p^*(\sigma^*) \ .$$

But $p^*(\sigma^*) \neq 0$ and hence for arbitrary $x,y \in Y \subsetneq A$,

$$P_Y(x)P^*_{Y-\{x\}}(y) = P^*_Y(y)P_{Y-\{y\}}(x)$$

which with the fact that $p(x,y) = p^*(y,x) \neq 0$ shows that the probabilities satisfy a concordant choice model.

Also for $X = \{x,y\} \subsetneq A$ and $\sigma = xy$, it follows from the definitions of F and of a reversible ranking model that

$$F(\sigma,X) \Big/ \sum_{\rho \in R(X)} F^*(\rho,X) = \frac{p(x,y)}{p(x,y)+p(y,x)} = p(x,y) = \mathfrak{p}(\rho)$$

and

$$F^*(\sigma,X) \Big/ \sum_{\rho \in R(X)} F^*(\rho,X) = p^*(\rho) \ .$$

Thus Eqs. 13 and 14 hold for subsets of size two, so we make the induction hypothesis that they hold for all sets $Y \subset A$ with less than n elements, $n > 2$. Let $X \subsetneq A$ be an n element set. Each $\sigma \in R(X)$ has the form $\sigma = x\mu$ for some $x \in X$, $\mu \in R(X-\{x\})$ and the reversible ranking model implies that $p(\sigma) = P_X(x)p(\mu)$. Using Theorem 2 and the first part of the present theorem to obtain the representation of $P_X(x)$ and the induction hypothesis for the form of $p(\mu)$, we obtain that

$$\mathfrak{p}(\sigma) = P_X(x)p(\mu)$$

$$= \frac{\sum\limits_{\rho\in R(x;X)} F(\rho,X)}{\sum\limits_{\rho\in R(X)} F(\rho,X)} \quad \frac{F(\mu,X-\{x\})}{\sum\limits_{\rho\in R(X-\{x\})} F(\rho,X-\{x\})}$$

$$= \frac{\sum\limits_{\rho\in R(X-\{x\})} f(x,X)F(\rho,X-\{x\})}{\sum\limits_{\rho\in R(X-\{x\})} F(\rho,X-\{x\})} \quad \frac{F(\mu,X-\{x\})}{\sum\limits_{\rho\in R(X)} F(\rho,X)}$$

$$= \frac{f(x,X)F(\mu,X-\{x\})}{\sum\limits_{\rho\in R(X)} F(\rho,X)}$$

$$= \frac{F(\sigma,X)}{\sum\limits_{\rho\in R(X)} F(\rho,X)} \; .$$

We can use similar arguments to derive the form of $p^*(\sigma)$ given by Eq. 14.

Now suppose that the preference and aversion probabilities for the subsets of A satisfy a concordant choice model and that Eqs. 13 and 14 hold for all rank orders. Then from Eqs. 13 and 14, the definition of F and the convention that $P_{\{z\}}(z) = 1 = P^*_{\{z\}}(z)$ for all $z \in A$, we obtain that for any set $X = \{x,y\} \subseteq A$ and $\sigma = xy$,

$$p(\sigma) = \frac{p(x,y)}{p(x,y)+p(y,x)}$$
$$= p(x,y)$$
$$= P_X(x)P_{X-\{x\}}(y)$$

and similarly that

$$p^*(\sigma) = p^*(x,y) = P^*_X(x)P^*_{X-\{x\}}(y) \ .$$

Theorem 2 implies that $p(x,y) = p^*(y,x) \neq 0$ which with the above shows that $p(\sigma) = p^*(\sigma^*) \neq 0$ and hence the preference, aversion and ranking probabilities for subsets of size two satisfy a reversible ranking model. So we suppose that A has at least three elements and make the induction hypothesis that the preference, aversion, and ranking probabilities for sets $Y \subset A$ with less than n elements, $n > 2$, satisfy a reversible ranking model. Let $X \subseteq A$ be an n element set. By assumption, $p(\sigma)$, $p(\mu)$ satisfy Eq. 13 and because the preference and aversion probabilities satisfy a concordant choice model, Theorem 2 implies that $P_X(\sigma_1)$ satisfies Eq. 8. The argument of the first part of the theorem shows that under these conditions $p(\sigma) = P_X(\sigma_1)p(\mu)$ and using the induction hypothesis that $p(\mu) = P_{X-\{\sigma_1\}}(\sigma_2)$ $P_{X-\{\sigma_1\sigma_2\}}(\sigma_3)\ldots p(\sigma_{n-1},\sigma_n)$, we obtain that

$$p(\sigma) = P_X(\sigma_1)P_{X-\{\sigma_1\}}(\sigma_2)\ldots p(\sigma_{n-1},\sigma_n) \ . \qquad (15)$$

In similar fashion it can be shown that

$$p^*(\sigma) = P^*_X(\sigma_1)P^*_{X-\{\sigma_1\}}(\sigma_2)\ldots p^*(\sigma_{n-1},\sigma_n) \ . \qquad (16)$$

Writing σ as $\sigma = \sigma_1 w \sigma_n$ where $w = \sigma_2\sigma_3\ldots\sigma_{n-2}\sigma_{n-1} \in R(X-\{\sigma_1\sigma_n\})$ then using Eqs. 15 and 16, the fact that the preference and aversion probabilities for the subsets of A satisfy a concordant choice model and the induction hypothesis that $p(\eta) = p^*(\eta^*)$ for $\eta \in R(Y)$, $Y \quad X$, we obtain that

$$
\begin{aligned}
p(\sigma) &= p(\sigma_1 w \sigma_n) \\
&= P_X(\sigma_1)\, p(w\sigma_n) \\
&= P_X(\sigma_1)\, p^*(\sigma_n w^*) \\
&= P_X(\sigma_1)\, P^*_{X-\{\sigma_1\}}(\sigma_n)\, p^*(w^*) \\
&= P^*_X(\sigma_n)\, P_{X-\{\sigma_n\}}(\sigma_1)\, p^*(w^*) \\
&= P^*_X(\sigma_n)\, P_{X-\{\sigma_n\}}(\sigma_1)\, p(w) \\
&= P^*_X(\sigma_n)\, p(\sigma_1 w) \\
&= P^*_X(\sigma_n)\, p^*(w^*\sigma_1) \\
&= p^*(\sigma_n w^* \sigma_1) \\
&= p^*(\sigma^*) \;.
\end{aligned}
$$

We have already shown that $p(\sigma) \neq 0$, hence the above result with Eqs. 15 and 16 proves that the preference, aversion and ranking probabilities for the subsets of X satisfy a reversible ranking model. QED.

It is immediate from Theorem 5 that any reversible ranking model is a concordant choice model but not conversely.

5. The Discard and Acceptance Conditions

Luce (1960) suggested that if a subject has to choose the most preferred element in the available set X then he may pick some element $y \in X$ as not worthy of further consideration thereby reducing his choice problem to the set $X - \{y\}$. He repeats the elimination process until a two element set remains, at which point he makes his final decision. Letting $Q_X^*(y)$, $y \in X$ denote the probability that the subject decides that y is not worthy of further consideration relative to the elements of X, then Luce's suggestion may be formalized as the assumption that for each $x \in X$,

$$P_X(x) = \sum_{y \in X-\{x\}} Q_X^*(y) P_{X-\{y\}}(x) . \tag{17}$$

The subject need not decide that y is the worst element in X in order to eliminate y as not worthy of further consideration but has simply to decide that some element $z \in X$ is better than y. Thus the probability $Q_X^*(y)$ need not equal the aversion probability $P_X^*(y)$.

However, if a set of preference and aversion probabilities for the subsets of a set A satisfy a concordant choice model, then for each $x,y \in X \subseteq A$,

$$P_X(x) P_{X-\{x\}}^*(y) = P_X^*(y) P_{X-\{y\}}(x)$$

and thus

$$P_X(x) = P_X(x) \sum_{y \in X-\{x\}} P_{X-\{x\}}^*(y)$$

$$= \sum_{y \in X-\{x\}} P_X(x) P_{X-\{x\}}^*(y)$$

$$= \sum_{y \in X-\{x\}} P^*_X(y) P_{X-\{y\}}(x) \ , \tag{18}$$

which is the special case of Eq. 17 in which the probability $Q^*_X(y)$ of choosing y as not worthy of further consideration relative to X is in fact equal to the aversion probability $P^*_X(y)$.

Luce did not consider the case in which the subject's task is to choose the least preferred element in the set X but a natural extension of his argument to this case leads to the assumption that there exists a distribution Q_X such that for each $x \in X$,

$$P^*_X(x) = \sum_{y \in X-\{x\}} Q_X(y) P^*_{X-\{y\}}(x) \ . \tag{19}$$

Using arguments similar to those used to obtain Eq. 18 we can show that if the preference and aversion probabilities for the subsets of a set A satisfy a concordant choice model then for $x \in X \subseteq A$,

$$P^*_X(x) = \sum_{y \in X-\{x\}} P_X(y) P^*_{X-\{y\}}(x) \tag{20}$$

which is the special case of Eq. 19 in which $Q_X \equiv P_X$.

We have seen that any concordant choice model satisfies Eqs. 18 and 20 and in this section we show that if the binary preference probabilities are all non-zero then Eqs. 18 and 20 are in fact equivalent to a concordant choice model.

5.1 Definitions and relations between the defined quantities.

Definition 5. A set of preference probabilities for the subsets of a set A satisfies the discard condition if there exist probability measures Q^*_X , $X \subseteq A$ such that for each $x \in X \subseteq A$,

$$P_X(x) = \sum_{y \in X-\{x\}} Q^*_X(y) P_{X-\{y\}}(x) .$$

A set of preference and aversion probabilities for the subsets of a set A satisfies the strong discard condition if it satisfies the discard condition with $Q^*_X \equiv P^*_X$ for each $X \subseteq A$ ..

Definition 6. A set of aversion probabilities for the subsets of a set A satisfy the acceptance condition if there exist probability measures Q_X , $X \subseteq A$ such that for each $x \in X \subseteq A$,

$$P^*_X(x) = \sum_{y \in X-\{x\}} Q_X(y) P^*_{X-\{y\}}(x) .$$

A set of preference and aversion probabilities for the subsets of a set A satisfies the strong acceptance condition if it satisfies the acceptance condition with $Q_X \equiv P_X$.

The strong discard and the strong acceptance conditions may appear tautological but examples are easily constructed which demonstrate that they are not. Such an example arises in the proof of Theorem 9.

We use the following result of Luce (1960, Theorem 1) in proving the next two theorems.

Lemma 4. Given a set A such that the binary choice probabilities $p(x,y) \neq 0$, $x,y \in A$ and the preference probabil-

<u>ities satisfy the choice axiom, then for any</u> n <u>element subset</u> $X \subseteq A$, $Q^*_X(x) = [1-P_X(x)]/(n-1)$ <u>is the unique solution of the discard condition.</u>

Because the discard condition does not involve the aversion probabilities and the acceptance condition does not involve the preference probabilities it is obvious that in general neither of these conditions implies the other. This result is not so obviously true when the binary choice probabilities for the set A satisfy $p(x,y) = p^*(y,x)$ for $x,y \in A$. Theorem 6 demonstrates that the result is also true in this case.

<u>Theorem 6</u>. <u>The discard condition neither implies nor is implied by the acceptance condition.</u>

Proof. Let $A = \{x,y,z\}$ and suppose that $P_A(x) = 1/6$, $p(y,z) = p^*(z,y) = 2/5$, $p(x,z) = p^*(z,x) = P^*_X(x) = P^*_X(y) = 1/4$, $p(x,y) = p^*(y,x) = P_A(y) = 1/3$, $P_A(z) = P^*_X(z) = 1/2$. The preference probabilities $\{P_X : X \subseteq A\}$ satisfy the choice axiom and Lemma 4 implies that the preference probabilities satisfy the discard condition. If we let $\underline{P}^* = [P^*_A(x), P^*_A(y), P^*_A(z)]$, $\underline{Q} = [Q_A(x), Q_A(y), Q_A(z)]$, $a = p^*(x,y)$, $b = p^*(x,z)$ and $c = p^*(y,z)$ then the acceptance condition becomes

$$\underline{P}^* = \underline{M}\underline{Q}$$

where

$$\underline{M} = \begin{bmatrix} 0 & b & a \\ c & 0 & 1-a \\ 1-c & 1-b & 0 \end{bmatrix} .$$

$\underline{M}^{-1}$ exists if and only if $K = a(1-b)c + (1-a)(1-c)b \neq 0$

in which case

$$\underline{M}^{-1} = \frac{1}{K} \begin{bmatrix} -(1-a)(1-b) & a(1-b) & (1-a)b \\ (1-a)(1-c) & -a(1-c) & ac \\ (1-b)c & b(1-c) & -bc \end{bmatrix}.$$

Computing $\underline{Q} = \underline{M}^{-1}P*$ with the values for a,b,c given above, we obtain

$$\underline{Q} = \left(\frac{35}{48}, \frac{40}{48}, -\frac{27}{48} \right),$$

which is not a probability distribution. Hence the acceptance condition does not hold.

A similar example shows that the acceptance condition does not imply the discard condition. QED.

Theorem 7. The strong discard condition is strictly stronger than the discard condition and the strong acceptance condition is strictly stronger than the acceptance condition.

Proof. Let $A = \{x,y,z\}$, $p(x,y) = p*(y,x) = 1/3$, $p(x,z) = p*(z,x) = 1/4$, $p(y,z) = p*(z,y) = 2/5$, $P_A(x) = P^*_A(x) = 1/6$, $P_A(y) = P^*_A(y) = 1/3$, $P_A(z) = P^*_A(z) = 1/2$. The preference probabilities satisfy the choice axiom and Lemma 4 implies that the preference probabilities satisfy the discard condition, but

$$P^*_A(x) = \frac{1}{6} \neq \frac{5}{12} = \frac{1-P_X(x)}{2} = Q^*_A(x)$$

which shows that the preference and aversion probabilities do not satisfy the strong discard condition.

A similar example can be constructed to show that the strong acceptance condition is strictly stronger than the acceptance condition. QED.

<u>Theorem 8</u>. <u>The strong discard condition neither implies nor is implied by the strong acceptance condition</u>.

Proof. Let $A = \{x,y,z\}$ and assume that $p(x,z) = p(y,z) = 1/3$, $p(x,y) = P_A(z) = 1/2$, $P_A(x) = P_A(y) = 1/4$ and for each $w \in X \subseteq A$, $P^*_X(w) = \frac{1-P_X(w)}{|X|-1}$ where $|X|$ denotes the number of elements in the set X. Then the preference probabilities satisfy the choice axiom and Lemma 4 implies that the preference and discard probabilities satisfy the strong discard condition. However,

$$P^*_A(x) = 3/8 \neq 5/12 = \sum_{w\in A-\{x\}} P_A(w) P^*_{A-\{w\}}(x)$$

which shows that the strong acceptance condition is not satisfied.

A similar example shows that the strong acceptance condition does not imply the strong discard condition. QED.

5.2 Concordant choice models and the strong discard and acceptance conditions.

The main result of this section is that if the binary preference probabilities between elements of a set A are all non-zero then the preference and aversion probabilities for the subsets of A satisfy a concordant choice model if and only if these probabilities satisfy the strong discard and the strong acceptance conditions.

For notational simplicity we define for any $x \in X$, $P_{X-\{x\}}(x) = 0 = P^*_{X-\{x\}}(x)$ and P_X is <u>positive</u> if $P_X(x) > 0$ for all $x \in X$. The proof of the main theorem is dependent on the following lemma.

<u>Lemma 5</u>. <u>Let</u> X <u>be a set with at least two elements and</u>

<u>let</u> P_Y , P_Y^* <u>be any positive probabilities for the sets</u> $Y = X-\{y\}$, $y \in X$. <u>If probability distributions</u> Q_X , Q_X^* <u>exist such that for each</u> $x \in X$

$$Q_X(x) = \sum_{y \in X-\{x\}} Q_X^*(y) P_{X-\{y\}}(x)$$

<u>and</u>

$$Q_X^*(x) = \sum_{y \in X-\{x\}} Q_X(y) P_{X-\{y\}}^*(x)$$

<u>then they are unique</u>.

Proof. Suppose that $X = \{x_1 x_2 \ldots x_n\}$. If Q_X , Q_X^* satisfy the conditions of the theorem then for each $x_j \in X$, $j = 1,2\ldots n$,

$$-Q_X(x_j) + \sum_{y \in X-\{x_j\}} Q_X^*(y) P_{X-\{y\}}(x_j) = 0$$

and

$$-Q_X^*(x_j) + \sum_{y \in X-\{x_j\}} Q_X(y) P_{X-\{y\}}^*(x_j) = 0 \ .$$

Writing these equations in matrix form gives $\underline{PQ} = 0$ where

$$\underline{P} = \begin{bmatrix} -1 & 0 & \cdots & 0 & 0 & P_{X-\{x_2\}}(x_1) & \cdots & P_{X-\{x_n\}}(x_1) \\ 0 & -1 & \cdots & 0 & P_{X-\{x_1\}}(x_2) & 0 & & P_{X-\{x_n\}}(x_2) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -1 & P_{X-\{x_1\}}(x_n) & P_{X-\{x_2\}}(x_n) & & 0 \\ 0 & P^*_{X-\{x_2\}}(x_1) & \cdots & P^*_{X-\{x_n\}}(x_1) & -1 & 0 & \cdots & 0 \\ P^*_{X-\{x_1\}}(x_2) & 0 & & P^*_{X-\{x_n\}}(x_2) & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & & \vdots \\ P^*_{X-\{x_1\}}(x_n) & P^*_{X-\{x_2\}}(x_n) & \cdots & 0 & 0 & 0 & \cdots & -1 \end{bmatrix}$$

and $\underline{Q}$ is a column vector with components

$$[Q_X(x_1), Q_X(x_2) \ldots Q_X(x_n), Q^*_X(x_1), Q^*_X(x_2) \ldots Q^*_X(x_n)] \quad .$$

The last row of the matrix $\underline{P}$ is equal to minus the sum of the remaining rows and we therefore study the reduced matrix in which the last row is identically zero and the other rows are the same as in the original matrix. Let μ_i, $i = 1,2...2n-1$ be any set of real numbers such that the weighted sum of the first $2n-1$ rows of the reduced matrix is zero. We prove that the μ_i are all zero, which implies that the reduced matrix has rank $2n-1$ and hence the initial matrix equation has at most one solution.

We first show that if $\mu_k = \max_{1 \leq i \leq 2n-1} \mu_i$ then there exists $1 \leq j < n$ such that $\mu_j = \mu_k$. The proof has three parts.

i) $k < n$. In this case the statement is clearly true.

ii) $k = n$. We know that $\mu_{n+1} \leq \mu_k$ and if $\mu_{n+1} < \mu_k$ then we obtain from the n^{th} column of the reduced matrix that

$$0 = -\mu_n + \sum_{i=1}^{n-1} \mu_{n+i} P^*_{X-\{x_n\}}(x_i)$$

$$< -\mu_n + \sum_{i=1}^{n-1} \mu_k P^*_{X-\{x_n\}}(x_i)$$

$$= -\mu_k + \mu_k \sum_{i=1}^{n-1} P^*_{X-\{x_n\}}(x_i)$$

$$= 0$$

which is a contradiction. Hence $\mu_{n+1} = \mu_k$.

Now if $\mu_j < \mu_k$ for all $j < n$ then applying the above equality to the $n+1^{st}$ column of the matrix, we obtain that

$$0 = -\mu_{n+1} + \sum_{j=1}^{n} \mu_j P_{X-\{x_1\}}(x_j)$$

$$< -\mu_{n+1} + \sum_{j=1}^{n} \mu_k P_{X-\{x_1\}}(x_j)$$

$$= -\mu_k + \mu_k \sum_{j=1}^{n} P_{X-\{x_1\}}(x_j)$$

$$= 0$$

which is a contradiction.

iii) $k > n$. If $\mu_j < \mu_k$ for all $j < n$ then we obtain from the k^{th} column of the matrix that

$$0 = -\mu_k + \sum_{i=1}^{n} \mu_i P_{X-\{x_{k-n}\}}(x_i)$$

$$< -\mu_k + \sum_{i=1}^{n} \mu_k P_{X-\{x_{k-n}\}}(x_i)$$

$$= 0$$

which is a contradiction.

It is immediate from i)-iii) that we may assume $k < n$. If the μ_i , $1 \leq i \leq 2n-1$ are not all zero, then without loss of generality we may assume that $\mu_k > 0$ and we obtain from the k^{th} column of the matrix that

$$0 = -\mu_k + \sum_{i=1}^{n-1} \mu_{n+i} P^*_{X-\{x_k\}}(x_i)$$

$$\leq -\mu_k + \sum_{i=1}^{n-1} \mu_k P^*_{X-\{x_k\}}(x_i)$$

$$= -\mu_k P^*_{X-\{x_k\}}(x_n)$$

$$< 0$$

which is a contradiction. QED.

Theorem 9. A set of preference and aversion probabilities for the subsets of a set A satisfy a concordant choice model if and only if

i) $p(x,y) \neq 0$ for $x,y \in A$

and

ii) the preference and aversion probabilities for the subsets of A satisfy the strong discard and the strong acceptance conditions.

Proof. It is part of the definition of a concordant choice model that the binary preference probabilities satisfy condition i) and we showed in the introduction to Sec. 5 that any concordant choice model satisfies condition ii). Therefore we have to show that conditions i) and ii) are sufficient for the preference and aversion probabilities to satisfy a concordant choice model.

The strong discard condition with condition i) implies that for $X = \{x,y\} \subseteq A$,

$$p(x,y) = p^*(y,x) \neq 0 \tag{21}$$

which shows that the preference and aversion probabilities for two element sets satisfy a concordant choice model. So we suppose that A has more than two elements and make the induction hypothesis that the preference and aversion probabilities for sets $Y \subset A$ with less than n elements, $n > 2$, satisfy a concordant choice model. Theorem 2 then implies that the preference and aversion probabilities for such sets satisfy Eqs. 8 and 9 and are non-zero. If $X \subseteq A$ is an n element set and P_X, P^*_X are the preference and aversion probabilities associated with X, then condition ii) implies that for $x \in X$,

$$P_X(x) = \sum_{y \in X-\{x\}} P^*_X(y) P_{X-\{y\}}(x) \tag{22}$$

and

$$P^*_X(x) = \sum_{y \in X-\{x\}} P_X(y) P^*_{X-\{y\}}(x) . \tag{23}$$

Let $\overline{P}_X$, $\overline{P}^*_X$ be the sets of choice probabilities given by Eqs. 8 and 9 when the arguments of F and F^* are $p(x,y)$ and $p^*(x,y)$, respectively, $x,y \in X \subseteq A$. Then the set $\{P_Y, P^*_Y; \overline{P}_X, \overline{P}^*_X : Y \subseteq X\}$ satisfies Eqs. 8, 9 and 21 and Theorem 2 implies that it satisfies a concordant choice model. It then follows from the first part of the present theorem that this set satisfies the strong discard and the strong acceptance condition and in particular for $x \in X$,

$$\overline{P}_X(x) = \sum_{y \in X-\{x\}} \overline{P}^*_X(y) P_{X-\{y\}}(x) \tag{24}$$

and

$$\overline{P}^*_X(x) = \sum_{y \in X-\{x\}} \overline{P}_X(y) P^*_{X-\{y\}}(x) . \tag{25}$$

But $P_{X-\{y\}}, P^*_{X-\{y\}}$, $y \in X$, are all positive distributions and Lemma 5 applied to Eqs. 22-25 implies that $P_X \equiv \bar{P}_X$, $P^*_X \equiv \bar{P}^*_X$, i.e. the preference and aversion probabilities for X satisfy Eqs. 8 and 9. By the induction hypothesis the preference and aversion probabilities for sets $Y \subset X$ satisfy Eqs. 8 and 9 and we know that Eq. 21 holds for the binary choice probabilities $p(x,y), p^*(x,y)$, $x,y \in X$. Theorem 2 then implies that the set of preference and aversion probabilities $\{P_Y, P^*_Y : Y \subseteq X\}$ satisfies a concordant choice model.

QED.

5.3 Regularity and the discard and acceptance conditions.

No necessary and sufficient conditions are known for a set of preference probabilities to satisfy the discard condition or for a set of aversion probabilities to satisfy the acceptance condition but Marley (1965) proved the following sufficiency theorem.

Theorem 10. *If a set of preference probabilities is regular, then it satisfies the discard condition but not conversely. If a set of aversion probabilities is regular, then it satisfies the acceptance condition but not conversely.*

The first part of Theorem 10 is Theorems 1 and 7 of Marley (1965). Similar techniques may be used to prove the second part.

We next prove that regularity is independent of the strong discard and acceptance conditions.

Theorem 11. *Neither the strong discard nor the strong acceptance condition on the subsets of a set* A *implies or is implied by regularity of the preference or aversion probabilities for the subsets of* A.

Proof. Assume that the preference and aversion probabilities for the subsets of A satisfy the strong discard and the strong acceptance condition and that $p(x,y) \neq 0$ for any $x,y \in A$. Theorem 9 shows that such a set of probabilities satisfies a concordant choice model and Theorem 4 implies that neither the preference nor the aversion probabilites need be regular. Thus neither the strong discard nor the strong acceptance condition alone can imply the regularity of the preference or aversion probabilities for the subsets of A.

Conversely, let $A = \{x,y,z\}$ and suppose that $p(r,s) = p^*(s,r) = 1/2$ for $r,s \in A$, $P_X(z) = P^*_X(z) = 1/2$, $P_X(y) = P^*_X(y) = 1/3$ and $P_X(x) = P^*_X(x) = 1/6$. Then both the preference and aversion probabilities for the subsets of A are regular but

$$P_X(x) = 1/6 \neq 5/12 = P^*_X(y)p(x,z) + P^*_X(z)p(x,y)$$

which contradicts the strong discard condition and

$$P^*_X(x) = 1/6 \neq 5/12 = P_X(y)p^*(x,z) + P_X(z)p^*(x,y)$$

which contradicts the strong acceptance condition. QED.

6. Joint Independent Random Utility Models

We have studied certain observable relations between the preference, aversion and ranking probabilities for the subsets of a set A but we have not suggested any psychological process which leads to choice probabilities satisfying these conditions. In this section we study a process which can generate sets of preference and aversion probabilities that satisfy either the strong discard or the strong acceptance conditions but not sets of probabilities that satisfy both conditions. It then follows from Theorem 9 that this process cannot generate sets of preference and aversion probabilities that satisfy a concordant choice model.

The process that we study is that which leads to the <u>random utility models</u> [Luce and Suppes (1965), Sec. 5.3, pp. 337-339]. Suppose that a subject's task is to choose the most preferred element in a set X. Then the random utility models assume that each element of X has a 'value' associated with it and the subject chooses that element which has the largest value at the time of choice, the value of each element of X varying with repeated choices according to some probability mechanism. Specifically, the independent random utility model assumes that with each element $x \in X$ there is associated a density $f_x(t)$ such that

$$P_X(x) = \int_{-\infty}^{\infty} f_x(t) \prod_{y \in X-\{x\}} F_y(t)\,dt\ , \tag{26}$$

where F_y is the cumulative distribution of the density f_y. If the subject's task were to choose the least preferred element in the set X then it is reasonable that the mechanism should be similar to that suggested for choice of the most preferred element. We therefore assume that the subject chooses as least preferred that element which has the smallest value at the time of choice and that the probability mechanism generating the values in this case is the same as in the previous case, i.e. for each $x \in X$,

$$P_X^*(x) = \int_{-\infty}^{\infty} f_x(t) \prod_{y \in X - \{x\}} [1 - F_y(t)]dt \,. \tag{27}$$

These conditions are formalized as:

Definition 7. A joint independent random utility model (JIRUM) is a set of preference and aversion probabilities for the subsets of a set A for which there exist densities f_x, $x \in A$ such that the preference and aversion probabilities satisfy Eqs. 26 and 27 for all $x \in X \subsetneq A$.

6.1 Two impossibility theorems.

There is only one result in the literature which concerns joint independent random utility models and this is due to Luce [(1959), Theorem 7, Pp. 57]. We now state this result.

Theorem 12. Let $A = \{x,y,z\}$ and assume that

i. both the preference and aversion probabilities for the subsets of A satisfy the choice axiom,

ii. $p(r,s) \neq 0$ for $r,s \in A$,

iii. $p(x,y) + p(x,z) \neq 1$

and iv. $p^*(x,y) = p(y,x)$,

then the preference and aversion probabilities for the subsets of A do not satisfy a joint independent utility model.

We use techniques similar to those of Luce to prove Theorem 13.

<u>Theorem 13</u>. <u>Let</u> $A = \{x,y,z\}$ <u>and assume that the preference and aversion probabilities for the subsets of</u> A <u>satisfy a concordant choice model with</u> $p(x,y) + p(x,z) \neq 1$. <u>Then the preference and aversion probabilities for the subsets of</u> A <u>do not satisfy a joint independent random utility model</u>.

Proof. Suppose the theorem is false, then by Eqs. 26 and 27 we have for $x \in A$,

$$\begin{aligned} P_A(x)-P^*_A(x) &= \int_{-\infty}^{\infty} f_x(t)\{F_y(t)F_z(t)-[1-F_y(t)][1-F_z(t)]\}dt \\ &= \int_{-\infty}^{\infty} f_x(t)[F_y(t)+F_z(t)-1] \\ &= p(x,y) + p(x,z) - 1 . \end{aligned} \tag{28}$$

The set $\{P_X,P^*_X: X \subseteq A\}$ satisfies a concordant choice model and it follows from Theorem 2 that

$$P_A(x)-P^*_A(x) = \frac{p(x,y)p(x,z)}{p(x,y)p(x,z)+p(y,x)p(y,z)+p(z,x)p(z,y)} - \frac{p^*(x,y)p^*(x,z)}{p^*(x,y)p^*(x,z)+p^*(y,x)p^*(y,z)+p^*(z,x)p^*(z,y)} .$$

Also, Theorem 2 shows that $p(r,s) = p^*(s,r)$ for $r,s \in A$, which is sufficient to prove that the denominators of the above expressions are equal and hence

$$\begin{aligned} P_A(x)-P^*_A(x) &= \frac{p(x,y)p(x,z)-p(y,x)p(z,x)}{p(x,y)p(x,z)+p(y,x)p(y,z)+p(z,x)p(z,y)} \\ &= \frac{p(x,y)p(x,z)-[1-p(x,y)][1-p(x,z)]}{p(x,y)p(x,z)+p(y,x)p(y,z)+p(z,x)p(z,y)} \\ &= \frac{p(x,y)+p(x,z)-1}{p(x,y)p(x,z)+p(y,x)p(y,z)+p(z,x)p(z,y)} . \end{aligned} \tag{29}$$

Comparing Eqs. 28 and 29 and remembering that $p(x,y)+p(x,z) \neq 1$ it follows that

$$p(x,y)p(x,z)+p(y,x)p(y,z)+p(z,x)p(z,y) = 1 .$$

We demonstrate that this is not possible. If $p(x,z) > p(y,z)$, then because $p(r,s) \neq 0$ for $r,s \in A$, we have

$$\begin{aligned} & p(x,y)p(x,z)+p(y,x)p(y,z)+p(z,x)p(z,y) \\ &= p(x,y)p(x,z)+[1-p(x,y)]p(y,z)+p(z,x)p(z,y) \\ &< \max\{p(x,z),p(y,z)\}+p(z,x)p(z,y) \\ &< p(x,z)+p(z,x) \\ &= 1 . \end{aligned}$$

The expression is also strictly less than one if $p(y,z) \geq p(x,z)$ and thus

$$p(x,y)p(x,z)+p(y,x)p(y,z)+p(z,x)p(z,y) < 1$$

for all binary probabilities satisfying the conditions of the theorem.

QED.

6.2 Regularity and joint independent random utility models.

<u>Theorem 14</u>. <u>If the preference and aversion probabilities for the subsets of a set A satisfy a joint random utility model then the preference and aversion probabilities for the subsets of A are regular but not conversely</u>.

Proof. If the preference and aversion probabilities satisfy a joint independent random utility model with densities $f_x(t)$, then for any $x \in X \subsetneq Y \subseteq A$,

$$P_X(x) = \int_{-\infty}^{\infty} f_x(t) \prod_{y \in X-\{x\}} F_y(t)\,dt$$

$$\geq \int_{-\infty}^{\infty} f_x(t) \prod_{y \in Y-\{x\}} F_Y(t)dt$$

$$= P_Y(x)$$

and

$$P^*_X(x) = \int_{-\infty}^{\infty} f_x(t) \prod_{y \in X-\{x\}} [1-F_y(t)]dt$$

$$\geq \int_{-\infty}^{\infty} f_x(t) \prod_{y \in Y-\{x\}} [1-F_y(t)]dt$$

$$= P^*_Y(x) \ .$$

Conversely, let $A = \{x,y,z\}$ and assume that the preference and aversion probabilities for the subsets of A satisfy the conditions of Theorem 12. Then Lemma 2 implies that the preference probabilities are regular and a similar result shows that the aversion probabilities are regular. However, Theorem 12 shows that the preference and aversion probabilities for the subsets of A do not satisfy a joint independent random utility model. QED.

6.3 Joint random utility models and the strong discard and acceptance conditions.

Theorem 14 with Theorem 10 shows that if a set of preference and aversion probabilities satisfy a joint independent random utility model, then the preference probabilities satisfy the discard condition and the aversion probabilities satisfy the acceptance condition. However, if the binary choice probabilities are non-zero and satisfy the constraints of Theorem 13, then Theorem 13 with Theorem 9 implies that no joint independent random utility model exists which implies both the strong discard and the strong acceptance condition; nonetheless, we show that each of the strong conditions is satisfied by some joint independent random utility model.

Theorem 12 shows that under quite general conditions no joint independent random utility model implies that both the preference and aversion probabilities for the subsets of a set A satisfy the choice axiom. This result would be true if no independent random utility model existed such that the preference probabilities satisfied the choice axiom or if no independent random utility model existed such that the aversion probabilities satisfied the choice axiom. However, Luce and Suppes [(1965), Theorem 32, P. 338] show that if a set of non-zero preference probabilities satisfy the choice axiom then they satisfy an independent random utility model and similar techniques may be used to prove that if a set of non-zero aversion probabilities satisfy the choice axiom then they satisfy an independent random utility model. These latter results show that the conditions of the next two theorems can be met non-vacuously.

Theorem 15. *If the preference and aversion probabilities for the subsets of a set* A *satisfy a joint independent random utility model such that the preference probabilities satisfy the choice axiom and the binary choice probabilities are all non-zero, then the preference and aversion probabilities for the subsets of* A *satisfy the strong acceptance condition but not the strong discard condition if* $p(x,y)+p(x,z) \neq 1$ *for some* $x,y,z \in A$.

Proof. When we write $r,s \in Y$ it is to be understood that $r \neq s$. If $Y \subseteq A$ is an n element set, then for any $x \in Y$,

$$P^*_Y(x) = \int_{-\infty}^{\infty} f_x(t) \prod_{y\in Y-\{x\}} [1-F_y(t)]dt$$

$$= 1+(-1)^{n-1}P_Y(x)+(-1)^{n-2}\sum_{z\in X-\{x\}} P_{Y-\{z\}}(x)$$

$$+(-1)^{n-3}\sum_{z,w\in X-\{x\}} P_{Y-\{z,w\}}(x)+\ldots -\sum_{z\in X-\{x\}} p(x,z) \quad . \qquad (30)$$

If $X \subseteq A$ has $n+1$ elements, then using the above expression for the expansion of $P^*_{X-\{y\}}(x)$, $y \in X-\{x\}$, we obtain that

$$\sum_{y\in X-\{x\}} P_X(y) P^*_{X-\{y\}}(x) = \sum_{y\in X-\{x\}} P_X(y) \Big[1+(-1)^{n-1}P_{X-\{y\}}(x)$$

$$+(-1)^{n-2}\sum_{z\in X-\{x,y\}} P_{X-\{z,y\}}(x)$$

$$+(-1)^{n-3}\sum_{z,w\in X-\{x,y\}} P_{X-\{y,z,w\}}(x)\ldots-\sum_{s\in X-\{x,y\}} p(x,s)\Big]$$

$$= [1-P_X(x)]+(-1)^{n-1}\sum_{y\in X-\{x\}} P_X(y)P_{X-\{y\}}(x)$$

$$+(-1)^{n-2}\sum_{y\in X-\{x\}}\ \sum_{z\in X-\{y,x\}} P_X(y)P_{X-\{y,z\}}(x)$$

$$+(-1)^{n-3}\sum_{y\in X-\{x\}}\ \sum_{z,w\in X-\{x,y\}} P_X(y)P_{X-\{y,z,w\}}(x)\ldots$$

$$\cdots \quad -\sum_{y\in X-\{x\}} \; \sum_{s\in X-\{y,x\}} p(x,s)\, P_X(y)$$

$$= [1-P_X(x)]+(-1)^{n-1}\sum_{y\in X-\{x\}} P_X(y)P_{X-\{y\}}(x)$$

$$+(-1)^{n-2}\sum_{r,s\in X-\{x\}} \; \sum_{v\in\{r,s\}} P_X(v)P_{X-\{r,s\}}(x)$$

$$+(-1)^{n-3}\sum_{r,s,t\in X-\{x\}} \; \sum_{v\in\{r,s,t\}} P_X(v)P_{X-\{r,s,t\}}(x) \cdots$$

$$-\sum_{y\in X-\{x\}} \; \sum_{v\in X-\{y,x\}} p(x,y)\, P_X(v)$$

Because the preference probabilities for the subsets of A satisfy the choice axiom with the binary choice probabilities all non-zero, it follows from Lemma 2 that for all sets $Y \subset X$

$$P_X(x) = \left[1 - \sum_{y\in Y} P_X(y)\right] P_{X-Y}(x)$$

i.e.

$$\sum_{y\in Y} P_X(y)P_{X-Y}(x) = P_{X-Y}(x) - P_X(x) \; .$$

Substituting these values in the above equation, we obtain that

$$\sum_{y\in X-\{x\}} P_X(y)P^*_{X-\{y\}}(x) = 1-P_X(x)+(-1)^{n-1}\sum_{y\in X-\{x\}} [P_{X-\{y\}}(x)-P_X(x)]$$

$$+(-1)^{n-2}\sum_{r,s\in X-\{x\}}[P_{X-\{r,s\}}(x)-P_X(x)]$$

$$+(-1)^{n-3}\sum_{r,s,t\in X-\{x\}}[P_{X-\{r,s,t\}}(x)-P_X(x)]\;.\;.\;.$$

$$-\sum_{y\in X-\{x\}}[p(x,y)-P_X(x)]$$

$$= 1+(-1)^{n-1}\sum_{y\in X-\{x\}}P_{X-\{y\}}(x)+(-1)^{n-2}\sum_{r,s\in X-\{x\}}P_{X-\{r,s\}}(x)$$

$$+(-1)^{n-3}\sum_{r,s,t\in X-\{x\}}P_{X-\{r,s,t\}}(x)\;.\;.\;.\;-\sum_{y\in X-\{x\}}p(x,y)\;.$$

$$-P_X(x)[1+\tbinom{n}{1}(-1)^{n-1}+\tbinom{n}{2}(-1)^{n-2}+\;.\;.\;.\;\tbinom{n}{n-1}(-1)]\;.$$

But $0 = (-1+1)^n = (-1)^n+\binom{n}{1}(-1)^{n-1}+\binom{n}{2}(-1)^{n-2}+\ldots\binom{n}{n-1}(-1)+1$,
and so $\binom{n}{1}(-1)^{n-1}+\binom{n}{2}(-1)^{n-2}+\ldots\binom{n}{n-1}(-1)+1 = (-1)^n$.
Substituting this result in the above equation and using Eq. 30 we obtain that

$$\sum_{y\in X-\{x\}}P_X(y)P^*_{X-\{y\}}(x) = 1+(-1)^n P_X(x)+(-1)^{n-1}\sum_{y\in X-\{x\}}P_{X-\{y\}}(y)$$

$$+(-1)^{n-2}\sum_{r,s\in X-\{x\}}P_{X-\{r,s\}}(x)\;.\;.\;.$$

$$-\sum_{y\in X-\{x\}}p(x,y)$$

$$= P^*_X(x)\;,$$

which is the strong acceptance condition.

Suppose now that for some $x,y,z \in A$, $p(x,y)+p(x,z) \neq 1$. By assumption, all the binary preference probabilities are non-zero and we have proved that the preference and aversion probabilities for the subsets of A satisfy the strong acceptance condition. If these probabilities also satisfy the strong discard condition then Theorem 9 shows that they satisfy a concordant choice model. But Theorem 13 shows that such a concordant choice model does not satisfy a joint independent random utility model, contradicting the hypothesis of the present theorem. QED.

Similar techniques may be used to prove the following theorem.

Theorem 16. If the preference and aversion probabilities for the subsets of a set A satisfy a joint independent random utility model such that the aversion probabilities satisfy the choice axiom and the binary choice probabilities are all non-zero then the preference and aversion probabilities satisfy the strong discard condition but not the strong acceptance condition if $p(x,y)+p(x,z) \neq 1$ for some $x,y,z \in A$.

Simple examples may be constructed to show that joint independent random utility models exist which do not satisfy the conditions of either Theorem 16 or Theorem 17.

Although the following result is an immediate consequence of earlier theorems we state it here for completeness.

Theorem 17. A joint independent random utility model neither implies nor is implied by either the strong acceptance or the strong discard condition.

Proof. Theorem 16 demonstrates a joint independent random utility model which does not satisfy the strong discard condition and Theorem 17 demonstrates a joint independent random utility model which does not satisfy the strong acceptance condition.

Now suppose that the preference and aversion probabilities for the subsets of the set $A = \{x,y,z\}$ satisfy the strong discard and the strong acceptance condition and all the binary preference probabilities are non-zero with the additional constraint that $p(x,y)+p(x,z) \neq 1$. Then Theorems 11 and 14 imply that the preference and aversion probabilities do not satisfy a joint independent random utility model and hence neither the strong discard or the strong acceptance condition alone can imply a joint random utility model. QED.

7. Binary Choice Between Two Component Objects

We have studied how choice and ranking probabilities for sets Y with more than two elements might be related to the binary choice probabilities $p(x,y)$, $x,y \in Y$ but we have not studied binary choices themselves in any detail. We next consider binary choices between elements of a set $A = X \times X^*$ where each element of A is called a <u>two component object</u>. Examples of such choices are binary choices of human subjects between bitter-sweet solutions [McLaughlin and Luce (1965)] and between uncertain outcomes when each uncertain outcome consists of two equally probable elementary outcomes [Davidson, Suppes and Siegel (1957)].

McLaughlin and Luce (1965) introduced certain theoretical ideas applicable to such situations which have not yet received the detailed attention that they deserve and in the remainder of this thesis we study their suggestions and others which follow quite naturally from earlier models for binary choice. These earlier models have been extended by Luce and Suppes (1965, pp. 331-367) and for convenience their results are summarized in Appendix A.

We first discuss certain models which impose mathematical structure over the choice set and then consider observable properties which are stated entirely in terms of the binary choice probabilities. We next consider which observable properties are necessary consequences of each model and relate these observable properties to those considered by Luce and Suppes (1965). Finally we discuss a set of conditions which is sufficient for a set of choice probabilities to satisfy one of these models.

We assume throughout this part of the thesis that the component sets X, X^* each contain at least two elements.

8. Models

Definition 8. A set of binary choice probabilities over a set $X \times X^*$ is a weak conjoint utility model provided there exist real valued functions f, g with domains X, X^*, respectively, such that for any $x,y \in X$, $x^*,y^* \in X^*$, $p[(x,x^*),(y,y^*)] \geq 1/2$ if and only if $f(x) + g(x^*) \geq f(y) + g(y^*)$.

A model discussed by Luce and Suppes (1965) is the strong utility model (Appendix A, Def. 2) which requires that there exist a real valued function u with domain $X \times X^*$ and a cumulative distribution Φ such that $\Phi(0) = 1/2$ and, for any $x,y \in X$, $x^*,y^* \in X^*$ with $p[(x,x^*),(y,y^*)] \neq 0,1$,

$$p[(x,x^*),(y,y^*)] = \Phi[u(x,x^*)-u(y,y^*)] . \quad (31)$$

When all the choice probabilities are different from 0 and 1, this implies the weak utility model (Appendix A, Def. 1), namely

$$p[(x,x^*),(y,y^*)] \geq 1/2 \text{ if and only if } u(x,x^*) \geq u(y,y^*) \quad (32)$$

and if the probabilities also satisfy a weak conjoint utility model then there exist real valued functions f, g such that

$$p[(x,x^*),(y,y^*)] \geq 1/2 \text{ if and only if } f(x) + g(x^*) \geq f(y) + g(y^*) . \quad (33)$$

Eqs. 32 and 33 imply [Aczel (1965)] that there exists a strictly monotonic increasing function K such that for any $(x,x^*) \in X \times X^*$,

$$u(x,x^*) = K[f(x) + g(x^*)]$$

and substituting this value in Eq. 31 we obtain that

$$p[(x,x^*),(y,y^*)] = \phi\{K[f(x)+g(x^*)]-K[f(y)+g(y^*)]\}$$

which is a possible stronger version of the strong utility model.

We summarize these ideas as formal definitions and then discuss the interrelations between them, the weak conjoint utility model (Def. 8), the weak utility model (Appendix A, Def. 1) and the binary strict utility model (Appendix A, Def. 3).

Definition 9. A set of binary choice probabilities over a set A is a strong utility model provided there exist a real valued function u over A and a strictly monotonic cumulative distribution ϕ such that $\phi(0) = 1/2$ and for all $a,b \in A$ with $p(a,b) \neq 0,1$,

$$p(a,b) = \phi[u(a)-u(b)] .$$

A strong utility model over a set $X \times X^*$ which is also a weak conjoint utility model over $X \times X^*$ is a strong conjoint utility model.

A strong conjoint utility model over a set $X \times X^*$ is of type K provided

i) K is a strictly monotonic increasing, real valued function

and ii) there exist real valued functions f, g with domains X, X^*, respectively, and a strictly monotonic cumulative distribution ϕ with $\phi(0) = 1/2$ such that for all $x,y \in X$, $x^*,y^* \in X^*$ with $p[(x,x^*),(y,y^*)] \neq 0,1$,

$$p[(x,x^*),(y,y^*)] = \Phi\{K[f(x)+g(x^*)]-K[f(y)+g(y^*)]\} .$$

The customary definition of a strong utility model does not require Φ to be strictly monotonic. The discussion prior to Def. 9 shows that any strong conjoint utility model is of some type and we later study a set of conditions under which the class of admissible transformations of the function K is known.

We next consider certain strict utility models which are also strong conjoint utility models.

Definition 10. A set of binary choice probabilities over a set $X \times X^*$ satisfy a (binary) strict conjoint utility model (of type K) provided

i) K is a strictly monotonic increasing, real valued function

and ii) there exist real valued functions f, g with domains X, X^* respectively, such that for $x,y \in X$, $x^*,y^* \in X^*$ with $p[(x,x^*),(y,y^*)] \neq 0,1$,

$$p[(x,x^*),(y,y^*)] = 1\Big/\Big[1+\exp\{K[f(y)+g(y^*)]-K[f(x)+g(x^*)]\}\Big] .$$

It is immediate from this definition that any binary strict conjoint utility model of type K is a strong conjoint utility model of type K, that there exist strong conjoint utility models which are not strict conjoint utility models and that any binary strict conjoint utility model is a binary strict utility model but not conversely.

The following counterexamples, plus the relevant definitions, determine the relations between the models that we have discussed, these relations being summarized in Fig. 3.

Counterexample 1. A weak conjoint utility model need not be a strong utility model.

A necessary condition for the binary choice probabilities over a set $X \times X^*$ to satisfy a strong utility model is the quadruple condition (Appendix A, Def. 6 and Fig. 4). If $X = \{x,y\}$, $X^* = \{x^*,y^*\}$ is such that $p(a,a) = 1/2$ for all $a \in X \times X^*$ and

$$\begin{aligned} 0 &< p[(x,x^*),(y,x^*)] = p[(x,x^*),(x,y^*)] < p[(x,x^*),(y,y^*)] \\ &< p[(y,x^*),(y,y^*)] = p[(x,y^*),(y,y^*)] < 1/2 \\ &= p[(y,x^*),(x,y^*)] \end{aligned}$$

then the probabilities satisfy a weak conjoint utility model with $f(x) = g(x^*) = 1$, $f(y) = g(y^*) = 2$ but they do not satisfy a strong utility model because $p[(x,x^*),(x,x^*)] = 1/2 > p[(y,x^*),(y,y^*)]$ whereas $p[(x,x^*),(y,x^*)] < p[(x,x^*),(y,y^*)]$.

Counterexample 2. A strong utility model need not be a weak conjoint utility model.

A necessary condition for a set of binary choice probabilities over a set $X \times X^*$ to satisfy a weak conjoint utility model is that if $\min\{p[(x,x^*),(y,y^*)],p[(y,x^*),(x,x^*)]\} \geq 1/2$ then $p[(x,x^*),(x,y^*)] \geq 1/2$. This is violated by the binary strict utility model over $\{x,y\} \times \{x^*,y^*\}$ in which $v(x,x^*) = v(y,y^*) = v(y,x^*) = 2v(x,y^*) \neq 0$ because

$$p[(x,x^*),(y,y^*)] = \frac{v(x,x^*)}{v(x,x^*)+v(y,y^*)} = 1/2 ,$$

$$p[(y,x^*),(x,x^*)] = \frac{v(y,x^*)}{v(y,x^*)+v(x,x^*)} = 1/2$$

whereas

$$p[(x,x^*),(x,y^*)] = \frac{v(x,x^*)}{v(x,x^*)+v(x,y^*)} = 1/3 < 1/2 .$$

But any binary strict utility model is also a strong one (Appendix A, Fig. 3).

Counterexample 3. A strong conjoint utility model need not be a binary strict utility model.

A necessary and sufficient condition for a set of binary choice probabilities over a set $X \times X^*$ to satisfy a strict utility model is that it satisfies the product rule over $X \times X^*$ (Appendix A, Def. 7) and it is easy to construct a set of choice probabilities that satisfy a strong conjoint utility model but not the product rule.

Counterexample 4. A binary strict utility model need not be a strong conjoint utility model.

The set of probabilities of Counterexample 2 shows that a binary strict utility model need not be a weak conjoint utility model, the latter being strictly weaker than a strong conjoint utility model.

8.1 A uniqueness theorem.

We know that any strong conjoint utility model is of some type K and we next give a set of conditions under which the admissible transformations of K are known. In the following theorem, $R(h)$ denotes the range of the function h.

Theorem 18. Suppose that a set of binary choice probabilities over a set $X \times X^*$ are different from 0 and 1 and satisfy strong conjoint utility models of type K_i, $i = 1,2$. Let ϕ_i, f_i and g_i, $i = 1,2$, be the real valued functions such that, for $x,y \in X$, $x^*,y^* \in X^*$ and $i = 1,2$,

$$p[(x,x^*),(y,y^*)] = \phi_i\{K_i[f_i(x)+g_i(x^*)]-K_i[f_i(y)+g_i(y^*)]\} . \quad (34)$$

Then if for $i = 1,2$

i) $R(f_i) = R(g_i) = R(K_i) =$ the real numbers, then there exist constants α, β, γ, ε, and $\delta > 0$ such that for any $(x,x^*) \in X \times X^*$ and real number a,

$$f_2(x) = \alpha f_1(x)+\beta , \quad (35)$$

$$g_2(x^*) = \alpha g_1(x^*)+\gamma , \quad (36)$$

$$\phi_2(\delta a) = \phi_1(a) , \quad (37)$$

and $$K_2[\alpha a+(\beta+\gamma)] = \delta K_1(a)+\varepsilon ; \quad (38)$$

ii) $R(f_i) = R(g_i) =$ the real numbers and $R(K_i) =$ the nonnegative real numbers then Eqs. 35-38 hold for all nonnegative real numbers a with $\varepsilon = 0$,

iii) $R(f_i) = R(g_i) =$ the nonnegative real numbers and $R(K_i) =$ the real numbers then Eqs. 35-38 hold for all real numbers a with $\beta = \gamma = 0$, $\alpha > 0$,

and

iv) $R(f_i) = R(g_i) = R(K_i) =$ the nonnegative real numbers then Eqs. 35-38 hold with $\beta = \gamma = \varepsilon = 0$, $\alpha > 0$.

Proof. Parts ii)-iv) are immediate from part i) and so we prove the latter. Eq. 34 implies that for $x,y \in X$, $x^* \in X^*$,

$$p[(x,x^*),(y,x^*)] \geq 1/2 \text{ is equivalent to } f_1(x) \geq f_1(y)$$
$$\text{and to } f_2(x) \geq f_2(y) .$$

Because the range of f_i , $i = 1,2$ is the real numbers, this pair of equivalences implies [Aczel (1965)] that there exists a continuous, strictly monotonic increasing function F such that for $x \in X$,

$$f_2(x) = Ff_1(x) \quad . \tag{39}$$

Similar arguments prove the existence of continuous, strictly monotonic increasing functions G and H such that for $x \in X$, $x^* \in X^*$,

$$g_2(x^*) = Gg_1(x^*) \tag{40}$$

and

$$f_2(x)+g_2(x^*) = H[f_1(x)+g_1(x^*)] \quad . \tag{41}$$

Using Eqs. 39-41 we obtain that for $a,b \in R(f_1) = R(g_1)$,

$$\begin{aligned} F(a)+G(b) &= H(a+b) \\ &= H(b+a) \\ &= F(b)+G(a) \end{aligned}$$

which for $a = 0 \in R(f_1)$ gives

$$G(b) = F(b)+[G(0)-F(0)] \quad . \tag{42}$$

Eqs. 41 and 42 then imply that

$$\begin{aligned} F(a+b)+G(0) &= H(a+b+0) \\ &= H(a+b) \\ &= F(a)+G(b) \\ &= F(a)+F(b)+[G(0)-F(0)] \end{aligned}$$

and so $$F(a+b) = F(a)+F(b)-F(0) \quad .$$

But F is continuous with domain the real numbers and hence there exists a constant α such that

$$F(a) = \alpha a+F(0)$$

and so for $x \in X$,

$$f_2(x) = F[f_1(x)] = \alpha f_1(x) + \beta ,$$

where $\beta = F(0)$. Substituting the above expression for F in Eq. 42 we obtain that

$$G(b) = F(b)+G(0)-F(0)$$
$$= \alpha a+G(0)$$

and so for $x^* \in X^*$,

$$g_2(x^*) = G[g_1(x^*)] = \alpha g_1(x^*)+\gamma$$

where $\gamma = G(0)$.

These expressions for f_2 , g_2 substituted in Eq. 34 give that for any real numbers a,b

$$\phi_1[K_1(a)-K_1(b)] = \phi_2\{K_2[\alpha a+(\beta+\gamma)]-K_2[\alpha b+(\beta+\gamma)]\}$$

and if we let $\psi = \phi_2^{-1}\phi_1$ then the above equation implies that for any real numbers a, b, c ,

$$\psi[K_1(a)-K_1(b)] = K_2[\alpha a+(\beta+\gamma)]-K_2[\alpha b+(\beta+\gamma)] , \qquad (43)$$

$$\psi[K_1(b)-K_1(c)] = K_2[\alpha b+(\beta+\gamma)]-K_2[\alpha c+(\beta+\gamma)]$$

and

$$\psi[K_1(a)-K_1(c)] = K_2[\alpha a+(\beta+\gamma)]-K_2[\alpha c+(\beta+\gamma)] .$$

But the sum of the right hand sides of the first two equations equals the right hand side of the third one and $R(K_1)$ = the real numbers. Hence for any real numbers d, e ,

$$\psi(d)+\psi(e) = \psi(d+e) .$$

The only bounded solution to this equation is $\psi(d) = \delta d$ for some constant δ and so $\phi_2(\delta d) = \phi_1(d)$. Because ϕ_2 is a cumulative distribution, we require $\delta > 0$. Substituting this value in Eq. 43 for some fixed b_o , we obtain that

$$K_2[\alpha a+(\beta+\gamma)] = \delta K_1(a)+\varepsilon$$

where $$\varepsilon = K_2[\alpha b_o+(\beta+\gamma)]-\delta K_1(b_o)$$ QED.

8.2 Strong conjoint utility models of exponential and identity type.

We next show that under certain interesting conditions only two classes of strong utility models arise, the first being strong conjoint models of exponential type, the other being strong conjoint models of identity type. We state these conditions in the form of an assumption and then discuss this assumption.

Assumption 1. X, X^* are sets and the binary choice probabilities over $X \times X^*$ satisfy a strong utility model such that, for $x,y \in X$, $x^*,y^* \in X^*$ with $p[(x,x^*),(y,y^*)] \neq 0,1$,

$$p[(x,x^*),(y,y^*)] = \Phi\{K[f(x),g(x^*)] - K[f(y),g(y^*)]\}$$

where i) each of the functions f,g is either a ratio scale with range the positive real numbers or an interval scale with range the real numbers,

ii) K is continuous in and dependent upon each of its arguments and has range the real numbers,

and iii) if $\underline{T}_f$, $\underline{T}_g$, $\underline{T}_K$ denote the class of admissible transformations of f, g and K respectively, then for any $x \in X$, $x^* \in X^*$, $T_1 \in \underline{T}_f$, $T_2 \in \underline{T}_g$ there exists $D(T_1,T_2) \in \underline{T}_K$ such that

$$K[T_1f(x),T_2g(x^*)] = D(T_1,T_2)K[f(x),g(x^*)] . \tag{44}$$

Assumption 1 without parts i) and iii) implies that K is an interval scale [Block and Marschak (1960), p. 104] and thus $\underline{T}_K$ of part iii) will be the class of linear transformations.

Part iii) of the assumption is strong. A similar constraint is implicit in dimensional analysis and Luce (1959, 1965) uses it explicitly. Expressing the constraint in words, we are assuming that any admissible transformation of the independent variables f and g effects only admissible transformations of the dependent variable K.

The functions f and g which arise in the above assumption will be called the <u>scale of</u> X and the <u>scale of</u> X^*, respectively, and we say that these scales are <u>independent</u> if we may choose any combination of their values and any combination of their admissible transformations.

<u>Theorem 19</u>. <u>If a set of binary choice probabilities over a set</u> $X \times X^*$ <u>satisfy Assumption</u> 1 <u>with the scales of</u> X <u>and</u> X^* <u>independent, then either the probabilities satisfy a strong conjoint utility model of exponential type or they satisfy a strong conjoint utility model of identity type</u>.

Proof. Luce (1965, Theorem 3) has shown that Eq. 44 can hold under the conditions of Assumption 1 only if f and g are ratio scales, in which case part i) of Assumption 1 requires that they be nonnegative functions. As pointed out above, K is an interval scale and Theorem 2 of Luce (1965) shows that under this set of conditions either there exist $\alpha \neq 0$, $\beta_1\beta_2 \neq 0$ and γ such that for each $(x,x^*) \in X \times X^*$,

$$K[f(x),g(x^*)] = \alpha[f(x)]^{\beta_1}[g(x^*)]^{\beta_2}+\gamma \tag{45}$$

or there exist $\beta_1, \beta_2 \neq 0$ and γ such that

$$K[f(x), g(x^*)] = \log[f(x)]^{\beta_1}[g(x^*)]^{\beta_2} + \gamma . \qquad (46)$$

In either case, define $f_1(x) = \log[f(x)]^{\beta_1}$, $g_1(x^*) = \log[g(x^*)]^{\beta_1}$ and $\phi_1(a) = \phi(\alpha a)$. Then if Eq. 45 holds we obtain that

$$p[(x,x^*),(y,y^*)] = \phi\{\alpha[f(x)]^{\beta_1}[g(x^*)]^{\beta_2} - \alpha[f(y)]^{\beta_1}[g(y^*)]^{\beta_2}\}$$

$$= \phi_1\{\exp[f_1(x)+g_1(x^*)] - \exp[f_1(y)+g_1(y^*)]\} ,$$

which is a strong conjoint utility model of exponential type, and if Eq. 46 holds we obtain that

$$p[(x,x^*),(y,y^*)] = \phi\{\log[f(x)]^{\beta_1}[g(x^*)]^{\beta_2} - \log[f(y)]^{\beta_1}[g(y^*)]^{\beta_2}\}$$

$$= \phi\{[f_1(x)+g_1(x^*)] - [f_1(y)+g_1(y^*)]\} ,$$

which is a strong conjoint utility model of identity type.

QED.

If $X \equiv X^*$ then clearly any admissible transformation which is applied to the scale of X must also be applied to the scale of X^* and the scales of X, X^* are not independent in the sense defined above. Luce (1965) did not study this case, but in Appendix B we use techniques similar to his to obtain the class of possible functional relations when the scales are restricted in the above manner. This class is quite large and in this paper we consider only one strong conjoint utility model of type K with K neither the exponential nor identity function.

Theorem 18 shows that under quite general conditions the type of a model is unique up to the class of transformations given in Eq. 38 and, in particular, under these conditions a strong conjoint utility model cannot be both of exponential type and of identity type. The following theorem proves this result without any assumptions concerning the range of f, g or K.

Theorem 20. A set of binary choice probabilities over a set $X \times X^*$ satisfies a strong conjoint utility model of exponential type and a strong conjoint utility model of identity type only if the scale of X or the scale of X^* is a constant function in both models.

Proof. If the probabilities satisfy a strong conjoint utility model of exponential type and one of identity type then there exist real valued functions f_i, g_i and strictly monotonic distributions ϕ_i, $i = 1,2$ such that for $x,y \in X$, $x^*,y^* \in X^*$ with $p[(x,x^*),(y,y^*)] \neq 0,1$,

$$p[(x,x^*),(y,y^*)] = \phi_1[f_1(x)g_1(x^*)-f_1(y)g_1(y^*)] \tag{47}$$

$$= \phi_2\{[f_2(x)+g_2(x^*)] -[f_2(y)+g_2(y^*)]\} . \tag{48}$$

Eq. 48 implies that

$$p[(x,x^*),(y,x^*)] = p[(x,y^*),(y,y^*)]$$

which with Eq. 47 and the strict monotonicity of ϕ_1 implies that

$$f_1(x)g_1(x^*)-f_1(y)g_1(x^*) = f_1(x)g_1(y^*)-f_1(y)g_1(y^*)$$

and so

$$[f_1(x)-f_1(y)][g_1(x^*)-g_1(y^*)] = 0 .$$

Because $x,y \in X$, $x^*,y^* \in X^*$ are arbitrary elements this result implies that either f_1 or g_1 is a constant function and then Aczel's (1965) result applied as in Theorem

18 implies that either f_2 or g_2, respectively, is a constant function. QED.

9. Observable Properties

Each conjoint utility model of the previous section involves at least two unknown functions which we do not expect to calculate from data. In this chapter we study a series of conditions which are stated entirely in terms of the observable choice probabilities and in the next chapter we show which observable properties are necessary consequences of each model.

We first generalize to probabilistic situations the algebraic "cancellation condition" discussed by Luce and Tukey (1965). This generalization is similar in nature to the well known generalizations of algebraic transitivity (Appendix A, Def. 5).

Definition 11. *Whenever* $\min\{p[(x,x^*),(y,y^*)],p[(y,z^*),(z,x^*)]\} \geq 1/2$ *the binary preference probabilities satisfy*

i. weak stochastic cancellation *provided that* $p[(x,z^*),(z,y^*)] \geq 1/2$,

ii. moderate stochastic cancellation *provided that* $p[(x,z^*),(z,y^*)] \geq \min\{p[(x,x^*),(y,y^*)],p[(y,z^*),(z,x^*)]\}$

iii. strong stochastic cancellation *provided that* $p[(x,z^*),(z,y^*)] \geq \max\{p[(x,x^*),(y,y^*)],p[(y,z^*),(z,x^*)]\}$.

These conditions were introduced by McLaughlin and Luce (1965) who studied whether they were satisfied by the binary choices of human subjects between certain bitter-sweet solutions. Their data provides almost complete support for weak cancellation, considerable support for moderate cancellation and throws considerable doubt on strong cancellation.

<u>Theorem 21</u>. <u>Strong stochastic cancellation is strictly stronger than moderate stochastic cancellation which in turn is strictly stronger than weak stochastic cancellation</u>.

Proof. For any set of binary choice probabilities p over a set $X \times X^*$ we introduce a new set of probabilities q via the definition: for each $x,y \in X$, $x^*,y^* \in X^*$,

$$q[(x,x^*),(y,y^*)] = p[(x,y^*),(y,x^*)] .$$

Inspection of the definitions of stochastic transitivity and stochastic cancellation shows that stochastic cancellation of p of a certain strength is equivalent to stochastic transitivity of q of the same strength. But strong stochastic transitivity is strictly stronger than moderate stochastic transitivity, which in turn is strictly stronger than weak stochastic transitivity (Appendix A, Fig. 4). QED.

The following is a kind of independence condition on the components of the objects under study.

<u>Definition 12</u>. <u>A set of binary choice probabilities over a set</u> $X \times X^*$ <u>satisfies the</u> transposition condition <u>if for</u> $a,b \in X \times X^*$, $x,y \in X$ <u>and</u> $x^*,y^* \in X^*$, $p[(x,x^*),a] \geq p[b,(y,y^*)]$ <u>is equivalent to</u> $p[(x,y^*),a] \geq p[b,(y,x^*)]$ <u>and to</u> $p[(y,x^*),a] \geq p[b,(x,y^*)]$.

We need the following lemma to prove the next theorem.

<u>Lemma 6</u>. <u>If a set of binary choice probabilities over a set</u> $X \times X^*$ <u>satisfies the transposition condition then for</u> $c,d \in X \times X^*$, $x,y \in X$ <u>and</u> $x^*,y^* \in X^*$, $p[c,(x,x^*)] \geq p[(y,y^*),d]$ <u>is equivalent to</u> $p[c,(x,y^*)] \geq p[(y,x^*),d]$ <u>and to</u> $p[c,(y,x^*)] \geq p[(x,y^*),d]$.

Proof. Obvious since $p[c,(x,x^*)] = 1-p[(x,x^*),c]$ QED.

We will occasionally use the symbol '$\leftrightarrow$' to denote the phrase 'if and only if.'

<u>Theorem 22</u>. <u>The transposition condition implies strong stochastic cancellation but not conversely</u>.

Proof. Suppose that the transposition condition is satisfied and $\min\{p[(x,x^*),(y,y^*)],p[(y,z^*),(z,x^*)]\} \geq 1/2$. Then the transposition condition and Lemma 6 imply that

$$
\begin{aligned}
& p[(x,x^*),(y,y^*)] \geq 1/2 = p[(z,z^*),(z,z^*)] \\
\leftrightarrow\ & p[(x,z^*),(y,y^*)] \geq p[(z,z^*),(z,x^*)] \\
\leftrightarrow\ & p[(x,z^*),(z,y^*)] \geq p[(y,z^*),(z,x^*)] \qquad (49)
\end{aligned}
$$

and

$$
\begin{aligned}
& p[(y,z^*),(z,x^*)] \geq 1/2 = p[(x,y^*),(x,y^*)] \\
\leftrightarrow\ & p[(y,y^*),(z,x^*)] \geq p[(x,y^*),(x,z^*)] \\
\leftrightarrow\ & p[(y,y^*),(x,x^*)] \geq p[(z,y^*),(x,z^*)]
\end{aligned}
$$

which is equivalent to

$$p[(x,z^*),(z,y^*)] \geq p[(x,x^*),(y,y^*)] . \qquad (50)$$

Combining Eqs. 49 and 50 we obtain that

$$p[(x,z^*),(z,y^*)] \geq \max\{p[(x,x^*),(y,y^*)], p[(y,z^*),(z,x^*)]\}$$

which is strong stochastic cancellation.

To show that the converse is false consider the set of probabilities given by Fig. 2. These probabilities satisfy strong stochastic cancellation but not the transposition because $p[(x,x^*),(y,y^*)] = 3/4 > 1/4 = p[(y,x^*),(x,y^*)]$ but $p[(x,y^*),(y,y^*)] = 3/4 = p[(y,x^*),(x,x^*)]$.

10. Relations Between the Models and the Observable Properties

We now consider which of the observable properties of Chapter 9 are necessary consequences of the strict and strong conjoint utility models of exponential, logarithmic and identity type. We need not consider every pair consisting of an observable property and a model because certain results are direct consequences of others. For example, we prove that the strict utility model of exponential type does not imply moderate stochastic cancellation and it is immediate from this result and the comment after Def. 10 that the strong conjoint utility model of exponential type does not imply moderate stochastic cancellation. We therefore prove a sequence of results which are sufficient to derive the relation of each model to each observable property, these relations being summarized in Fig. 3.

The study of the strict and strong conjoint utility models is clearly incomplete since we only consider models of the exponential, logarithmic and identity type. We still need a classification of other strict and strong conjoint utility models in terms of which observable properties of Chapter 9 are necessary consequences of each.

We assume throughout this chapter that all the probabilities are different from 0 and 1.

Theorem 23. Any weak conjoint utility model satisfies weak stochastic cancellation but not conversely.

Proof. We have sets X, X^* such that for $\alpha,\beta \in X$, $\alpha^*,\beta^* \in X^*$,

$$p[(\alpha,\alpha^*),(\beta,\beta^*)] \geq 1/2 \text{ if and only if } f(\alpha)+g(\alpha^*) \geq f(\beta)+g(\beta^*)$$

and hence if $\min\{p[(x,x^*),(y,y^*)],p[(y,z^*),(z,x^*)]\} \geq 1/2$

then $\quad f(x)+g(x^*) \geq f(y)+g(y^*)$

and $\quad f(y)+g(z^*) \geq f(z)+g(x^*)$

which implies that

$$f(x)+g(z^*) \geq f(z)+g(y^*)$$

and so $\quad p[(x,z^*),(z,y^*)] \geq 1/2$.

The set of probabilities given in Fig. 2 satisfies weak stochastic cancellation but because

$$p[(x,x^*),(y,x^*)] = 1/4 < 1/2 < \min\{p[(x,x^*),(y,y^*)],p[(y,y^*),(y,x^*)]\}$$

it does not satisfy weak transitivity, which is a necessary condition for a weak conjoint utility model. QED.

<u>Theorem 24</u>. <u>Any strong conjoint utility model of identity type satisfies the transposition condition</u>.

Proof. Using the notation of Def. 9, for $a = (\alpha,\alpha^*) \in X \times X^*$ define $h(a) = f(\alpha)+g(\alpha^*)$. Then for $a,b,(x,x^*),(y,y^*) \in X \times X^*$,

$$p[(x,x^*),a] \geq p[b,(y,y^*)]$$

$$\leftrightarrow \phi\{[f(x)+g(x^*)]-h(a)\} \geq \phi\{h(b)-[f(y)+g(y^*)]\}$$

$$\leftrightarrow \quad [f(x)+g(x^*)]-h(a) \geq h(b)-[f(y)+g(y^*)]$$

$$\leftrightarrow \quad [f(x)+g(y^*)]-h(a) \geq h(b)-[f(y)+g(x^*)]$$

$$\leftrightarrow \phi\{[f(x)+g(y^*)]-h(a)\} \geq \phi\{h(b)-[f(y)+g(x^*)]\}$$

$$\leftrightarrow p[(x,y^*),a] \geq p[b,(y,x^*)]$$

which is the first part of the transposition condition. The second part is proved similarly. QED.

It is not known whether the converse of Theorem 24 is true but in Chapter 12 we show that the transposition condition plus an existence condition is sufficient for a set of binary choice probabilities to satisfy a strong conjoint utility model of identity type.

Theorem 25. Strong stochastic cancellation does not imply the weak conjoint utility model.

Proof. The set of probabilities given in Fig. 2 satisfies strong stochastic cancellation but because $p[(x,x^*),(y,x^*)] = 1/4 < 1/2 < \min\{p[(x,x^*),(y,y^*)],p[(y,y^*),(y,x^*)]\}$ it does not satisfy weak stochastic transitivity, which is a necessary condition for a weak conjoint utility model.

QED.

Theorem 26. Any strong conjoint utility model of logarithmic type satisfies moderate but not strong stochastic cancellation and strong stochastic cancellation does not imply the strong conjoint utility model of logarithmic type.

Proof. We may without loss of generality suppose that X, X^* are subsets of the positive real numbers and that for $x,y \in X$, $x^*,y^* \in X^*$,

$$p[(x,x^*),(y,y^*)] = \phi[\log(x+x^*)-\log(y+y^*)]$$

$$= \phi\left[\log\left(\frac{x+x^*}{y+y^*}\right)\right] .$$

If $\min\{p[(x,x^*),(y,y^*)],p[(y,z^*),(z,x^*)]\} \geq 1/2$ then

$$\frac{x+x^*}{y+y^*} \geq 1 \quad , \quad \frac{y+z^*}{z+x^*} \geq 1 \qquad (51)$$

and without loss of generality suppose that

$$\frac{x+x^*}{y+y^*} \geq \frac{y+z^*}{z+x^*} . \qquad (52)$$

Then,

$$\begin{aligned}(x+z^*)(z+x^*) &= (z+x^*)[(x+x^*)-(x^*-z^*)] \\ &= (z+x^*)(x+x^*)-(z+x^*)(x^*-z^*) \\ &\geqq (y+z^*)(y+y^*)-(z+x^*)(x^*-z^*) \qquad \text{(by Eq. 52)} \\ &= (y+z^*)[(z+y^*)-(z-y)]-(z+x^*)(x^*-z^*) \\ &= (y+z^*)(z+y^*)-[(y+z^*)(z-y)+(z+x^*)(x^*-z^*)] \ . \end{aligned} \tag{53}$$

But Eq. 51 implies that

$$\begin{aligned}(y+z^*)(y-z)+(z+x^*)(z^*-x^*) &\geqq (z+x^*)(y-z)+(z+x^*)(z^*-x^*) \\ &= (z+x^*)[(y-z)+(z^*-x^*)] \\ &= (z+x^*)[(y+z^*)-(z+x^*)] \\ &\geqq 0 \ .\end{aligned}$$

Substituting this result in Eq. 53 we obtain that

$$(x+z^*)(z+x^*) \geqq (y+z^*)(z+y^*)$$

and so

$$\frac{x+z^*}{z+y^*} \geqq \frac{y+z^*}{z+x^*}$$

which with Eq. 52 implies that

$$p[(x,z^*),(z,y^*)] \geqq \min\{p[(x,x^*),(y,y^*)], \\ p[(y,z^*),(z,x^*)]\}$$

which is moderate stochastic cancellation.

However, if $x = y^* = 3$, $y = x^* = 1$, then

$$p[(x,x^*),(y,y^*)] = \phi[\log(4/4)] = \phi(0)$$

$$p[(y,y^*),(y,x^*)] = \phi[\log(4/2)] = \phi(\log 2)$$

whereas

$$p[(x,y^*),(y,y^*)] = \phi[\log(6/4)] < \phi(\log 2)$$

which contradicts strong stochastic cancellation.

If strong stochastic cancellation implied the strong conjoint utility model of logarithmic type, then it would follow from earlier results and those summarized in Appendix A (see Fig. 3) that strong stochastic cancellation implied the weak conjoint utility model, which contradicts Theorem 25. QED.

Theorem 27. A strict conjoint utility model of logarithmic type neither implies nor is implied by strong stochastic cancellation.

Proof. If we choose $\phi(a) = 1/(1+e^{-a})$ in the counterexample of Theorem 26 then that counterexample shows that the strict conjoint utility model of logarithmic type does not imply strong stochastic cancellation and an argument similar to that of the last paragraph of Theorem 26 shows that strong stochastic cancellation does not imply the strict conjoint utility model of logarithmic type. QED.

Theorem 28. Any binary strict conjoint utility model of exponential type satisfies weak but not moderate stochastic cancellation and moderate stochastic cancellation does not imply the strict conjoint utility model of exponential type.

Proof. It is deducible from earlier theorems (see Fig. 3) that the binary strict conjoint utility model of exponential type satisfies weak stochastic cancellation. Now suppose that X is real numbers not less than unity, X^* the positive reals and for $x,y \in X$, $\alpha,\beta \in X^*$

$$p[(x,\alpha),(y,\beta)] = \frac{x^\alpha}{x^\alpha+y^\beta} .$$

This set of probabilities satisfies a strict conjoint utility model of exponential type such that $p[(x,\alpha),(y,\beta)] \geq p[(z,\gamma),(w,\delta)]$ if and only if $x^{\alpha}/y^{\beta} \geq z^{\gamma}/w^{\delta}$. Thus if $x = y = \alpha = 2$, $z = 1/2$, $\beta = 1/4$, $\gamma = 1$, then

$$\frac{x^{\alpha}}{y^{\beta}} = 2^{7/4} > 1 \text{ , hence } p[(x,\alpha),(y,\beta)] > 1/2 ,$$

$$\frac{y^{\gamma}}{z^{\beta}} = 8 > 1 \text{ , hence } p[(y,\gamma),(z,\alpha)] > 1/2$$

whereas

$$\frac{x^{\gamma}}{z^{\beta}} = 2^{5/4} < \min\left(2^{7/4},2^{3}\right) = \min\left(\frac{x^{\alpha}}{y^{\beta}}, \frac{y^{\gamma}}{z^{\alpha}}\right)$$

and so

$$p[(x,\gamma),(z,\beta)] < \min\{p[(x,\alpha),(y,\beta)],p[(y,\gamma)(z,\alpha)]\}$$

which contradicts moderate stochastic cancellation.

An argument similar to that of the last paragraph of Theorem 26 shows that moderate stochastic cancellation does not imply the strict conjoint utility model of exponential type. QED.

Because any strict conjoint utility model of type K is a strong conjoint utility model of type K , Theorems 26 and 28 prove that a strong conjoint utility model of exponential type is not necessarily also of logarithmic type; Theorems 24 and 26 prove that a strong conjoint utility model of logarithmic type is not necessarily also of identity type; and Theorem 20 proves that there exist strong conjoint utility models which are of exponential but not of identity type and vice versa. The questions remain whether there exist strong conjoint utility models of identity type which are also of logarithmic type and whether there exist strong conjoint utility models of logarithmic type which are also of exponential type. If the conditions of Theorem 18 are satisfied then these questions can be answered in the negative but we do not know of much weaker conditions for which this is the case.

11. Relations To the Results of Luce and Suppes

If a set of binary preference probabilities over a set $X \times X^*$ satisfy weak stochastic cancellation, then in particular if $x,y,z \in X$ and $w^* \in X^*$ are such that $\min\{p[(x,w^*),(y,w^*)],p[(y,w^*),(z,w^*)]\} \geq 1/2$, then $p[(x,w^*),(z,w^*)] \geq 1/2$, i.e. if $a,b,c \in X \times \{w^*\}$ are such that $\min\{p(a,b),p(b,c)\} \geq 1/2$ then $p(a,c) \geq 1/2$, which is weak stochastic transitivity on $X \times \{w^*\}$. Similar arguments show that weak, moderate and strong stochastic cancellation on $X \times X^*$ imply weak, moderate and strong stochastic transitivity, respectively, on $\{w\} \times X^*$, $X \times \{w^*\}$ for each $(w,w^*) \in X \times X^*$. However, we next prove that, provided X, X^* each contain at least two elements the converse of each of the above statements is false. We then discuss the relations between the observable properties introduced in this part of the thesis and those of Luce and Suppes (see Appendix A) when all the properties hold over a set $X \times X^*$ where X, X^* each contain at least two elements.

<u>Theorem 29</u>. <u>Weak stochastic cancellation neither implies nor is implied by the product rule</u>.

Proof. The product rule is equivalent to the binary strict utility model (Appendix A, Fig. 4) and it is immediate from earlier results (see Fig. 3) that weak stochastic cancellation does not imply the binary strict utility model.

Counterexample 2 (p. 53) exhibits a set of choice probabilities that satisfies a binary strict utility model but does not satisfy weak stochastic cancellation. QED.

Because the product condition is stronger than weak, moderate, and strong stochastic transitivity (Appendix A,

Fig. 4) and weak stochastic cancellation is weaker than moderate and strong stochastic cancellation, Theorem 29 is sufficient to show that the converse of each of the implications mentioned at the beginning of this chapter is false.

Theorem 30. The transposition condition implies the quadruple condition but not the product rule and the product rule does not imply the transposition condition.

Proof. If the transposition condition holds and

$$p[r,(z,z^*)] \geq p[(w,w^*),s]$$

then with this expression as the first inequality of Lemma 6, the second inequality becomes

$$p[r,(z,w^*)] \geq p[(w,z^*),s]$$

and with this expression as the first inequality of Lemma 6, the third inequality becomes

$$p[r,(w,w^*)] \geq p[(z,z^*),s] .$$

The equivalence of this inequality and the initial inequality shows that the quadruple condition holds.

Theorem 24 shows that any strong conjoint utility model of identity type in which the choice probabilities are different from 0 and 1 satisfies the transposition condition but examples can easily be constructed which show that such a model need not satisfy the product rule and hence the transposition condition does not imply the product rule.

If the product rule implied the transposition condition then it would follow from Theorems 21 and 22 that the product rule implies weak stochastic cancellation which contradicts Theorem 29. QED.

<u>Theorem 31</u>. <u>Strong stochastic cancellation neither implies nor is implied by either weak stochastic transitivity or the triangle condition</u>.

Proof. We have already pointed out in Theorem 25 that the set of probabilities given in Fig. 2 satisfy strong stochastic cancellation but not weak stochastic transitivity. It also does not satisfy the triangle condition because

$$p[(x,x^*),(y,x^*)]+p[(y,x^*),(x,y^*)] = 1/2 < 3/4 = p[(x,x^*),(x,y^*)] .$$

Conversely, if the choice probabilities satisfy the product rule then they satisfy both weak stochastic transitivity and the triangle condition (Appendix A, Fig. 4) whereas Theorem 29 shows that they need not satisfy weak stochastic cancellation, which is strictly weaker than strong stochastic cancellation.

<u>Theorem 32</u>. <u>The weak conjoint utility model neither implies nor is implied by the triangle condition</u>.

Proof. If $\{x,y\} \times \{x^*,y^*\}$ is such that $\max\{p[(x,x^*),(y,x^*)],p[(y,x^*),(y,y^*)]\} < 1/2 = p[(y,x^*),(x,y^*)]$, $p[(x,x^*),(x,y^*)] = p[(x,y^*),(y,y^*)] = 1/8$ and $p[(x,x^*),(y,y^*)] = 3/8$ then the weak conjoint utility model holds with $f(x) = g(x^*) = 1$, $f(y) = g(y^*) = 2$ but the triangle condition does not because

$$p[(y,y^*),(x,y^*)] = 7/8 > 6/8 = p[(y,y^*),(x,x^*)] + p[(x,x^*),(x,y^*)] .$$

If the triangle condition implied the weak conjoint utility model then the triangle condition would imply the weak utility model, which contradicts the results of Luce and Suppes (Appendix A, Fig. 4). QED.

12. Representation Theorems

In this section we study certain conditions which are sufficient for a set of binary preference probabilities to satisfy a strong conjoint utility model of identity type and also other conditions that imply the existence of functions whose properties are very similar to those of the functions which occur in the weak conjoint utility model. We first state two lemmas, the first due to Debreu (1958) and the second due to Aczel (1965).

Lemma 7. Suppose that a set of binary choice probabilities over a set A is such that

i) it satisfies the quadruple condition

and ii) for any $a,b,c \in A$ and $q \in [0,1]$ with $p(b,a) \leq q \leq p(c,a)$ there exists $e \in A$ such that $p(e,a) = q$,

then there exists a real valued function u with domain A such that

$$p(a,b) \geq p(c,d) \text{ if and only if } u(a)-u(b) \geq u(c)-u(d) .$$

Lemma 8. For two real valued functions g, h with domains S, T, respectively, and for a real valued function f with domain $S \times T$ the fact that for any $x,y \in S$, $p,q \in T$,

$$f(x,p) \geq f(y,q) \text{ if and only if } g(x)+h(p) \geq g(y)+h(q)$$

implies the existence of a strictly monotonic map k such that

$$f(x,p) = k[g(x)+h(p)] .$$

Theorem 33. Suppose that a set of binary choice probabilities over a set $X \times X^*$ is such that

i) <u>it satisfies the transposition condition; and</u>

ii) <u>for any</u> $a,b,c \in X \times X^*$ <u>and</u> $q \in [0,1]$ <u>with</u> $p(b,a) \le q \le p(c,a)$ <u>there exists</u> $e \in X \times X^*$ <u>such that</u> $p(e,a) = q$,

<u>then the probabilities satisfy a strong conjoint utility model of identity type</u>.

Proof. Theorem 30 shows that the transposition condition implies the quadruple condition and thus the conditions of Lemma 7 are satisfied, which implies that there exists a real valued function u with domain $X \times X^*$ such that for any $a,b,c,d \in X \times X^*$

$$p(a,b) \ge p(c,d) \text{ if and only if } u(a)-u(b) \ge u(c)-u(b) .$$

Applying Lemma 8 to this equivalence we obtain that there exists a strictly monotonic increasing function ϕ such that for $a,b \in X \times X^*$,

$$p(a,b) = \phi[u(a)-u(b)] . \tag{54}$$

If $(x_o,x_o^*) \in X \times X^*$ is a fixed element then Eq. 54 implies that for any $(x,x^*) \in X \times X^*$,

$$p[(x,x^*),(x,x^*)] = p[(x_o,x_o^*),(x_o,x_o^*)]$$

which with the transposition condition implies that

$$p[(x,x_o^*),(x,x^*)] = p[(x_o,x_o^*),(x_o,x^*)] .$$

From the above result and the representation given by Eq. 54 we obtain that

$$u(x,x_o^*)-u(x,x^*) = u(x_o,x_o^*)-u(x_o,x^*)$$

and hence

$$u(x,x^*) = u(x,x_o^*)+u(x_o,x^*)-u(x_o,x_o^*) . \tag{55}$$

If for any $z \in X$, $z^* \in X^*$ we define $f(z) = u(z,x_o^*)$ and $g(z^*) = u(x_o,z^*)$, then substituting Eq. 55 in Eq. 54 we obtain that for $x,y \in X$, $x^*,y^* \in X^*$,

$$\begin{aligned} p[(x,x^*),(y,y^*)] &= \phi[u(x,x^*)-u(y,y^*)] \\ &= \phi\{[u(x,x_o^*)+u(x_o,x^*)-u(x_o,x_o^*)] - \\ &\qquad u(y,x_o^*)+u(x_o,y^*)-u(x_o,x_o^*)]\} \\ &= \phi\{[f(x)+g(x^*)]-[f(y)+g(y^*)]\} , \end{aligned}$$

which is a strong conjoint utility model of identity type.

QED.

I have not found sufficient conditions for a set of probabilities to satisfy strong conjoint utility models of other types.

Theorem 34. If a set of binary choice probabilities over a set $X \times X^*$ satisfy strong stochastic transitivity and weak stochastic cancellation and for any $a \in X \times X^*$, $x \in X$, $x^* \in X^*$ there exists $z \in X$, $z^* \in X^*$ such that $p[a,(z,x^*)] = 1/2 = p[a,(x,z^*)]$ then there exist real valued functions u, f, g with domains $X \times X^*$, X, X^* respectively such that for $x,y \in X$, $x^*,y^* \in X^*$,

i. $p[(x,x^*),(y,y^*)] \geq 1/2$ if and only if $u(x,x^*) \geq u(y,y^*)$

ii. $p[(x,x^*),(y,x^*)] \geq 1/2$ if and only if $f(x) \geq f(y)$

and iii. $p[(x,x^*),(x,y^*)] \geq 1/2$ if and only if $g(x^*) \geq g(y^*)$.

Proof. i. This part of the proof is due to Luce (1964). For a fixed $\alpha \in X \times X^*$ let $u(a) = p(a,\alpha)$ for each $a \in X \times X^*$. If $p(a,b) \geq 1/2$ then there are three cases to consider.

1. $p(b,\alpha) \geq 1/2$, then by strong stochastic transitivity $u(a) = p(a,\alpha \geq p(b,\alpha) = u(b)$.

2. $p(\alpha,a) \geq 1/2$, then by strong stochastic transitivity $p(\alpha,b) \geq p(\alpha,a)$ which implies that $u(a) = p(a,\alpha) \geq p(b,\alpha) = u(b)$.
3. Neither 1 nor 2 holds i.e. $u(a) = p(a,\alpha) > 1/2 > p(b,\alpha) = u(b)$.

Thus the weak utility model holds.

ii. Now suppose that $x,y \in X$, $z^* \in X^*$ are such that $p[(x,z^*),(y,z^*)] \geq 1/2$. We show that $p[(x,w^*),(y,w^*)] \geq 1/2$ for all $w^* \in X^*$. Given $w^* \in X^*$, $w^* \neq z^*$ choose $w \in X$ such that $p[(w,w^*),(x,z^*)] = 1/2$. Then by weak transitivity

$$p[(w,w^*),(y,z^*)] \geq 1/2$$

which with

$$p[(x,z^*),(w,w^*)] = 1/2$$

implies by weak stochastic cancellation that $p[(x,w^*),(y,w^*)] \geq 1/2$, which is the desired result.

Let $x_o^* \in X^*$ be fixed and for each $x \in X$ define $f(x) = u(x,x_o^*)$. The the above discussion shows that for $x,y \in X$, $x^* \in X^*$,

$$p[(x,x^*),(y,x^*)] \geq 1/2 \text{ is equivalent to } p[(x,x_o^*),(y,x_o^*)] \geq 1/2 ,$$

which by part i. holds if and only if $f(x) = u(x,x_o^*) \geq u(y,x_o^*) = f(y)$ i.e. $p[(x,x^*),(y,x^*)] \geq 1/2$ if and only if $f(x) \geq f(y)$.

Similar arguments may be used to prove part iii. QED.

Although the above theorem does not give sufficient conditions for a weak conjoint utility model the representation obtained is very similar to that required of such a model.

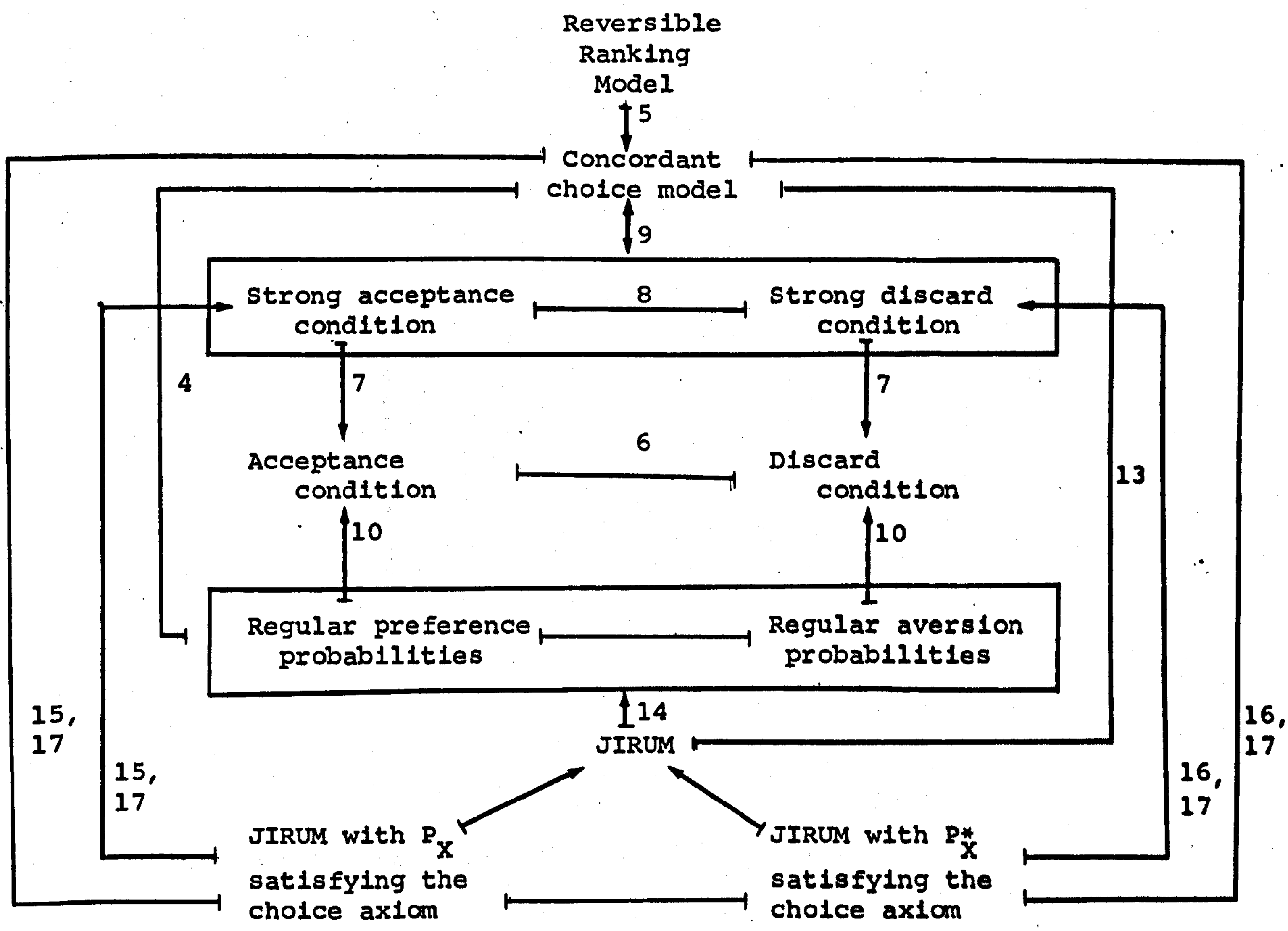


Fig. 1. A summary of the implications ($\rightarrow$) and failures of implications ($\rightarrow\!\!|$) among models and observable probabilities on the assumption that $p(x,y) = p^*(y,x) \neq 0$ for any elements x, y being considered. The number beside a line indicates the relevant theorem and those lines with no number are results which are immediate from definitions. JIRUM denotes a joint independent random utility model.

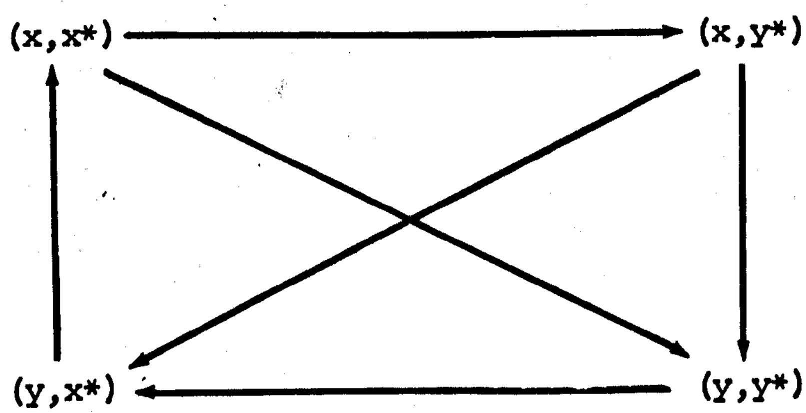
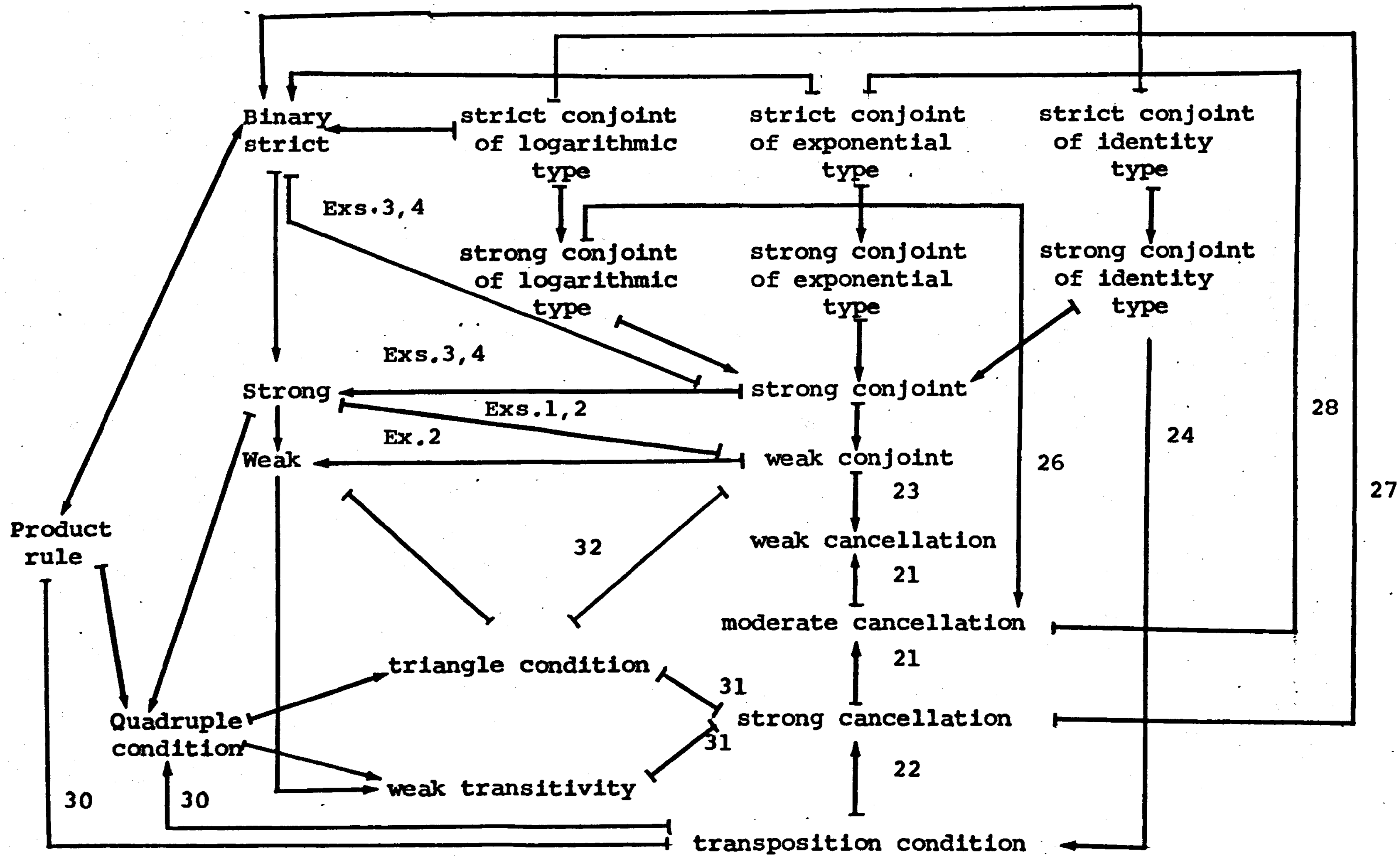


Figure 2. $(\alpha,\alpha^*) \rightarrow (\beta,\beta^*)$ indicates that

$p[(\alpha,\alpha^*),(\beta,\beta^*)] = 3/4$.



Binary strict
strict conjoint of logarithmic type
strict conjoint of exponential type
strict conjoint of identity type
strong conjoint of logarithmic type
strong conjoint of exponential type
strong conjoint of identity type
Exs.3,4
Exs.3,4
Strong
strong conjoint
Exs.1,2
Ex.2
Weak
weak conjoint
23
weak cancellation
21
moderate cancellation
21
strong cancellation
22
transposition condition
24
26
28
27
32
Product rule
triangle condition
31
31
Quadruple condition
weak transitivity
30
30

Figure 3

Figure 3. A summary of the implications ($\rightarrow$) and failures of implications ($\dashv$) among models and observable properties for binary choices between elements of a set $X \times X^*$ on the assumptions that X, X^* each contain at least two elements and that the choice probabilities are different from 0 and 1. The numbers beside a line indicate the relevant theorems or counterexamples.

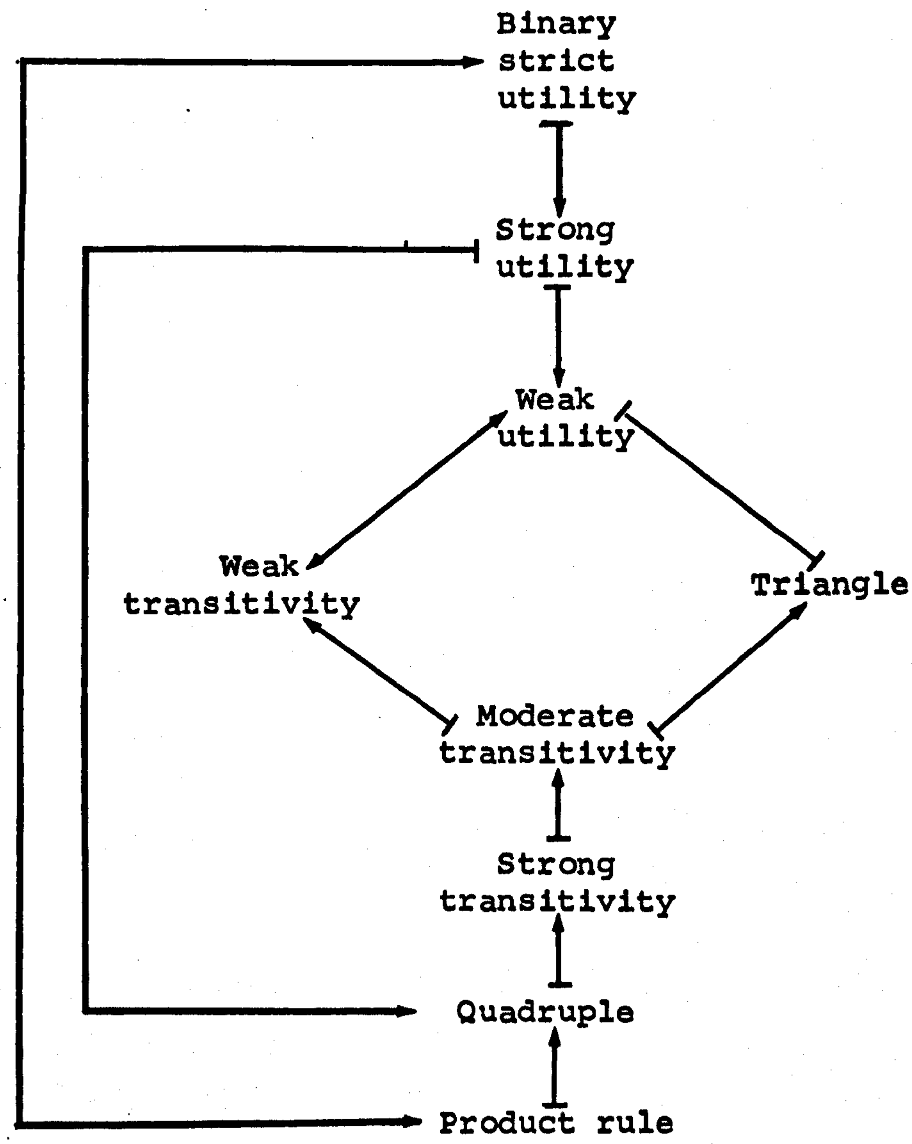


Fig. 4. A summary of the implications (→) and failures of implications (⊣) among models and observable properties for binary choices between elements of a finite set A on the assumption that the choice probabilities are different from 0 and 1. The relation between any two concepts can be deduced using only the transitivity of implication.

Appendix A

The following definitions and Fig. 4 are taken from Luce and Suppes [1965, pp. 333-344].

Definition 1. A weak (binary) utility model is a set of binary preference probabilities for which there exists a real-valued function w over A such that

$$p(x,y) \geq 1/2 \text{ if and only if } w(x) \geq w(y),\ x,y \in A .$$

Definition 2. A strong (or Fechnerian) (binary) utility model is a set of binary preference probabilities for which there exist a real-valued function u over A and a cumulative distribution function ϕ such that

(i) $\phi(0) = 1/2$ and

(ii) for all $x,y \in A$ for which $p(x,y) \neq 0$ or 1,

$$p(x,y) = \phi[u(x)-u(y)] .$$

Definition 3. A strict binary utility model is a set of binary preference probabilities for which there exists a positive real-valued function v over A such that for all $x,y \in A$ for which $p(x,y) \neq 0$ or 1,

$$p(x,y) = \frac{v(x)}{v(x)+v(y)} .$$

Definition 4. A set of binary preference probabilities satisfies the triangle condition if for every $x,y,z \in A$,

$$p(x,y)+p(y,z) \geq p(x,z) .$$

Definition 5. Whenever $\min[p(x,y),p(y,z)] \geq 1/2$, the binary preference probabilities are said to satisfy

(i) weak (stochastic) transitivity provided that

$$p(x,z) \geq 1/2 \;;$$

(ii) moderate (stochastic) transitivity provided that

$$p(x,z) \geq \min[p(x,y),p(y,z)] \leq$$

(iii) strong (stochastic) transitivity provided that

$$p(x,z) \geq \max[p(x,y),p(y,z)] \;.$$

Definition 6. A set of binary preference probabilities satisfy the quadruple condition provided that $p(w,x) \geq p(y,z)$ implies $p(x,y) \geq p(x,z)$.

Definition 7. A set of binary preference probabilities not equal to 0 or 1 satisfy the product rule if for every set of three distinct elements $x_1, x_2, x_3 \in A$,

$$p(x_1,x_2)p(x_2,x_3)p(x_3,x_1) = p(x_1,x_3)p(x_3,x_2)p(x_2,x_1) \;.$$

Appendix B

Under the conditions of Theorem 19 any strong conjoint utility model is necessarily of either exponential or identity type. However, if $X \equiv X^*$ then the conditions of that theorem cannot be met and in this Appendix we study a set of conditions which are applicable to this case, these conditions being summarized as:

<u>Assumption A</u>. R_i, $i = 1,2$, <u>are subsets of the real numbers and</u> $u: R_1 \times R_1 \to R_2$ <u>is a continuous, real valued function which is dependent upon each of its arguments with</u>

i) $R_i =$ $= R =$ <u>the real numbers if variable</u> $i = 1,2$ <u>is measured on an interval scale</u>,

$= R_+ =$ <u>the nonnegative real numbers if variable</u> $i = 1,2$ <u>is measured on a ratio scale</u>,

<u>and</u> ii) <u>if</u> $\underline{T}_i$, $i = 1,2$, <u>denotes the class of admissible transformations of variable</u> i <u>then for any</u> $x,y \in R_1$, $T \in \underline{T}_1$, <u>there exists</u> $D(T) \in T_2$ <u>such that</u>

$$u(Tx,Ty) = D(T)u(x,y) \; . \tag{1}$$

Assumption A holds throughout this Appendix and we consider only those cases in which each variable is either a ratio or interval scale.

<u>Theorem 1</u>. <u>If both the independent and dependent variables are ratio scales then there exist a continuous function</u> f <u>and a constant</u> k <u>such that</u>

$$u(x,y) = y^k f(x/y) \; .$$

Proof. In this case Eq. 1 asserts that for $x,y,z \in R_+$ there exists $D(z) \in R_+$ such that

$$u(zx,zy) = D(z)u(x,y) \tag{2}$$

and hence for arbitrary $z,x,r,s \in R_+$,

$$u(zxr,zxs) = D(zx)u(r,s)$$
$$= D(z)D(x)u(r,s)$$

which implies that either $u(r,s) \equiv 0$, which is not the case since u depends on both its arguments, or for all $z,x \in R_+$

$$D(zx) = D(z)D(x) \ .$$

D is a continuous function and the solution of the above equation is $D(z) = z^k$ for some constant k .
Using Eq. 2 with $z = 1/y$ we obtain that

$$u(x/y,1) = D(1/y)u(x,y)$$

and hence

$$u(x,y) = \frac{1}{D(1/y)} u(x/y,1)$$

$$= y^k f(x/y)$$

where $f(x/y) = u(x/y,1)$ QED.

This result is a special case of the $\prod$-theorem of dimensional analysis [see, for example, Sedov (1959), Pp. 16-20].

Theorem 2. If the independent variables are ratio scales and the dependent variable an interval scale, then either there exist a continuous function f and constant β such that

$$u(x,y) = f(x/y)+\beta \log y$$

or there exist a continuous function f and constants k, γ such that

$$u(x,y) = y^k[f(x/y)-\gamma]+\gamma \ .$$

Proof. In this case Eq. 1 asserts that for $x,y,z \in R_+$ there exists $D(z) \in R_+$, $C(z) \in R$ such that

$$u(zx,zy) = D(z)u(x,y)+C(z) \ . \tag{3}$$

Using this equation in two ways we obtain that

$$\begin{aligned} u(zx,zx) &= D(z)u(x,x)+C(z) \\ &= D(z)[D(x)u(1,1)+C(x)]+C(z) \\ &= D(x)u(z,z)+C(x) \\ &= D(x)[D(z)u(1,1)+C(z)]+C(x) \end{aligned}$$

which implies that

$$D(z)C(x)+C(z) = D(x)C(z)+C(x)$$

i.e. $$C(x)[1-D(z)] = C(z)[1-D(x)] \ .$$

Since x,z are arbitrary either $C \equiv 0$, $D \equiv 1$ or $C(x) = \gamma[1-D(x)]$ for all $x \in R_+$ and some constant γ. If $C \equiv 0$ then u is in fact a ratio scale and Theorem 1 applies.

We now consider the other two cases.

1. $D \equiv 1$.

In this case, using Eq. 3 in two ways we obtain that

$$\begin{aligned} u(zx,zx) &= D(zx)u(1,1)+C(zx) \\ &= u(1,1)+C(zx) \\ &= D(z)u(x,x)+C(z) \\ &= D(z)[D(x)u(1,1)+C(x)]+C(z) \\ &= u(1,1)+C(x)+C(z) \end{aligned}$$

and hence $C(z)+C(x) = C(zx)$ for all $z,x \in R_+$. C is continuous and the solution of this functional equation is $C(x) = \beta \log x$ where β is a constant. Using Eq. 3 with $z = 1/y$ we obtain that

$$u(x/y,1) = D(1/y)u(x,y)+C(1/y)$$

$$= u(x,y)+\beta \log (1/y)$$

and hence $u(x,y) = f(x/y)+ \beta \log y$

where $f(x/y) = u(x/y,1)$.

2. If $C(x) = \gamma[1-D(x)]$ for all $x \in R_+$ then in particular

$$C(1) = \gamma[1-D(1)] \tag{4}$$

whereas Eq. 3 implies that for arbitrary $y \in R_+$,

$$u(1,y) = D(1)u(1,y)+C(1)$$

and hence

$$C(1) = u(1,y)[1-D(1)] . \tag{5}$$

Comparing Eqs. 4 and 5 we see that u does not depend on its second argument unless $D(1) = 1$ in which case there exists some $w \in R_+$ such that $u(1,w) \neq \gamma$.

For this w and arbitrary $z,x \in R_+$, Eq. 3 implies that

$$\begin{aligned} u(zx,zxw) &= D(zx)u(1,w)+C(zx) \\ &= D(zx)A+\gamma \end{aligned} \tag{6}$$

where $A = u(1,w)-\gamma \neq 0$.

Using Eq. 3 it can also be shown that

$$u(zx,zxw) = D(z)D(x)A+\gamma \tag{7}$$

and Eqs. 6 and 7 imply that

$$D(zx) = D(z)D(x)$$

for all $z \in R_+$. D is continuous and the solution of this equation is $D(z) = z^k$ for some k .

Using Eq. 3 with $z = 1/y$, we obtain that

$$\begin{aligned} u(x/y,1) &= D(1/y)u(x,y)+C(1/y) \\ &= D(1/y)u(x,y)+\gamma[1-D(1/y)] \end{aligned}$$

and hence

$$\begin{aligned} u(x,y) &= \{u(x/y,1)-\gamma[1-D(1/y)]\} \frac{1}{D(1/y)} \\ &= y^k[u(x/y,1)-\gamma(1-\frac{1}{y^k})] \end{aligned}$$

$$= y^k[u(x/y,1)-\gamma]+\gamma$$

$$= y^k[f(x/y)-\gamma]+\gamma$$

where $f(x/y) = u(x/y,1)$ QED.

Theorem 3. <u>If the independent variables are interval scales and the dependent variable a ratio scale then there exist constants</u> k <u>and</u> $\alpha,\beta > 0$ <u>such that</u>

$$u(x,y) = \begin{cases} \alpha(x-y)^k & x \geq y \\ \beta(y-x)^k & x < y . \end{cases}$$

Proof. In this case Eq. 1 asserts that for $x,y,\alpha \in R$, $z \in R_+$ there exists $D(z,\alpha) \in R_+$ such that

$$u(zx+\alpha,zy+\alpha) = D(z,\alpha)u(x,y) . \tag{8}$$

When $\alpha = 0$ Eq. 8 reduces to Eq. 2 and hence the arguments of Theorem 1 show that $D(z,0) = z^k$ for some constant k. Also Eq. 8 with $z = 1$ implies that for arbitrary $x,y,\alpha,\beta \in R$

$$\begin{aligned} u(x+\alpha+\beta,y+\alpha+\beta) &= D(1,\alpha+\beta)u(x,y) \\ &= D(1,\beta)u(x+\alpha,y+\alpha) \\ &= D(1,\beta)D(1,\alpha)u(x,y) \end{aligned}$$

and hence either $u \equiv 0$ or $D(1,\alpha+\beta) = D(1,\beta)D(1,\alpha)$ for all $\alpha,\beta \in R$. D is continuous and hence $D(1,\alpha) = \exp A\alpha$ for some constant A. Now assume that $x \geq y$. Eq. 8 then implies that

$$\begin{aligned} u(x,y) &= u[(x-y)+y,0+y] \\ &= D(1,y)u(x-y,0) \\ &= D(1,y)D(x-y,0)u(1,0) \\ &= e^{Ay}(x-y)^k u(1,0) . \end{aligned}$$

Because u depends on both its arguments, the above equation implies that $u(1,0) \neq 0$. Substituting the expression for $u(x,y)$ into Eq. 8 with $\alpha = 0$ and equating terms, we obtain that $A = 0$ and hence for $x \geq y$,

$$u(x,y) = (x-y)^k u(1,0) .$$

A similar argument shows that for $x < y$,

$$u(x,y) = (y-x)^k u(0,1) .$$

QED.

We have not yet obtained a satisfactory proof of the following conjecture.

<u>Conjecture</u>. <u>If both the independent and dependent variables are interval scales then either there exist</u> α,β,γ <u>and</u> k <u>such that for</u> $x,y \in R$,

$$u(x,y) = \begin{cases} \alpha(x-y)^k + \gamma & \underline{\text{for}} \quad x \geq y \\ \beta(y-x)^k + \gamma & \underline{\text{for}} \quad y < x \end{cases}$$

<u>or there exist</u> α, β <u>and</u> γ <u>such that</u>

$$u(x,y) = \alpha x + \beta y + \gamma$$

<u>or there exist</u> α, β <u>and</u> γ <u>such that</u>

$$u(x,y) = \begin{cases} \alpha \log(x-y) + \gamma & \underline{\text{for}} \quad x \geq y \\ \beta \log(y-x) + \gamma & \underline{\text{for}} \quad y < x . \end{cases}$$