Discussion Paper No. 498

MEASUREMENT ERROR THEORIES FOR VON NEUMANN-MORGENSTERN UTILITY FUNCTIONS

by

Jehoshua Eliashberg* and John Hauser**

September 1981

* Northwestern University

** Massachusetts Institute of Technology

ABSTRACT

Von Neumann-Morgenstern (vN-M) utility theory is the dominant model of risk preference in decision analysis and in marketing science. Most applications in both disciplines assess utility parameters by asking a series of questions which deterministically specify the parameters. However, any questioning procedure introduces measurement error and any model is, at best, an approximation.

This paper develops a measurement error theory for vN-M utility estimation. We examine the relationship among question formats and provide maximum likelihood estimators for the parameters of vN-M utility functions. Uncertainty in utility parameters induces uncertainty in the decision outcome. Thus, we provide estimates of the probability that a given alternative maximizes expected utility.

Our general results may require numerical methods, but for the two most widely used functional forms, we provide analytic formulae that are easy to use. Since the unique aspect of vN-M theory is risk modeling, we emphasize uniattributed utility functions then discuss multiattributed extensions.

Numerical examples illustrate the results.

1. PERSPECTIVE

The measurement and modeling of how individuals form preferences among risky alternatives is an important problem in decision analysis, marketing, and psychology.

In decision analysis, the focus is prescriptive. The decision maker faces a complex and important problem, say the decision of whether or not to rely on nuclear power. The outcomes of this decision are characterized by multiple objectives such as economic, national defense, environment and safety considerations, and by risk in the sense that the amount of achievement of each objective is uncertain before the decision is made. The preference modeling task is to assess the decision maker's utility function over the risky consequences in order to help the decision maker decide among his strategic options.

In marketing, the focus is descriptive. The marketing scientist wishes to predict which product consumers will purchase. Risk becomes important for major purchases such as a home heating system where the outcomes, such as annual cost, fuel availability, cleanliness, reliability, and heating comfort may be uncertain to the consumer. The preference modeling task is to estimate consumer utility functions to describe how consumers will respond to marketing strategies such as product modifications or advertising designed to remove perceived risk.

In psychology, the focus is investigative. The psychologist wants to understand how individuals really process information. Thus, a paramorphic model is not sufficient. The psychologist wants to distinguish the process that best represents cognitive information processing. The preference modeling task is to distinguish one processing model from another.

Clearly, the focus and use of preference models varies across the disciplines. However, each discipline has found von Neumann-Morgenstern utility theory to provide a theoretical foundation from which to address the task of modeling preferences among risky alternatives. In decision analysis, Keeney and Raiffa [38] and Farquhar [16] describe over 20 applications, while Fishburn and Kochenberger [18] examine the commonalities in 30 applications. In marketing, Hauser and Urban [30] have used vN-M theory to measure functions to describe consumer response to health care delivery and Eliashberg [11] has described consumer response to the student housing market. In psychology, Goodman, Saltzman, Edwards, and Krantz [21], Lee [41], and Tversky [52] have investigated the descriptive validity of the theory. Results are encouraging and such investigations have led to improved models such as prospect theory (Kahneman and Tversky [34]).

To date, in all three disciplines, the theory has been used to investigate the conditions under which vN-M measurement should apply and to investigate which functional forms are appropriate if behavioral assumptions apply. Despite the power of the axiomatic theory, measurement has proceeded deterministically. For example, Keeney and Raiffa [38] provide many procedures for assessing utility functions, but each of these procedures is deterministic in the sense that measurement error is not explicitly modeled. The typical situation requires k questions to assess the k parameters of a selected utility function. Additional questions check consistency but to quote Keeney and Raiffa [38, p. 197]: "Unfortunately, no general procedure exists either for determining whether a given set of qualitative and quantitative assessments are consistent or indicating an appropriate functional form of the utility function when assessments are consistent". Meyer and Pratt [45] do provide a procedure for "fairing" a smooth function through a set of points but their procedure is a fitting procedure, not an estimation based on an explicit error theory. Even in marketing, which explicitly acknowledges measurement error, "estimation" of vN-M functions has proceeded by the solution of algebraic equations, not by statistical inference. See Hauser and Urban [31,appendix] for illustrative example. Redundant questions are used to test independence conditions, not to estimate parameters or to estimate measurement error.

Our research seeks to develop the foundations of an error theory for vN-M utility theory. The key feature of vN-M theory that distinguishes it from other utility estimation procedures (e.g., conjoint analysis, preference regression, logit analysis, expectancy value measurement, preference intensity measures) is that it measures preferences over risky alternatives. Thus, we can illustrate much of the properties of an error theory with the estimation of a uniattributed function with one unknown parameter which measures risk preference. For this case we provide maximum likelihood estimators for various functional forms, questionning formats, and assumptions about the nature of measurement error. Since any uncertainty in the parameter of a vN-M utility function induces uncertainty in the utility function and hence the decision outcome, we also provide a means to characterize this uncertainty and estimate the probabilities that a decision indeed maximizes expected utility. In marketing, this probability is interpreted as the probability a consumer will choose an alternative. In decision analysis, this probability is the probability that a given decision is the "correct" decision.

We extend the single parameter, uniattributed results to multiple parameters and the multiple attributes. Analytic results are less tractable, but they are usable, particularly if we allow a clustered questioning format.

After an expository illustration of the theory, we discuss the limitations of our results and suggest directions for future research on this topic.

We begin a brief review of the literature.

2. EXISTING LITERATURES

The utility assessment or estimation litereatures in decision analysis, marketing, and psychology are each extensive. This review is purposively concise emphasizing the previous work relevant to the development of a vN-M error theory. For more extensive reviews, see Green and Srinivasan [25], Keeney and Raiffa [38] and Krantz, Luce, Suppes and Tversky [39].

Decision Analysis

We draw on decision analysis for axiomatic foundations, derived functional forms and empirical experience.

Axiomatic Foundations. VN-M utility theory takes its name from a 1947 book by von-Neumann and Morgenstern [55] who provided three fundamental axioms which, if satisfied, imply the existence and uniqueness (subject to a scaling change) of a utility function that summarizes an individual's preference over risky alternatives. Such functions have proven useful in decision analysis because, if the axioms apply, the most preferred strategy is that strategy which maximizes expected utility. Thus, independent quantification of the probability distribution over outcomes and the utility function indicates which strategy should be selected. These axioms have been reformulated by Friedman and Savage [19], Herstein and Milner [32], Jensen [33], Marshak [43] and possibly others, but the axioms remain the foundation for prescriptively modeling preferences among risky alternatives.

Derived Functional Forms. As vN-M theory was applied and extended from preferences over money to preferences over multiple objectives, researchers recognized that they must provide a structure with which to assess an individual's utility function. This area of research, summarized by Keeney and Raiffa [38], provides a set of theorems which derive specific functional forms from qualitative assumptions.

Uniattributed functions are derived from assumptions about how an individual's risk preference changes as his assets increase. For example, we expect a firm to be less risk averse for lotteries involving thousands of dollars if their assets are \$100 million than if their assets are \$100,000. A key concept is local risk aversion, R(x), which indicates how an individual's risk attitude depends on assets, x. The value this function takes on for a given asset level indicates whether a utility function is risk prone or risk averse at that asset level. Besides providing a monotonic scale of risk aversion, R(x) uniquely determines a utility function subject only to a change of scale. In particular, if u(x) is the utility function, R(x) is given by Pratt [48]:

$$R(x) = -\frac{d^2u(x)}{dx^2} / \frac{du(x)}{dx} \qquad (1a)$$

Eliashberg and Winkler [12,13] show that this concept is focal for modeling multiple-person decision making such as that involved in industrial products. See also discussion in Choffray and Lilien [8].

Another important concept in vN-M utility theory is proportional risk aversion, S(x). It measures an individual's risk attitude when consequences of risky decisions are expressed in proportion to assets. (E.g., a new product launch with an uncertain return on investment.) The relationship between local risk aversion, R(x), and proportional risk aversion, S(x), is given by:

$$S(x) = xR(x) \qquad (1b)$$

For clarity we refer to R(x) as "absolute" risk aversion for the remainder of this paper.

We will find R(x) and S(x) useful in the development of an error theory. Table 1 summarizes some of the commonly used functions and the assumptions upon which they are based. The last column, "inverse function", will be explained in Section 3. Note that forms 1 to 4 require only one unknown parameter, r, whereas form 5 requires more than one unknown parameter.

TABLE 1

SELECTED UNIATTRIBUTED vN-M UTILITY FUNCTIONS

Behavioral Assumption	Functional Form	Range	Inverse Function
1. Risk aversion independent of asset position, i.e., constant absolute risk aversion.	$1 - e^{-rx}$ $r > 0$	$0 \le x \le \infty$	$-(1/x)\ln(1-p)$
	$\frac{[1-e^{-r(x-x_0)}]}{[1-e^{-r(x_*-x_0)}]}$	$x_0 \le x \le x_*$	implicit except for special cases
2. Risk aversion (seeking) inversely proportional to asset position above x_0, i.e., constant proportional risk aversion	$(x-x_0)^r/(x_*-x_0)^r$ $r>0$,	$x_0 \le x \le x_*$	$\log p/\log [(x-x_0)/(x_*-x_0)]$
	$-\log (x-x_0)$ $r=0$		see note
	$\sim (x-x_0)^r$ $r<0$		~ indicates strategic equivalence
3. Risk neutral	$(x-x_0)/(x_*-x_0)$	$x_0 \le x \le x_*$	special case of (1) where $r \to 0$ and (2) where $r \to 1$
4. Decreasingly risk averse I	$\frac{\log[(x+r)/(x_0+r)]}{\log [(x_*+r)/(x_0+r)]}$ $r > -x_0$	$x_0 \le x \le x_*$	implicit except for special cases
5. Decreasingly risk averse II	$\frac{\left(1+r_3-e^{-r_1 x}-r_3 e^{-r_2 x}\right)}{(1+r_3)}$ $r_1, r_2, r_3 > 0$	$0 \le x \le \infty$	---

Note: For $r \le 0$ in form 2 the form is undefined for $x \to x_0$. Throughout this paper, we assume $r > 0$. Similar results can be derived for $r \le 0$ if we deal with $x \ge x_0 + \varepsilon$ where ε is a small positive finite number.

Multiattributed functions are derived from assumptions of independence among attributes. There is a portfolio of functions which allow the analyst to make tradeoffs among generality (less restrictive assumptions, more complex assessment) and parsimony (more restrictive assumptions, less complex assessment). See, for example, Farquhar [15,16]. One widely used set of functions, developed by Keeney [35,37] and Fishburn [17], provides a reasonable tradeoff among generality and parsimony. Keeney [35] shows that if risk preferences for each attribute do not depend upon other attributes then the utility function, $U(\underline{x})$, over a vector of attributes can be written as:

$$U(\underline{x}) = \Sigma_k w_k u_k(x_k) + \sum_{\ell>k} \Sigma_k w_{\ell k} u_k(x_\ell) u_k(x_k)$$

$$+ \ldots + w_{1\ldots K} u_1(x_1) u_2(x_1) \ldots u_K(x_K) \qquad (2)$$

where u (u) is a uniattributed utility function scaled from 0.0 to 1.0 and w_k, $w_{\ell k}$,... are scaling constants. This is unknown as the "multilinear" form. Keeney [37] then shows that if tradeoffs among any two attributes are independent of the other attributes, $w_{\ell k} = W w_\ell w_k$ where W is a constant. This is known as the "multiplicative" form. Fishburn [17] identifies a condition, known as marginality, which, when true, implies that W = 0. When marginality applies, the higher order terms in equation (2) disappear and the utility function becomes an additive representation. Finally, Richard [49] shows that W can be interpreted as a measure of multiattributed risk aversion.

Empirical Experience. VN-M utility assessment has been applied to a wide range of problems in both the public and private sector. See Keeney and Raiffa [38, Chapter 7] for a review. Fishburn and Kochenberger [18] have examined 30 of the assessments which deal with a uniattributed utility function for wealth. (Their sources are Barnes and Reinmuth [2], Grayson [22], Green [23], Halter and Dean [27], and Swalm [50].) They conclude that:

(1) There is usually a monetary value, referred to as the target level, at which the utility function changes shape. The target level is often the zero-gain point.

(2) Below target utility (usually the loss region) is frequently risk prone (convex).

(3) Above target utility (usually the gain region) is frequently risk averse (concave).

(4) The utility is almost always steeper below target than above.

Fishburn and Kochenberger [18] then used a least squares criterion to identify the single parameter for the constant risk averse, constant proportional risk averse, and linear utility functions (forms 1, 2, and 3 in Table 1) where separate fits were used above and below target. They found that the non-linear forms (1 and 2) fit substantially better than the linear form (3), that the constant risk averse form (1) fits better for concave below, convex above utilities and the constant proportional risk averse form (2) fits better for convex below, concave above utilities. However, Fishburn and Kochenberger did not develop an explicit error theory.

There is no equivalent study for multiattributed preferences although researchers found that over certain ranges a multiplicative or additive function was a reasonable tradeoff among generality and parsimony. As discussed in Keeney and Raiffa [38], these applications include air pollution control (Ellis and Keeney [14]), fire department operation (Keeney [36]), corporate objectives, siting and licensing of nuclear power facilities (Gros [26]), policies for frozen blood (Bodily [4]), treatment for cleft palate (Krischner [40]), water quality (O'Connor [47]), foreign policy (Brown and Peterson [6]), and forest pest management (Bell [3]).

Marketing Science

While much work in marketing science has addressed the estimation of consumer preference functions, only recently have researchers addressed the descriptive validity of vN-M theory.

Preference Estimation. The measurement and modeling of consumer preferences is a necessary step in addressing many marketing problems, particularly product development, and there are hundreds of applications yearly. The methods include conjoint analysis (Green and Srinivasan [25]), preference regression (Urban [53]), logit analysis (McFadden [44]), intensity theory (Hauser and Shugan [29]), and expectancy values (Wilkie and Pessemier [56]). See Urban and Hauser [54, Chapter 10] for a review. All methods except expectancy value are based on a statistical error theory which enables the analyst to obtain estimates of the underlying parameters. However, since none of these methods explicitly consider uncertainty in the attributes (risk), vN-M theory is valuable when risk is an important consideration.

VN-M Applications. Hauser [28] has shown that with the addition of a choice axiom, vN-M theory applies axiomatically to descriptive modeling. Based on this theoretical result, Hauser and Urban [31] assessed a multiplicative, constantly risk averse utility function for the attributes of health care plans. They report that for their sample, vN-M utility assessment predicts better (50% correct prediction) than preference regression (47%) or logit analysis (43%) and significantly better than an equal weighting scheme (40%). More importantly, vN-M theory correctly predicted market shares while all other methods were in error at the .05 level of statistical significance. Eliashberg [11] assessed multilinear and additive piecewise linear utility functions for the attributes of student housing. He also reports reasonable prediction (53% correct). Although both studies collected redundant information, the information was used to test the adequacy of the behavioral assumptions rather than to estimate parameters.

Mathematical Psychology

Mathematical psychologists are concerned with the question of whether vN-M theory is an adequate description of human information processing. That is, they wish to identify and test qualitative properties which imply the existence of a decision calculus based on expected utility where the expectation is over subjective probability estimates. Goodman, Saltzman, Edwards, and Krantz [21], Lee [41], and Tversky [52] have all investigated the descriptive validity of expected utility and Krantz, Luce, Suppes, and Tversky [39] have investigated axiom systems and representation theorems. They concluded that although some researchers have investigated alternative formulations, " expected utility theory has reigned as the major theory of individual decision making under uncertainty". [39, p. 398].

Recently, Kahneman and Tversky [34] have relaxed some of the assumptions of expected utility theory. They (1) allow preference functions which are convex in losses but concave in gains and (2) extend the axioms to include a transformation of the probabilities used in calculating the expectation. Their formulation, called "prospect theory", overcomes many of the psychologists' theoretical objections to vN-M theory.

Discussion. VN-M theory has received much attention in the literature as the dominant method for modeling preferences with respect to risky alternatives. However, an error theory can improve applications in each literature.

In decision analysis, assessment questions are often difficult to answer leading to potential measurement error. Our error theory enables the analyst to ask multiple questions to infer the true values (mean) and variances of the parameters of the uniattributed utility functions.

In marketing science, the analyst is limited in the number of questions he can ask in a consumer interview. Our error theory allows him either (1) to infer risk parameters from multiple questions asked of each individual or (2) to characterize the risk parameters of a consumer population as a probability distribution inferred from questions which may or may not vary across individual consumers. Once the parameters are estimated, our error theory allows the marketing scientist to forecast the probability that a risky alternative is chosen.

In mathematical psychology, our error theory provides statistical statements about the distribution of risk parameters. Such statistics are useful in testing when vN-M is applicable and when a more general theory, such as propect theory, must be used. Finally, prospect theory also requires the estimation of utility-like functions. An error theory for vN-M utility functions could form the basis for an error theory for the more general prospect functions, once elicitation procedures are developed.

We begin our exposition with the development of an error theory for single-parameter, uniattributed utility functions.

3. SINGLE-PARAMETER UNIATTRIBUTED UTILITY FUNCTIONS

We provide estimators and choice probabilities for general single-parameter functions. In the general case, such results may require numerical techniques for inverting functions and integration. However, since the empirical evidence cited in section 2 suggests that forms 1, 2, and 3 in table 1 are the most commonly used utility functions we provide analytic estimators and choice probabilities for these special cases. Since these results are analytically tractable and forms 1 and 2 are quite flexible, we expect these results to be useful for most applications in decision analysis and marketing.

Our analysis applies at the individual level or at the population level.

At the individual level, variation in the unknown parameter represents uncertainty in estimating that parameter. At the population level, variation in the unknown parameter represents heterogeneity among individual consumers. Since these two problems are mathematically equivalent we base our theory on the assumption that $i = 1,2,\ldots, I$ questions are asked. The index, i, can characterize variation within or across individuals. For simplicity of exposition we discuss variation within individuals.

Following Fishburn and Kochenberger [18], we proceed with separate estimations in the "gains" and in the "loss" regions. Thus, we can assume that utility function is either concave or convex throughout the region, $x_0 \le x_i \le x_*$. We state the results for concave risk averse utility functions since the results for convex utility functions are quite similar. Without loss of generality, we assume that the attribute of interest, x, has been scaled such that preference is monotonically increasing in x over the region of estimation.

Estimation

Since the distinguishing feature of vN-M utility theory is its ability to model preference for risky alternatives, measurement incorporates at least one alternative with an uncertain level of the attribute. Since the vN-M axioms enable us to model risk by standard gambles, we use such gambles as our basic unit of measurement. The basic conceptual measurement is shown in Figure 1. The individual is given a lottery which has an outcome of x_* with probability, p_i, and an outcome of x_0 with probability, $(1-p_i)$. He is asked to compare this lottery to a "certainty equivalent" of x_i. The researcher selects x_0, x_* and either x_i or p_i. If x_i is given, the individual is asked to provide the probability, p_i, such that he is indifferent among the two alternatives. If p_i is given, the individual supplies x_i. See

Keeney and Raiffa [38] for specific questions for decision makers and Hauser and Urban [31] and Eliashberg [11] for specific questions for consumers.

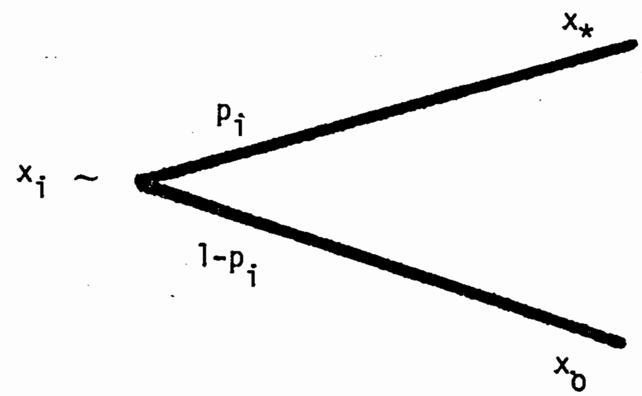

Figure 1: Standard Lottery Used as Basic Risk Measure

Let r be the unknown parameter of the utility function, $u(x,r)$. The standard procedure of assessment is to solve an algebraic equation derived from the expected utility property by scaling the utility function such that $u(x_*,r) = 1.0$ and $u(x_0,r) = 0.0$. Then, if there were no error, the implication of Figure 1 would be:

$$u(x_i,r) = p_i u(x_*,r) + (1-p_i)u(x_0,r) = p_i \qquad (3)$$

The utility parameter, r, is determined by inverting equation (3). For example, if $u(x_i,r)$ is given by a special case of form 2, x_i^r, for $x_i \in [0,1]$, then $r = \log p_i/\log x_i$. In general, we can define an inverse function, $r(x_i,p_i)$ into a real value, r_i, determined by the solution of equation (3). In some cases, such as the power function (form 2 in Table 1) and the infinite range exponential function (form 1 in Table 1) the inverse function is analytic; in other cases, such as the logarithmic function (form 4 in Table 1), the function exists but is not closed form. Various inverse functions are given in Table 1.

Each time we ask a question we introduce error. We choose to model this error as inducing a probability distribution on our observation of the individual's risk preference, r. This assumption is consistent with "random utility" error theories, e.g., Thurstone [51] and Luce and Suppes [42], but modified to emphasize the unique strength of vN-M theory--risk preference.

Our assumption implies that the observed value, r_i, of the risk parameter is a random draw from some probability density function, $f(r_i|\lambda)$, where λ is a parameter to be estimated. (We allow λ to be vector valued.) The density function for the risk parameter, $f(r_i|\lambda)$, in turn, induces a density function on the utility function, $g(u_j|\lambda)$ where u_j is the value of the utility function for any given value of the attribute, x_j. In fact:

$$g(u_j|\lambda) = f[\mathcal{r}(x_j,u_j)|\lambda]|d\mathcal{r}(x_j,u_j)/du_j| \qquad (4)$$

where $\mathcal{r}(x_j,u_j)$ is obtained by solving the equation, $u_j = u(x_j,r)$, for r. (See Mood and Greybill [46, p.224].) Since $\mathcal{r}(x_j,u_j) = r(x_j,p_j)$ for $u_j = p_j$ we drop the distinction for the remainder of the paper.

Before we estimate λ we must further specify $f(r_i|\lambda)$. The most natural assumption is that $f(r_i|\lambda)$ is normal with mean, μ, and variance, σ^2. Thus, we derive estimators for a normal error theory.

While a normal error theory is appropriate for most cases, it does have one drawback. Theoretically, if r is normally distributed, it can take on any value in the range, $(-\infty,\infty)$, but some functional forms for $u(x,r)$ are restricted by qualitative judgements to a more specific range, usually $(0,\infty)$. For example, if an individual is judged risk averse with risk aversion independent of asset position (form 1 in Table 1), we expect r to be positive. When $\mu >> \sigma$, this case can be handled with normal error theory since negative values of r are unlikely. Alternatively, we can model the error as inducing a negative exponential distribution on r for the range, (r_0,∞). I.e., $f(r|\lambda) = (\lambda-r_0)^{-1} \exp[-(r-r_0)/(\lambda-r_0)]$.

We develop an error theory for both normal and exponential assumptions. The former is useful when r is unrestricted or when $\mu >> \sigma$. The latter is useful when we restrict r to be greater than r_0 by qualitative assumptions, and we are willing to accept the consequences of an exponential distribution which has its maximum value at r_0.

The first issue we address is the question format. Hauser and Urban [31] advocate setting x_i and having the individual specify p_i. The stated reason is that the scale properties of x_i may not be known. Specifying p_i simultaneously scales x_i and $u(x_i, r)$ thus making it easier to interpret what "utility of x_i" means. On the other hand, Keeney and Raiffa [38] and Eliashberg [11] advocate setting p_i and having the individual specify x_i. They believe that this is the easier task.

Our first result suggests that the estimators do not depend on question format.

Theorem 1: If the utility function, $u(x,r)$*, is a bijection from the range of its single parameter,* r*, onto* $[0,1]$*, if measurement errors are independent across measurements, and error is modeled as a density function on* r*, the maximum likelihood estimators of the parameters charaterizing the distribution of* r *are invariant with respect to the questionning format.*

Proof. Suppose that the researcher specifies x_i and the individual provides p_i for each of I measurements. If measurement errors are independent across measurements, then for any vector of certainty equivalents, $\underline{x} = (x_1, x_2, \ldots, x_I)$, the joint probability, $F(\underline{p}|\underline{x}, \lambda)$, of observing a vector of answers, $\underline{p} = (p_1, p_2, \ldots, p_I)$, is given by

$$F(\underline{p}|\underline{x},\lambda) = \Pi_i \; f[r(x_i,p_i)|\lambda] \left|\frac{d}{dp_i} r(x_i,p_i)\right| \qquad (5)$$

where $f(\cdot|\lambda)$ is the distribution of r, $r(x_i,p_i)$ is the inverse function, and

$|dr(x_i,p_i)/dp_i|$ is the absolute value of the Jacobian transformation (Mood and Greybill [46, p.224]).

To obtain the maximum likelihood estimators of the parameter, $\hat{\lambda}$, we maximize $F(\underline{p}|\underline{x},\lambda)$ with respect to λ. Since the Jacobian is independent of λ, this is equivalent to maximizing the following likelihood function:

$$L(\lambda|\underline{x},\underline{p}) = \Sigma_i \log f[r(x_i,p_i)|\lambda] \qquad (6)$$

Finally, by symmetry it is clear that we obtain the same likelihood function if the researcher provides the probability, p_i, and the individual answers with the certainty equivalent, x_i.

Theorem 1 does not resolve the debate over question format, but the result is important because it indicates that choice among formats is not a statistical issue but rather a debate over the ease of interviewing and interpretation and the magnitude of error introduced by alternative formats.

Our second result specifies the appropriate maximum likelihood estimators for normal and exponential error theories. Although the equation 7 and 8 are well known (Mood and Greybill [46, p. 183], Giri [20, p. 77], we state them as a theorem for emphasis.

Theorem 2: According to the assumptions of theorem 1, the maximum likelihood estimators for normal error theory are:

$$\hat{\mu} = (1/I)\ \Sigma_i\ r(x_i,p_i) \qquad (7)$$

$$\hat{\sigma}^2 = (1/I)\ \Sigma_i [r(x_i,p_i) - \hat{\mu}]^2 \qquad (8)$$

The maximum likelihood estimators for exponential error theory are:

$$\hat{\lambda} = (1/I)\ \Sigma_i\ r(x_i,p_i) \qquad (9)$$

Proof. Equations 7,8, and 9 are obtained by maximizing $L(\lambda|\underline{x},\underline{p})$ in equation 6.

Since the results are well known, we do not repeat the detailed maximization here.[1]

Equations 7, 8, and 9 are obvious with hindsight. Once we realize the power of the inverse function, $r(x_i,p_i)$, to transform a difficult statistical problem into standard statistical problem we proceed with standard estimation theory. This simplicity does not diminish the usefulness of the result as we illustrate with the following example.

Suppose that you are considering replacing your antiquated home heating system with a new oil, gas, electric or solar system. You are uncertain about unit fuel cost, about heating efficiency, and weather. Thus, the annual savings of each new system over your present system is characterized by a probability distribution over the range of $200 to $1200. You assess your utility function by supplying a probability, p_i, for each certainty equivalent (in $00's) between x_0 and x_*. If your underlying preferences are given by $u(x,r) = (x-200)^{1/2}/(1000)^{1/2}$ and you "round-off" to the nearest .05 when giving p_i, you would get the answers in Table 2. Since your preferences correspond to form 2 in Table 1, we use the inverse function in column 4 of Table 1 to compute $r(x_i,p_i)$. This is shown in Table 2. Finally, using Theorem 2 we estimate $\hat{\mu} = .50$ and $\hat{\sigma} = .03$ for normal errors and $\hat{\lambda} = .50$ for exponential errors where we have assumed $r_0 \equiv 0$. Figure 2 compares the estimated utility function and the observed points from Table 2.

---

[1]Equation 8 is the maximum likelihood estimate for σ^2, but it is biased for small I. The more common unbiased estimator is $[I/(I-1)]\sigma^2$. Also, if we want to estimate r_0 rather than set it by qualitative judgment, its estimator is $\hat{r}_0 = \min_i[r(x_i,p_i)]$.

TABLE 2

EXAMPLE ASSESSMENT FOR THE ANNUAL SAVINGS

OF A HOME HEATING SYSTEM

x_i (dollars)	p_i	$r(x_i,p_i)$ (power function)
300	.30	.52
400	.45	.50
500	.55	.50
600	.65	.47
700	.70	.51
800	.75	.56
900	.85	.46
1000	.90	.47
1100	.95	.49

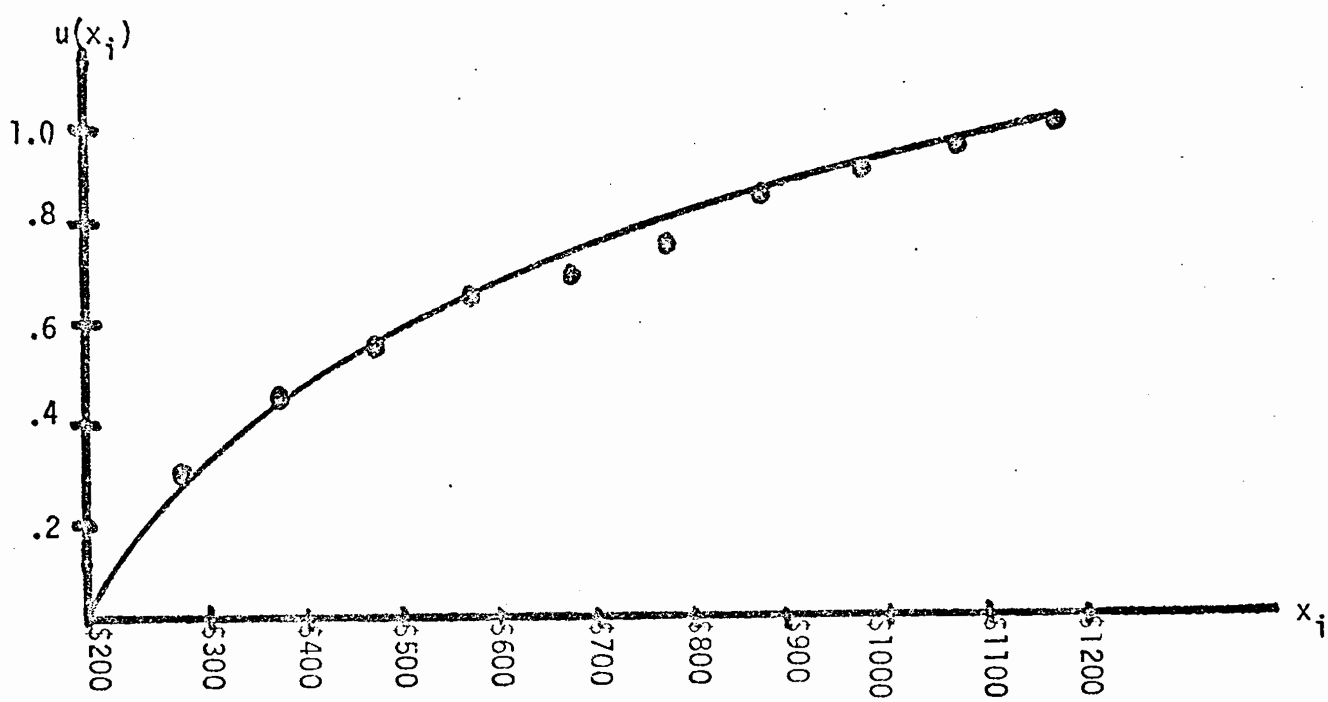

Figure 2: Maximum-likelihood Estimate of Assessed Utility Function

Because results 1 and 2 transform a vN-M error theory into a known statistical problem, we can apply many of the results of classical statistical inference. For example, if normal theory applies and the analyst believes that the true value of μ is μ^t, he can test this hypothesis by recognizing that $(\hat{\mu} - \mu^t)(I - 1)^{1/2}/\hat{\sigma}$ is distributed as a t-statistic with I-1 degrees of freedom. For example, for the power function in Table 2, a 95% confidence interval for μ is [.48, .52]. These confidence intervals for μ are useful since they become tighter as the number of observations, I, gets larger. Similarly, we can use a chi-squared distribution (Mood and Greybill [46, p.254]) to compute that a 95% confidence interval for $\hat{\sigma}$ is [.02, .06].

For exponential error theory the sampling distribution for $\hat{\lambda}$ has a gamma density with mean, λ, and variance, λ^2/n.

Probability of Choice

Once we have estimated the parameters of a utility function, we use that utility function to model decisions among risky alternatives. If r were known with certainty, we would know u(x,r) for any x and could compute directly the expected utility of any alternative. However, r is a random variable inducing randomness on u(x,r), and hence uncertainty on any decision among risky alternatives. Thus, with measurement error we can at most compute the probability, P_j, that alternative j is chosen. In marketing we interpret these probabilities as forecasts of behavior and in decision analysis, as the confidence that a decision is correct. Without loss of generality, we focus on alternative 1.

Single random draw. For most marketing problems and some decision analysis problems, we can model randomness as a single random draw from $f(r|\lambda)$. The interpretation is that our knowledge of the individual's utility function is uncertain but he is consistent in the sense of using the same utility function to evaluate all alternatives in his choice set.

Suppose there are J alternatives, each characterized by some probability density function, $h_j(x)$, on the single attribute. Then if r were known with certainty, the individual would select alternative that maximizes $\int u(x,r)h_j(x)dx$. However, r is a random variable and as it varies, different alternatives are likely to maximize utility. Thus define

$$\delta(r) = \begin{cases} 1 \text{ if } \int u(x,r)h_1(x)dx \geq \int u(x,r)h_j(x)dx \text{ for } j = 1,2,\ldots J \\ 0 \text{ otherwise.} \end{cases}$$

Then, the probability that alternative 1 is chosen is simply the probability that $\delta(r) = 1$. I.e.,

$$\hat{P}_1 = \int \delta(r)f(r|\hat{\lambda})dr \qquad (10)$$

In principal we can obtain $\delta(r)$ by numerical methods for any set of $h_j(x)$'s. Thus, in principal, $\hat{P}_1$ is computable. However, for practical problems we must obtain more analytic results.

We begin by examining the binary choice problem with dichotomous outcomes illustrated in Figure 3. Without loss of generality assume $x_1 > x_2$. This problem contains the essence of risky choice; the individual must decide among a potentially greater payoff, alternative 1, and a greater likelihood of the payoff, alternative 2. (We assume $\beta > \alpha$, otherwise a risk averse individual would always select alternative 1.) For this problem we can obtain an analytic expression for $\hat{P}_1$. (Let $\Phi[\cdot]$ be the standardized cumulative normal distribution.)

Figure 3: Binary Choice Problem

Theorem 3: For the binary choice problem in Figure 3, assume that $\mu >> \sigma$ *for normal and* $r_0=0$ *for exponential theory. Then for NORMAL theory:*

$\hat{p}_1 \cong \Phi[(\hat{\mu}-\kappa^{-1}\log(\beta/\alpha))/\hat{\sigma}]$ *for constant proportional risk aversion*

$$\hat{p}_1 \cong \begin{cases} \Phi[(r_c-\hat{\mu})/\hat{\sigma}] \text{ if } \alpha x_1 > \beta x_2 & \text{for constant absolute risk aversion (infinite range, } 0 \le x < \infty) \\ 0 \text{ otherwise} \end{cases}$$

where $\kappa = \log[(x_1-x_0)/(x_2-x_0)]$ *and* r_c *solves the equation*

$$\beta \exp(-r_c x_2) - \alpha \exp(-r_c x_1) = \beta - \alpha$$

For EXPONENTIAL ERROR theory:

$\hat{p}_1 = [\beta/\alpha]^{-\frac{1}{\lambda\kappa}}$ *for constant proportional risk aversion*

$$\hat{p}_1 = \begin{cases} 1-\exp[r_c/\hat{\lambda}] \text{ if } \alpha x_1 > \beta x_2 & \text{for constant absolute risk aversion (infinite range, } 0 \le x < \infty) \\ 0 \text{ otherwise} \end{cases}$$

Proof. We scale $u(x,r)$ such that $u(x_0,r) = 0$ and $u(x_*,r) = 1$. For the binary choice problem, $\delta(r) = 1$ if $\alpha u(x_1,r) > \beta u(x_2,r)$. Substituting the constant proportional risk aversion utility function, $u(x,r) = (x-x_0)^r/(x_*-x_0)^r$, yields $\delta(r) = 1$ iff $r \ge \log(\beta/\alpha)/\log[(x_1-x_0)/(x_2-x_1)] \equiv \kappa^{-1}\log(\beta/\alpha)$. Recognizing $r \sim N(\hat{\mu},\hat{\sigma}^2)$ and $\Phi((\mu-z)/\sigma) = \text{Prob}[r \ge z]$ yields the result. The result is only approximate since we ignore $r < 0$ which occurs with low probability when $\mu >> \sigma$. Now substituting the infinite range constant absolute risk aversion utility function, $u(x,r) = 1 - e^{-rx}$, into $\alpha u(x_1,r) = \beta u(x_2,r)$ yields the equation for r_c. Note that as $r \to \infty$, $u(x,r) \to 1$, and alternative 2 will be preferred since $\beta > \alpha$. As $r \to 0$, alternative 1 will be preferred if $\alpha x_1 > \beta x_2$ since $u(x,r)$ approaches linearity. Thus, $\delta(r) = 1$ for $0 \le r \le r_c$ and $\delta(r) = 0$ for $r > r_c$, if there is only one solution to the equation for $r_c \ge 0$. We provide a proof of this fact in Lemma 1. See appendix.

If $\alpha x_1 \leq \beta x_2$, then $\alpha u(x_1,r) < \beta u(x_2,r)$ for $r \geq 0$, hence $\hat{P}_1 \cong 0$. This also is formalized in Lemma 1. The results for exponential theory follow the same arguments except $\text{Prob}[r \geq z] = \exp(-z/\hat{\lambda})$. The result is exact since $\text{Prob}[r \leq 0] = 0$.

Theorem 3 derives analytic results for alternatives with dichotomous outcomes, that is, when $h_j(x)$ is a Binomial distribution. Theorem 3 can be extended to continuous distributions on $h_j(x)$. For example, if $h_j(x)$ is $N(\mu_j, \sigma_j^2)$ and the individual is constantly absolute risk averse, then $\int u(x,r)h_j(x)dx = 1- \exp(-r\mu_j+r^2\sigma_j^2/2)$. With a little algebra it is easy to show that alternative 1 will be chosen if $r < 2(\mu_1-\mu_2)/(\sigma_1^2-\sigma_2^2)$. Thus, if we substitute $r_c = 2(\mu_1-\mu_2)/(\sigma_1^2-\sigma_2^2)$ into Theorem 3, we have an estimate for $\hat{P}_1$ when $h_j(x)$ is normal. (Keeney and Raiffa [38, p. 202] provide a table of $\int u(x,r)h_j(x)dx$ for Beta, Binomial, Cauchy, Exponential, Gamma, Geometric, Normal, Poisson, and Uniform distributions for $h_j(x)$.)

Theorem 3 is a very usable result. For example, consider the hypothetical alternatives in Figure 4. Alternative 1 is oil heat where the high risk reflects volatile supplies. Alternative 2 is gas heat where the risk reflects only uncertainty in the heating characteristics of the home.

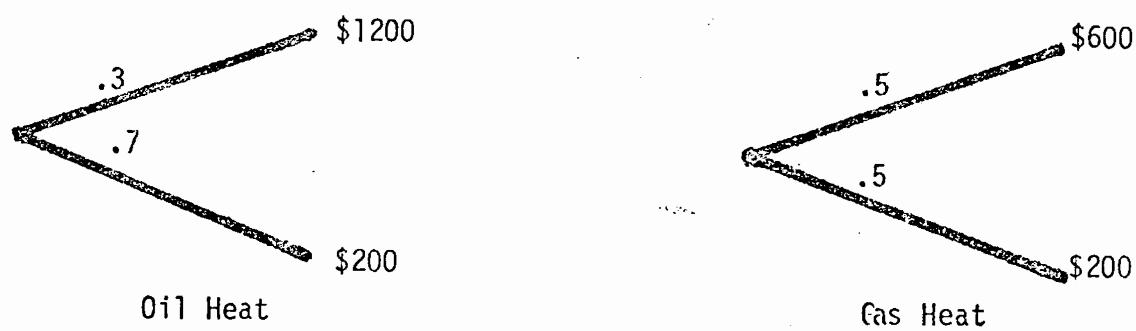

Figure 4: Hypothetical Characteristics of the Risk Involved for Two Home Heating Systems

Using the utility function in Table 2, $\hat{\mu} = .5$, $\hat{\sigma} = .03$. From Figure 4 $\alpha = .3$, $\beta = .5$, $x_1 = 1200$, and $x_2 = 600$. Assuming a constant proportional risk averse utility function and substituting these values in Theorem 3 yields $P_1 \cong \Phi[\{.5 - \log(.5/.3)/\log(1000/400)\}/.03] \cong \Phi(-1.92) \cong .027$ for normal theory. If this were a marketing application we would assign a .027 value to the probability that the consumer would choose oil heat. If this were a decision analysis application, we would recommend that the decision maker choose gas heat. We would also ask him if he were comfortable with a 2.7 percent probability of making a "wrong" decision.

In Theorem 2, we obtained the same estimators of the true value (mean value) of the risk parameter independent of the error theory we used. I.e., the estimator for both normal and exponential error theories is the sample mean of the inverse function. However, the probability of choice depends on the complete distribution of $f(r|\lambda)$. Hence, we can expect that the estimate of $\hat{P}_1$ will depend upon the error theory that is assumed. In the above example, the sample standard deviation, $\hat{\sigma} = .03$, is much smaller than the sample mean, $\hat{\mu} = .5$. Normal error theory captures this phenomena. However, exponential error theory assumes that the mean of $f(r|\lambda)$ equals the standard deviation for $r_o = 0$. Hence, we would expect exponential error theory to be less certain in identifying gas heat as the maximum utility alternative. Using Theorem 3, we compute $\hat{P}_1 = .33$ for exponential error theory implying that there is less than a 70% chance that gas heat is the best alternative.

This example demonstrates that the choice of an error theory can strongly affect predictions of choice. Since the appropriate distribution for $f(r|\lambda)$ is an empirical question, we caution the reader to examine a histogram of the raw data before selecting an error theory for use in Theorem 3. For example, a histogram (across individuals) of Fishburn and Kochenberger's data [18, p. 511] suggests that the risk parameters for constant proportional risk averse utility functions are normally distributed and the risk parameters for constant absolute risk averse utility functions are exponentially distributed.

Multiple random draws. In Theorem 3 the individual uses the same utility function to evaluate all alternatives. The uncertainty is due to our lack of knowledge about his utility function. The other viewpoint is that errors are involved every time he evaluates an alternative. In this case the appropriate model is a series of independent draws from $f(r|\hat{\lambda})$ for every alternative.

Again, if r were known with certainty, he would select the alternative to maximize $\int u(x,r)h_j(x)dx$. However, since r_j is now an independent draw for every alternative, the probability that alternative 1 is chosen is:

$$P_1 = \text{Prob}[\int u(x,r)h_1(x)dx \geq \int u(x,r)h_j(x)dx \text{ for } j = 1,2,\ldots J] \qquad (11)$$

where $u(x,r)$ is a random variable for every x in the range (x_0,x_*). Since the effect on $u(x,r)$ of the randomness on r is obtained directly from equation 4, the probability in equation 11 is computable, in principal, but like equation 10, numerical integration may be necessary.

Fortunately, for the two most used functional forms, the distribution of $u(x,r)$ can be obtained analytically.

Theorem 4: Suppose we have obtained maximum likelihood estimates for the risk parameters of the constant proportional risk averse utility function (form 2, table 1) and the infinite range ($0\leq x\leq\infty$) constant absolute risk averse utility function (form 1, table 1) by using result 2. Then, if errors are modeled as NORMAL, the utility functions have lognormal distributions. In particular,

$$u(x,r) \sim \Lambda(-k\hat{\mu},k^2\hat{\sigma}^2) \text{ for constant proportional risk aversion}$$

$$1 - u(x,r) \sim \Lambda(-x\hat{\mu},x^2\hat{\sigma}^2) \text{ for constant absolute risk aversion}$$

where $k = \log[(x_* - x_0)/(x - x_0)]$

$\Lambda(a,b)$ = *a lognormal distribution with parameters* a, b.

If errors are modeled as EXPONENTIAL, the utility functions have Beta distributions. In particular,

$u(x,r) \sim \text{Beta}(1/\hat{\lambda}k, 1)$ *for constant proportional risk aversion*

$u(x,r) \sim \text{Beta}(1, 1/\hat{\lambda}x)$ *for constant absolute risk aversion*

where Beta(c, d) = *a Beta distribution with parameters c, d.*

Proof. We begin with Normal error theory. By definition, if z is a normal random variable with mean, μ, and variance, σ^2, and if z=logy, then y is a lognormal random variable with parameters μ and σ^2, designated $y \sim \Lambda(\mu,\sigma^2)$. See Aitchison and Brown [1]. For constant proportional risk aversion, $r(x,u) = -k^{-1}\log u$ or $\log u = -kr(x,u)$. If $r(x,u) \sim N(\mu,\sigma^2)$, then $-kr(x,u) \sim N(-k\mu,k^2\sigma^2)$ which is our first result. For constant absolute risk aversion $\log(1-u) = -xr(x,y)$ yielding the second result.

Now assume exponential error with $r_0 = 0$. For constant proportional risk aversion $f(r|\hat{\lambda}) = \hat{\lambda}^{-1}\exp(r/\hat{\lambda})$, $r(x,u) = -k^{-1}\log u$, and $|dr/du| = 1/ku$ Substituting in equation 4 yields $g(u|\hat{\lambda}) = (\hat{\lambda}k)^{-1} u^{(1/\hat{\lambda}k)-1}$ which we recognize as a Beta distribution which parameters $(1/\hat{\lambda}k)$ and 1. (A Beta distribution is proportional to $u^{c-1}(1-u)^{d-1}$.) For constant absolute risk aversion $r(x,u) = -(1/x)\log(1-u)$ and $|dr/du| = 1/[x(1-u)]$. Substituting in equation 4 yields $g(u|\hat{\lambda}) = (\hat{\lambda}x)^{-1}(1-u)^{(1/\hat{\lambda}x)-1}$ which we recognize as a Beta distribution with parameters 1 and $(1/\hat{\lambda}x)$.

Theorem 4 is very useful for practical applications. If we are evaluating riskless alternatives, then equation 11 becomes a quantal choice problem similar to logit or probit analysis (McFadden [44]) except we use lognormal or Beta distributions rather than the Weibull and normal distribution used in logit and probit analysis respectively. For example, Boyd and Mellman [5] estimated a quantal choice model for automobile choice in which the distribution of the utility function was lognormal mixture of Weibull distributions. Since details of quantal choice models are discussed elsewhere e.g., McFadden [44], we do not discuss them here. For some discussion of numerical techniques see Daganzo [9].

Alternatively, if we are modeling risky alternatives, equation 11 suggests that $h_j(x)$ be chosen from an appropriate family of distributions. That is, we as modelers choose $h_j(x)$ such that the random variable, $\int h_j(x)u(x,r)dx$, is analytic or tabled. See Proposition 1. For further discussions on selecting distributions see DeGroot [10] or Keeney and Raiffa [38]. The appropriate selection of $h_j(x)$ again reduces equation 11 to a quantal choice problem. We do not imply that equation 11 is always an easy computation. It is not. But an extensive literature, reviewed by McFadden [44], exists on how to use equation 11 once the distribution of $\int h_j(x)u(x,r)dx$ is specified. The contribution of Theorem 4 is to specify the distribution for $u(x,r)$.

We close this section with a proposition that illustrates how Theorem 4 and equation 11 are used for a constant proportional risk averse utility function when the choice problem consists of two dichotomous alternatives. The proof is in the appendix.

Proposition 1. Consider the binary choice problem in Figure 3 and assume that the utility function is a constant proportional risk averse utility function. Assume further that the risk parameter is independently drawn from $f(r|\lambda)$ *for each alternative, then the probability that alternative 1 is chosen is given by:*

$$\hat{P}_1 = \Phi\{[\hat{\mu}(k_2-k_1) - \log(\beta/\alpha)]/\eta\hat{\sigma}\} \qquad \text{NORMAL ERROR THEORY}$$

where $\eta = \sqrt{k_1^2+k_2^2}$ and

$$k_i = \log[(x_*-x_0)/(x_i-x_0)]$$

$$\hat{P}_1 = [k_2/(k_1+k_2)][\beta/\alpha]^{-(1/\hat{\lambda}k_2)} \qquad \text{EXPONENTIAL ERROR THEORY}$$

Summary

This completes our discussion of single parameter, uniattributed utility functions. Together our four theorems provide a foundation of an error theory. Theorem 1 states that our estimators apply to the question formats used in marketing and in decision analysis. Theorem 2 provides maximum likelihood estimators (MLE) for both normal and exponential error theories. By using MLE's we make efficient use of redundant questionning to obtain a "best guess" for the utility function.

Theorems 3 and 4 then use these MLE's to estimate choice probabilities. Theorem 3 provides analytic estimates of $\hat{P}_1$ for the choice problem in Figure 3. Since this problem captures the essense of risky choice, it should approximate many decision problems. Furthermore, the results are readily extended to other choice problems. The more general result, Theorem 4, does not compute $\hat{P}_1$ analytically, but transforms the von Neumann-Morgenstern problem so that it is complementary to the well-studied quantal choice problem. The importance of Theorem 4 is that it characterizes analytically the distributions of the utility functions once we have the MLE estimates of the utility parameters. Equation 11 plus Theorem 4 provide a means to compute $\hat{P}_1$ for the more general problem. This is illustrated with proposition 1.

Throughout our development we have tried to provide the analyst with flexibility by (1) specifying two error theories, (2) deriving estimators and choice probabilities for the general case, and (3) deriving practical analytic results for the two most widely used functional forms.

To select among error theories, we suggest that the analyst obtain $r(x_i,p_i)$ for sufficiently large I and plot its histogram. If the histogram is symmetric and unimodel, normal theory is likely to be a good model; if the histogram is unimodel and skewed with $\hat{\sigma} \cong \hat{\mu}$, then exponential theory is likely to be a good model.

To select among functional forms we suggest either (1) qualitative judgment and questionning or (2) statistical fit. The former method is discussed in Keeney and Raiffa [38, p. 191]. For example, if the individual suggests that his risk aversion is constant over the range of x, form 1 is a good model; whereas if risk aversion decreases as assets increase, form 2 is the better model. The latter method can depend on minimum variance of $r(x_i, p_i)$ or another statistical criterion.

4. MULTIPLE PARAMETER UNIATTRIBUTED UTILITY FUNCTIONS

While the class of single parameter utility functions is quite flexible and can accommodate a wide range of interesting problems, occasionally a researcher may wish to estimate the parameters of a utility function that is more complex. For example, Keeney and Raiffa [38,p. 209] report that a computer program which has been used at the Harvard Business School since 1966 is based on the decreasingly risk averse three parameter function in Table 1 (form 5). In general, the computation of choice probabilities is not analytically tractable for multiple parameter functions, but equations 10 and 11 still apply for numerical solutions. Since a researcher choosing a multiple parameter utility function may be willing to sacrifice analytic simplicity for greater flexibility, we provide a means to estimate the parameters of the utility function recognizing numerical integration may be necessary for choice probabilities.

We provide two methods. The first method requires clustered questioning, but provides maximum likelihood estimates. The second method relaxes the clustering requirement, but provides a regression approximation.

Clustered Questions

Suppose a marketing scientist wishes to estimate a population distribution for the three parameter function (form 5) in Table 5. One procedure might be to ask each consumer three questions and use that information to characterize the population by a multivariate distribution on r_1, r_2, r_3.

Alternatively, a decision analyst may ask I clusters of three questions in order to estimate the decision maker's utility function. (In general, a cluster will contain K questions if K parameters are to be estimated.)

Let (x_{ki}, p_{ki}) be the certainty equivalent and lottery probability associated with the ith cluster and kth question within the cluster. By theorem 1, we can specify either x_{ki} or p_{ki} while the individual responds with the other value. Let $\underline{x}_i = (x_{1i}, x_{2i},\ldots x_{Ki})$ and $\underline{p}_i = (p_{1i}, p_{2i},\ldots,p_{Ki})$. If a vector-valued function, $\underline{r}(\underline{x}_i,\underline{p}_i)$, exists mapping the vectors $\underline{x}_i$ and $\underline{p}_i$ onto the range of the K unknown parameters of the utility function, and if we assume that errors cause $\underline{r}(\underline{x}_i,\underline{p}_i)$ to be distributed with a multivariate normal distribution with mean, $\underline{\mu}$, and covariance matrix, Σ, then the maximum likelihood estimators, $\hat{\underline{\mu}}$ and $\hat{\Gamma}$, are simply the multivariate extensions of the univariate estimators in theorem 2. That is,

$$\hat{\underline{\mu}} = (1/I)\ \Sigma_i \underline{r}(\underline{x}_i,\underline{p}_i) \tag{12}$$

$$\hat{\Gamma} = (1/I)\ \Sigma_i [\underline{r}(\underline{x}_i,\underline{p}_i) - \hat{\underline{\mu}}][\underline{r}(\underline{x}_i,\underline{p}_i) - \hat{\underline{\mu}})]^T \tag{13}$$

For a formal proof see Giri [20, chapter 15]. As before we can construct confidence regions with the multivariate extension of a t-test. For example, the appropriate statistic for $\hat{\underline{\mu}}$ is Hotelling's T^2 statistic (Giri[]0, chapter 7] and Green [24, p. 257]).

Similar results apply for exponential error.

Independent Questions

If, for whatever reasons, the analyst feels that clustered questions are not appropriate for his situation, he may wish to ask K x I independent questions. In this case, without further specifying the interrelationships of the question formats we can not obtain maximum likelihood estimators. However, we can obtain a practical regression approximation.

Following Pratt [42] and Keeney and Raiffa [38, p. 160] define a risk premium, π_i, as the amount by which the certainty equivalent, x_i, exceeds the expected value of the lottery, $\bar{x}_i$. For the measurement in Figure 1:

$$\pi_i(x_i,p_i) = x_i - p_i x_* - (1-p_i)x_0 \qquad (14)$$

Keeney and Raiffa [38, p. 161] then consider variation about the expected value of the lottery and show by Taylor's series expansion that the local risk averion, $R(x,\underline{r})$, is approximately proportional to π_i. ($R(x,\underline{r})$ is defined by equation (1) where we have added the unknown parameters, $\underline{r}$, to the notation.) In particular, Keeney and Raiffa show

$$\pi_i(x_i,p_i) = (1/2)v_i^2 \, R(x_i,\underline{r}) + \varepsilon \qquad (15)$$

where v_i^2 is the variance of the lottery, $v_i^2 = (1/2)(x_*-\bar{x}_i)^2 + (1/2)(x_0-\bar{x}_i)^2$, and ε indicates higher order terms that are assumed to be negligible. Rearranging terms yields:

$$R^0(x_i,p_i) = R(x_i,\underline{r}) + \tilde{e}_i \qquad (16)$$

where $R^0(x_i,p_i) = 4(x_i - p_i x_* - (1-p_i)x_0)/[(x_*-\bar{x}_i)^2 + (x_0-\bar{x}_i)^2]$ is a function of known data because $\bar{x}_i = p_i x_* + (1-p_i)x_0$. Note that we have incorporated the Taylor's series error, ε, in the measurement error, $\tilde{e}_i$.

Equation (16) is now in the form of a regression equation. If $R(x_i,\underline{r})$ is linear in its parameters, ordinary least squares regression applies. Alternatively, a researcher may use non-linear techniques for non-linear $R(x_i,\underline{r})$. Once the parameters of $R(x_i,\underline{r})$ are estimated, we can recover $u(x_i,\underline{r})$ from equation (1) by integration since $u(x,\underline{r}) = f_1 \int \exp[-\int R(x,\underline{r})dx]dx + f_2$ where f_1 and f_2 are constants chosen to scale the utility function.

For example, we might wish to consider utility functions which combine constant absolute and proportional risk aversion (forms 1 and 2 in Table 1). In this case, $R(x_i,\underline{r}) = r_1+r_2(x_i-x_o)^{-1}$ is linear in the unknown parameters. Since equation (16) does not require an inverse function, we can allow x_* and x_o to vary across measurements, i.

5. MULTIATTRIBUTED UTILITY FUNCTIONS

The theory in section 3 and 4 provides us with an error theory to estimate and use uniattributed utility functions. For many applications, such as decisions among alternative financial investments, a uniattributed theory suffices. However, there are many applications in both decision analysis and marketing where it is necessary to model prescriptively or descriptively decisions involving multiple attributes, each of which is risky. For example, the decision to buy a home heating system might involve reliability, safety, and cleanliness as well as annual dollar savings. (Choffray and Lilien [7] illustrate empirically a multiattributed preference problem for solar air-conditioning.)

The most general problem is to estimate the parameters, $\underline{w}$, of an M-valued function, $U(\underline{x},\underline{w})$, mapping the levels of the attributes, $\underline{x}$, onto a utility measure. In this section we let $\underline{x}_{\ell i} = (x_{\ell i1}, x_{\ell i2},\ldots x_{\ell iM})$ be the levels of the M attributes for the certainty equivalent in the ℓth question in the ith cluster. In assessing $U(\underline{x},\underline{w})$ we specify either (1) all of $\underline{x}_{\ell i}$ or (2) $p_{\ell i}$ and all but one of the $\underline{x}_{\ell i}$. A multiattributed extension of theorem 1 lets us choose $p_{\ell i}$ or any of the $\underline{x}_{\ell i}$ as the value specified by the individual.

Estimation. If we ask our questions in I clusters of L questions and if a computable vector-valued inverse function, $\underline{W}(\underline{\underline{x}}_i, \underline{p}_i)$, exists mapping the question set onto the unknown parameters, then the multiattributed problem is isomorphic to the multiple parameter uniattributed problem. ($\underline{\underline{x}}_i$ is the matrix with rows $\underline{x}_{\ell i}$ and $\underline{p}_i$ is the vector with elements $p_{\ell i}$.) Equations 12 and 13 can be used to estimate the mean and covariance of a multivariate normal distribution on $\underline{w}$. Confidence intervals are computed with Hotelling's T^2 statistic.

Probability of Choice. Choice probabilities can be obtained with the appropriate modification of equations 10 and 11, although numerical techniques may be necessary. For example, one might use equation 10 by sampling from the multivariate normal distribution then using the sampled $\tilde{\underline{w}}$ to compute the expected utility of each option. Choice probabilities are then the percent of times an alternative is chosen in, say 1000, draws. This computation method is similar to methods used in probit analysis, e.g., Dagonzo [9], and has proven feasible in that context.

Estimation Example

We illustrate the multiattribute extension with a hypothetical home heating system example. Suppose that besides annual savings, x_1, the individual is concerned with reliability as measured by the probability, x_2, that no repair will be needed each year. We suppose that the individual plans to purchase a service contract (a form of insurance policy) such that only negative effect of a repair is inconvenience (not dollar cost). We wish to model the decision maker's preferences by a constant proportional risk averse, multilinear utility function. I.e.:

$$U(\underline{x},\underline{w}) = w_3 u_1(x_1) + w_4 u_2(x_2) + (1-w_3-w_4)u_1(x_1)u_2(x_2) \tag{17}$$

$$u_1(x_1) = [(x_1 - 200)/1000]^{w_1} \;;\; u_2(x_2) = x_2^{w_2}$$

We estimate the four unknown parameters, $\underline{w} = (w_1, w_2, w_3, w_4)$ by asking the lottery questions shown in Table 3. In each question, the decision maker is asked to give a probability, $p_{\ell i}$, such that he is indifferent between a certainty equivalent, $(x_{\ell i1}, x_{\ell i2})$ and a lottery where the system is described by $(x''_{\ell i1}, x''_{\ell i2})$ with probability $p_{\ell i}$, and by $(x'_{\ell i1}, x'_{\ell i2})$ with probability, $(1-p_{\ell i})$. In other words, the standard lottery shown in figure 5:

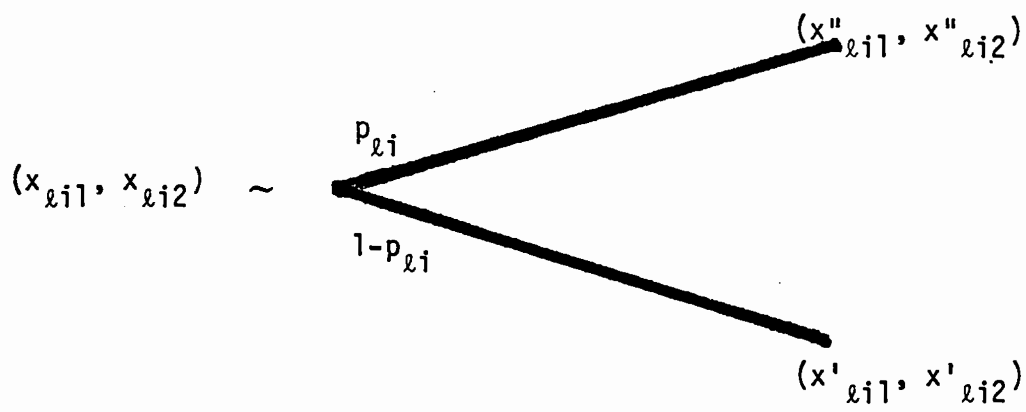

Figure 5: Schematic of Multivariate Lottery

The reader will note that we have constructed the questions in Table 3 for easy computation of the inverse function, $\underline{w}(\underline{\underline{x}}_i, \underline{p}_i)$.

$$w_{1i} = \log(p_{1i})/\log[(x_{1i1} - 200)/1000] \tag{18}$$

$$w_{2i} = \log(p_{2i})/\log(x_{2i2})$$

$$w_{3i} = p_{3i}/p_{1i}$$

$$w_{4i} = p_{4i}/p_{2i}$$

TABLE 3

EXAMPLE ASSESSMENT FOR THE COST AND RELIABILITY

OF A HOME HEATING SYSTEM

Certainty Equivalent				"Win"		"Loss"		Probability	Parameters			
i	ℓ	$x_{\ell i1}$	$x_{\ell i2}$	$x''_{\ell i1}$	$x''_{\ell i2}$	$x'_{\ell i1}$	$x'_{\ell i2}$	$p_{\ell i}$	w_{1i}	w_{2i}	w_{3i}	w_{4i}
1	1	400	.20	1200	.20	200	.20	.45	.50			
	2	400	.20	400	1.00	400	.00	.60		.32		
	3	400	.00	1200	1.00	200	.00	.20			.44	
	4	200	.20	1200	1.00	200	.00	.50				.83
2	1	600	.40	1200	.40	200	.40	.65	.47			
	2	600	.40	600	1.00	600	.00	.75		.31		
	3	600	.00	1200	1.00	200	.00	.25			.38	
	4	200	.40	1200	1.00	200	.00	.60				.80
3	1	800	.60	1200	.60	200	.60	.75	.56			
	2	800	.60	800	1.00	800	.00	.80		.44		
	3	800	.00	1200	1.00	200	.00	.30			.40	
	4	200	.60	1200	1.00	200	.00	.65				.81
4	1	1000	.80	1200	.80	200	.80	.90	.47			
	2	1000	.80	1000	1.00	1000	.00	.95		.23		
	3	1000	.00	1200	1.00	200	.00	.35			.39	
	4	200	.80	1200	1.00	200	.00	.75				.79
								$\hat{\mu}$ =	.50	.33	.40	.81

x_1 = savings in dollars

x_2 = probability that no repair is needed in a given year

This simplicity is for ease of exposition. Tradeoff questions as well as lotteries can be used and the inverse function can vary with i as long as it is computable for all i. Even with our simplification, the sixteen questions provide the decision maker with a variety of questions. The "answers", $p_{\ell i}$, to the lottery questions were "constructed" by assuming $\underline{w}$ = (.50, .33, .40, .80) and rounding off to the nearest .05.

Examination of Table 3 reveals that the estimated parameters, $\underline{w}^T$, recover the "known" values quite well. The covariance matrix, $\hat{\Gamma}$, and the corresponding correlation matrix, $\hat{C}$, can be readily computed with equations 12 and 13.

$$\hat{\Gamma} = \begin{bmatrix} .0054 & .0102 & .0009 & .0009 \\ & .0226 & .0010 & .0020 \\ & & .0021 & .0010 \\ & & & .0009 \end{bmatrix} \qquad \hat{C} = \begin{bmatrix} 1.0 & .93 & .27 & .41 \\ & 1.0 & .15 & .44 \\ & & 1.0 & .74 \\ & & & 1.0 \end{bmatrix}$$

We note that the high off-diagonal elements in $\hat{C}$ are partially due to the small sample size, I=4, and partially due to structural intercorrelation in equation 18. E.g., p_{1i} appears in the equations for both w_1 and w_3. Such correlations can be avoided with larger sample sizes and judicious choice of question format.

6. CONCLUSIONS AND FUTURE DIRECTIONS

This completes our development of an initial error theory for von Neumann-Morgenstern utility theory. Our emphasis is on uniattributed single parameter functions since they illustrate the unique advantage of vN-M theory. Our results are practical and flexible. The analyst can choose among question formats, error assumptions, and functional forms. Our analytic results make the theory usable. Our more general results are feasible, if difficult numerically, and hence encourage further development of analytic theory and computer software.

Many interesting problems remain including (1) practical algorithms for quantal choice analysis with lognormal and Beta distributions, (2) analytic distributions for $\int h_j(x)u(x,r)dx$, and (3) analytic solutions for choice probabilities with selected multiattribute forms.

Assessment has proven feasible in both decision analysis and marketing. This experience suggests the error theory is based on practical question formats. Furthermore, once the data is obtained, our illustrative examples show that the estimators can recover known utility functions. However, many empirical questions remain which will ultimately be resolved through practice. Accumulated experience can suggest conditions when normal (or exponential) is the better error assumption and indicate which functional form best represents decision makers and consumers.

APPENDIX

Lemma 1: Assume $\beta > \alpha$ *and* $x_1 > x_2$*, then the equation,* $\beta \exp(-r_c x_2) - \alpha \exp(-r_c x_1) = \beta - \alpha$*, has at most one solution for* $r_c > 0$.

Proof: First, rewrite the equation in functional form:

$$E(r) = \alpha[1-\exp(-rx_1)] - \beta[1-\exp(-rx_2)] \quad \text{(A1)}$$

recognizing $x_1 > x_2$ and $\beta > \alpha$. Alternative 1 will be chosen whenever $E(r) \geq 0$.

By a Taylor expansion $E(r) \cong \alpha x_1 - \beta x_2$ as $r \to 0$. Let $E(0) = \lim_{r_c \to 0} E(r_c)$ and let $E(\infty) = \lim_{r_c \to \infty} E(r_c)$. Then $E(0) > 0$ if $\alpha x_1 > \beta x_2$ and $E(0) \leq 0$ if $\alpha x_1 \leq \beta x_2$. By direct substitution $E(\infty) = \alpha - \beta < 0$ since $\alpha < \beta$.

Now differentiate $E(r)$ yielding

$$E'(r) = dE(r)/dr = \alpha x_1 \exp(-rx_1) - \beta x_2 \exp(-rx_2)$$

Setting the derivative equal to zero yields $r^* = (\log \alpha x_1 - \log \beta x_2)/(x_1 - x_2)$. Since $x_1 > x_2$, $r^* > 0$ iff $\alpha x_1 > \beta x_2$. Furthermore $E'(0) = \alpha x_1 - \beta x_2$ thus $E'(0) > 0$ iff $\alpha x_1 > \beta x_2$.

Assume $\alpha x_1 < \beta x_2$ then $E(0) < 0$, $E(\infty) < 0$, and $E(r)$ is monotonic in the range $(0,\infty)$. Thus there is no solution to A1 for $r_c > 0$. If $\alpha x_1 = \beta x_2$ the only solution for $r \geq 0$ is $r_c = 0$.

Assume $\alpha x_1 > \beta x_2$, then $E(0) > 0$, $E'(0) > 0$, and $r^* > 0$. Thus $E(r) > 0$ for $r \leq r^*$. Now $E(r^*) > 0$, $E(\infty) < 0$, and $E(r)$ is monotonic in the range (r^*,∞). Thus, there is exactly one solution to A1 for $r_c > 0$ and it occurs in the range (r^*,∞). Note that we have also proven that $E(r) > 0$ for $r \in (0,r_c)$ and $E(r) < 0$ for $r \in (r_c,\infty)$, thus alternative 1 will only be chosen in the range $(0,r_c)$.

PROOF TO PROPOSITION 1

Alternative 1 will be chosen if $\alpha u(x_1,r_1) > \beta u(x_2,r_2)$. Consider first normal error theory. Rearranging terms this condition becomes $\log u(x_1,r_1) - \log u(x_2,r_2) > \log(\beta/\alpha)$. Using theorem 4, the left hand side of the inequality is distributed as $N[\hat{\mu}(k_2-k_1), \hat{\sigma}^2(k_1{}^2+k_2{}^2)]$ and the result follows.

Now consider exponential error theory. Again rearranging terms indicates that alternative 1 will be chosen if $u(x_1,r_1)/u(x_2,r_2) > \beta/\alpha$. Let $u_i = u(x_i,r_i)$ then by Theorem 4 and the assumption of independent draws $g(u_1,u_2)$ is given by:

$$g(u_1,u_2) = (\hat{\lambda}^2 k_1 k_2)^{-1}\,(u_1)^{(1/\hat{\lambda}k)-1}\,(u_2)^{(1/\hat{\lambda}k_2)-1}$$

Define $z=u_1/u_2$ and $t=u_2$ then the p.d.f. of z and t is obtained using a Jacobian transformation:

$$f_{zt}(z,t) = q_1 q_2 (z)^{q_1-1}\,(t)^{q_1+q_2-1}$$

$$\text{where } q_i = (1/\hat{\lambda}k_i)$$

Integrating out t yields the marginal distribution for z:

$$f_z(z)=\begin{cases} \dfrac{q_1 q_2}{q_1+q_2}\, z^{q_1-1} & 0 \le z \le 1 \\[2ex] \dfrac{q_1 q_2}{q_1+q_2}\, z^{-q_2-1} & z \ge 1 \end{cases}$$

Since $(\beta/\alpha) > 1$, $\hat{P}_1 = \text{Prob}[z > \beta/\alpha] = \int_{\beta/\alpha}^{\infty} f_z(z)dz = [q_1/(q_1+q_2)](\beta/\alpha)^{-q_2}$. Finally, substituting $q_i = (1/\hat{\lambda}k_i)$ into the above expression yields the result.

REFERENCES

1. J. Aitchison and J.A.C. Brown, The Lognormal Distribution with Special Reference to its Uses in Economics (England: Cambridge University Press, 1969).

2. Barnes, J.D. and J.E. Reinmuth, "Comparing Inputed and Actual Utility Functions in a Competitive Bidding Setting," Decision Sciences, Vol. 7, No. 4 (October 1976), pp. 801-812.

3. Bell, D.E., "A Decision Analysis for Objectives of a Forest Pest Problem," R.R.-75-43, International Institute for Applied Systems Analysis, Laxenburg, Austria.

4. Bodily, S.E., "The Utilization of Frozen Red Cells in Blood Banking Systems: A Decision Theoretic Approach," Technical Report No. 94, Operations Research Center, M.I.T., Cambridge, MA, 1974.

5. J.H. Boyd and R.E. Mellman, "The Effect of Fuel Economy Standards on the U.S. Automobile Market: A Hedonic Demand Analysis," Transportation Research, Vol. 14A, No. 5-6 (October-December 1980), pp. 367-378

6. Brown, R.V. and C. Peterson, "An Analysis of Alternative Mideastern Oil Agreements," Technical Report, Decisions and Designs, Inc., McLean, Virginia, 1975.

7. J.M. Choffray and G.L. Lilien, "The Market for Solar Cooling: Perceptions, Response, and Strategy Implications," Studies in Management Science, Vol. 10, 1978, pp. 209-226.

8. J.M. Choffray and G.L. Lilien, Market Planning for New Industrial Products, New York, John Wiley & Sons, 1980.

9. Daganzo, C., Multinomial Probit: The Theory and its Applications to Demand Forecasting (New York: Academic Press, 1979).

10. Degroot, M.H., Optimal Statistical Decisions, (McGraw Hill, New York), 1970.

11. Eliashberg, Jehoshua, "Consumer Preference Judgments: An Exposition with Empirical Applications," Management Science, Vol. 26, No.1 (January 1980), pp. 60-77.

12. Eliashberg, J. and Winkler, Robert L., "The Role of Attitude Toward Risk in Strictly Competitive Decision-Making Situations," Management Science, Vol. 24, No. 12, August 1978, pp. 1231-1241.

13. Eliashberg, J. and Winkler, R.L., "Risk Sharing and Group Decision-Making," Management Science, Vol. 27, No. 11, November 1981.

14. Ellis, H.M. and R.L. Keeney, "A Rational Approach for Government Decisions Concerning Air Pollution," in A.W. Drake, R.L. Keeney, and P.M. Morse, eds., Analysis of Public Systems, (Cambridge, MA: MIT Press, 1972).

15. Farquhar, P.H., "A Fractional Hypercube Decomposition Theory for Multi-attribute Utility Functions," Operations Research, Vol. 23, No. 5 (September-October 1975), pp. 941-967.

16. Farquhar, P.H., "A Survey of Multiattribute Utility Theory and Applications," Studies in Management Science, Vol. 6 (1977), pp. 59-89.

17. Fishburn, P.C., "von Neumann-Morgenstern Utility Functions on Two Attributes," Operations Research, Vol. 22, No. 1 (January-February 1974), pp. 35-45.

18. Fishburn, P.C., and G.A. Kochenberger, "Two-Piece von Neumann-Morgenstern Utility Functions," Decision Sciences, Vol. 10 (1979), pp. 503-518.

19. Friedman, M. and L.J. Savage, "The Expected-Utility Hypothesis and the Measureability of Utility," Journal of Political Economy, Vol. 60 (1952), pp. 463-474.

20. Giri, N.C., Multivariate Statistical Inference, (Academic Press, NY, 1977).

21. Goodman, B., M. Saltzman, W. Edwards and D.H. Krantz, "Prediction of Bids for Two-Outcome Gambles in a Casino Setting," Organizational, Behavioral and Human Performance, Vol. 24 (1979), pp. 382-399.

22. Grayson, C.J., Decisions under Uncertainty: Drilling Decisions by Oil and Gas Operators (Cambridge, MA: Graduate School of Business, Harvard University, 1960).

23. Green, P.E., "Risk Attitudes and Chemical Investment Decisions," Chemical Engineering Progress, Vol. 59, No. 1 (January 1963), pp. 35-40.

24. Green, P. E., Analyzing Multivariate Data (Hinsdale, IL: The Dryden Press, 1978).

25. Green, P.E. and V. Srinivasan, "Conjoint Analysis in Consumer Research: Issues and Outlook," Journal of Consumer Research, Vol. 5, No. 2 (September 1978), pp. 103-123.

26. Gros, J.G., "A Paretian Environmental Approach to Power Plant Siting in New England," Doctoral Dissertation, Harvard University, Cambridge, MA, 1974.

27. Halter, A.N. and G.W. Dean, Decisions Under Uncertainty (Cincinnati, OH: Southwestern Publishing Co., 1971).

28. Hauser, J.R., "Consumer Preference Axioms: Behavioral Postulates for Describing and Predicting Stochastic Choice," Management Science, Vol. 24, No. 13 (September 1978), pp. 1331-1341.

29. Hauser, J.R. and S.M. Shugan, "Intensity Measures of Consumer Preference," Operations Research, Vol. 20, No. 2 (March-April 1980), pp. 278-320.

30. Hauser, J.R. and G.L. Urban, "A Normative Methodology for Modeling Consumer Response to Innovation," Operations Research, Vol. 25, No. 4 (July-August 1975), pp. 579-619.

31. Hauser, J.R. and G.L. Urban, "Direct Assessment of Consumer Utility Functions: von Neumann-Morgenstern Theory Applied to Marketing," Journal of Consumer Research, Vol. 5 (March 1979), pp. 251-262.

32. Herstein, I.N. and J. Milnor, "An Axiomatic Approach to Measureable Utility," Econometrica, Vol. 21 (1953), pp. 291-297.

33. Jensēn, N.E., "An Introduction to Bernoullian Utility Theory. I. Utility Functions," Swedish Journal of Economics, Vol. 69 (1967), pp. 163-183.

34. Kahneman, D. and A. Tversky, "Prospect Theory: An Analysis of Decision Under Risk," Econometrica, Vol. 47, No. 2 (March 1979), pp. 263-291.

35. Keeney, R.L., "Utility Functions for Multiattributed Consequences," Management Science, 18 (1972), pp. 276-287.

36. Keeney, R.L., "A Utility Function for the Response Times of Engines and Ladders to Fires," Urban Analysis, Vol. 1 (1973), pp. 209-222.

37. Keeney, R.L., "Multiplicative Utility Functions," Operations Research, Vol. 22, No. 1 (January 1974), pp. 22-33.

38. Keeney, R.L., and H. Raiffa, Decisions with Multiple Objectives: Preferences and Value Tradeoffs (New York: John Wiley & Sons, 1976).

39. Krantz, D.H., R.D. Luce, P. Suppes, and A. Tversky, Foundations of Measurement (New York: Academic Press, 1971).

40. Krischer, J.P., "An Analysis of Patient Management Decisions as Applied to Cleft Palate," TR-12-74, Center for Research in Computing Technology, Harvard University, Cambridge, MA, 1974.

41. Lee, W., Decision Theory and Human Behavior (New York: John Wiley & Sons, 1971).

42. Luce, R.D. and P. Suppes, "Preference, Utility, and Subjective Probability," in R.D. Luce, R.R. Bush, and E. Galanter (eds.) Handbook of Mathematical Psychology, Vol. 3 (New York: John Wiley & Sons, 1965), pp. 249-410.

43. Marschak, J., "Rational Behavior, Uncertain Prospects, and Measureable Utility," Econometrica, Vol. 18 (1950), pp. 111-141.

44. McFadden, D., "Quantal Choice Analysis: A Survey," Annals of Economics and Social Measurement, (1976), pp. 363-390.

45. Meyer, R.F. and J.W. Pratt, "The Consistent Assessment and Fairing of Preference Functions," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-4, No. 3 (September 1968), pp. 270-278.

46. Mood, A.M. and F.A. Graybill, Introduction to the Theory of Statistics (New York: McGraw-Hill Book Co., 1963).

47. O'Connor, M.F., "The Application of Multiattribute Scaling Procedures to the Development of Indices of Water Quality," Report 7339, Center for Mathematical Studies in Business and Economics, University of Chicago, Chicago, IL, 1973.

48. Pratt, J.W., "Risk Aversion in the Small and the Large," Econometrica, Vol. 32 (1964), pp. 122-136.

49. Richard, S.F., "Risk Aversion, Utility Independence, and Separable Utility Functions," Management Science, Vol. 22 (1975), pp. 12-21.

50. Swalm, R.D., "Utility Theory - Insights into Risk Taking," Harvard Business Review, Vol. 47 (November-December 1966), pp. 123-136.

51. Thurstone, L., "A Law of Comparative Judgment," Psychological Review, 34 (1927), pp. 273-286.

52. Tversky, A., "Additivity, Utility, and Subjective Probability," Journal of Mathematical Psychology, Vol. 4 (1967), pp. 175-201.

53. Urban, G.L., "PERCEPTOR: A Model for Product Positioning," Management Science, Vol. 21, No. 8 (April 1975), pp. 858-871.

54. Urban, Glen L. and John R. Hauser, Design and Marketing of New Products (Englewood Cliffs, N.J.: Prentice-Hall, 1980).

55. von Neumann, J. and O. Morgenstern, Theory of Games and Economic Behavior Princeton, N.J.: Princeton University Press, 1947.

56. Wilkie, W.L. and E.A. Pessemier, "Issues in Marketing's Use of Multi-attribute Attitude Models," Journal of Marketing Research, Vol. 10 (November 1973), pp. 428-441.