Pub King

Risk Modeling in Agriculture: Retrospective and Prospective

Program proceedings for the annual meeting of the Technical Committee of S-232, held March 24-26, 1994, Gulf Shores State Park, Alabama.

Iowa State University
Ames, Iowa 50011-1070
August 1994

THE EFFECTS OF SIMILARITY ON CHOICE AND DECISION EFFORT*

David E. Buschena**

This paper explores risky choice patterns in an experimental setting and develops and tests a model of decision using time as a measure of effort. Patterns of decision time are tested statistically, shown to depend on characteristics of the risky alternatives reflecting the costs and benefits of choice, and are shown to correspond with the predictions of an optimization model.

There are two relevant fields of literature for this paper. The first addresses the effects of the similarity of risky alternatives on choice (Rubinstein 1988; Aipurua et al. 1993; Leland 1990, 1992a, 1992b, and 1992c; Buschena 1992). Similarity models offer an explanation for violations of the Expected Utility model (EU). These models' predictions are motivated through tradeoffs between the costs and benefits of evaluation effort akin to arguments of Bounded Rationality (Simon 1960; March 1978).

The second area of relevant literature uses evaluation time to measure decision effort, finding significant relationships between evaluation effort and the importance of choice. Specifically, these models hold that evaluation time depends significantly on measures of the differences in utility between the alternatives. Work of particular interest for this paper are empirical efforts by Dashiell (1937); Petrusic and Jamieson (1978); Payne, Bettman, and Johnson (1993); and Wilcox (1993a, b).

I. Previous Literature

A. Similarity Effects on Risky Choice

Recent work by a number of researchers (Rubinstein; Azipurua et al.; Leland (1990, 1992a, 1992b, 1992c); and Buschena) have implicitly incorporated the effects of decision costs or bounds on rationality in models of risky choice. These effects were incorporated through measures of the similarity between alternatives. Models explicitly accounting for decision costs or bounds [as called for in Simon (1960) and in March (1978)] have been developed in work by Lipman (1991), Conlisk (1988), Heiner (1988), Encarnaciòn (1987), Ng (1975), and Radner (1975). These similarity models are also related to work by Luce (1956) and by Payne, Bettman, and Johnson (1993). The importance of similarity effects on the occurrence of EU Independence Violations was shown to be significant in Buschena (1992).

---

*David Zilberman, Brian Wright, Mike Roberts, Dan Benjamin, and Thomas Stratmann provided useful comments regarding this study. Special thanks to Barbara Mellers and Shi-Jie Chang for discussions and for the arrangement and development of the experimental framework. Sheila Smith provided excellent technical support.

**Assistant Professor, Department of Agricultural Economics and Economics, Montana State University.

Two models describing the similarity of risky alternatives are addressed here. The predictions of these models form the basis for tests discussed in the empirical section of this paper. To illustrate these similarity models, consider the following risky choice pairs of a type known as prospects. For prospects, one risky alternative gives a single non-zero outcome (x) with a positive probability (p) while the other also gives a single non-zero outcome (y) with positive probability (q).

Rubinstein's ε-difference model of similarity for risky choice postulates intra-dimensional similarity for prospects, given as:[1]

R1. Similarity in the outcomes: xS_Xy iff $|x-y| < \varepsilon$,
R2. Similarity in the probabilities: pS_Pq iff $|p-q| < \varepsilon$.

That is, Rubinstein's ε-difference similarity proposes separable similarity effects for the choices, where absolute differences in the alternatives' dimensions define similarity.

Rubinstein further specifies a choice procedure (*), having four evaluation algorithms depending on the intra-dimensional similarity:

1. If both the outcome and the probability dimensions are similar (xS_Xy and pS_Pq), the choice procedure is unspecified.

2. If the outcome dimension is similar (xS_Xy) but the probability dimension is not ($\neg pS_Pq$), then the probability dimension determines choice and the risky alternative carrying the highest probability is selected.

3. If the probability dimension is similar (pS_Pq) but the outcome dimension is not ($\neg xS_Xy$), then the outcome dimension determines choice and the risky alternative carrying the highest outcome is selected.

4. If neither dimension is similar ($\neg xS_Xy$ and $\neg pS_Pq$), then choice is unspecified.

This paper makes this choice procedure operational by asserting choice follows expected utility under condition 4; choice under condition 1 remains unspecified.

Rubinstein shows that the "*" choice procedure based on the ε-difference similarity allows for intransitive indifference. Azipurua et al. allow for this intransitivity and expand the definition of similarity in an appealing way.

Under Rubinstein's ε-difference similarity measure, the probability dimension is normalized while the outcome dimension is not. Azipurua et al. argue that it is reasonable to model similarity (as it affects individuals' choices) so that the critical probability differences depend on the outcome level for R2:

A1. Similarity in the outcomes: xS_Xy iff $|x-y| < \varepsilon_X$,
A2. Similarity in the probabilities: pS_Pq iff $|p-q| < \varepsilon_{P(X)}$.

The intuitive support for the structure in A2 is that, as the outcomes increase, the decision maker takes a particular probability differences more seriously. This relationship is reproduced in Figure 1 from Azipurua et al., where the critical level of probability similarity decreases with the size of the outcomes.

Tests of the empirical validity of Rubinstein's and Azipurua's definitions of similarity will be carried out in the empirical section. These tests use individual's reported similarity "perceptions" as an appropriate indicator of similarity as it applies to risky choice.

The hypothesized effects of similarity on choice are modeled here through the use of approximation methods. The predictions of these methods may differ in varying degree from complete EU maximization. That is, the use of these approximation methods may bias some pairs' evaluation and subsequent choice relative to EU maximization. Consider choice between prospects (AB) and (CD) from Kahneman and Tversky:

Select between A and B:

A: gives $3000 for certain (probability 1.0) **B: gives $4000 with probability .80, otherwise $0**

Select between C and D:

C: gives $3000 with probability .25, otherwise $0 **D: gives $4000 with probability .20, otherwise $0**

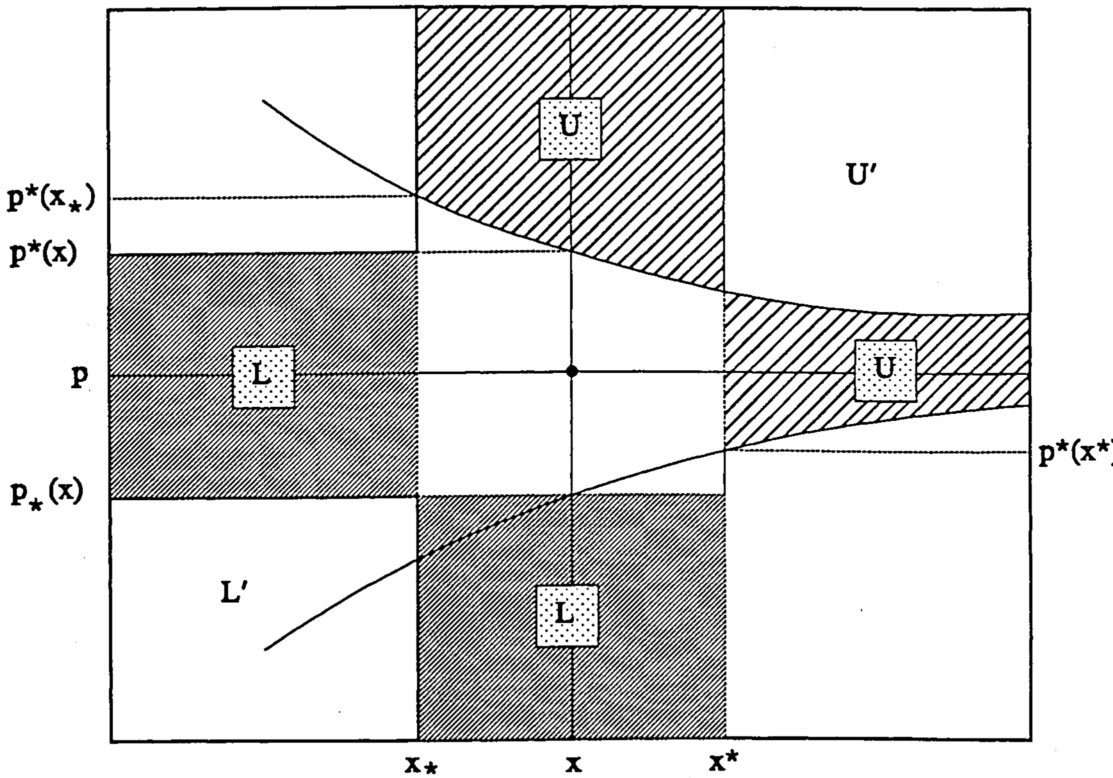
A violation of EU can occur in two ways for choices between pairs (AB) and (CD). If the individual selects A and then D, an EU violation occurs; if the individual selects B and then C, a violation occurs. Note that choice patterns A and C or B and D are consistent with EU.

Suppose that the outcome dimensions are dissimilar as defined in R1 and in A1. Given that a violation of EU occurs, the similarity models of Rubinstein (1988) and Azipurua et al. (1993) predict the selection of A in the first pair but D in the second.[2] The prediction of D over C holds that choice is degenerate, depending only on the outcomes ($3000, $4000) because the probabilities (.25, .20) are similar. Conversely, the probabilities defining A and B (1.0, .8) are sufficiently different that the alternatives are taken to be fully evaluated in step 4 of process "*", here assumed to be evaluated through EU preferences. Leland (1990) proposes a discrete step function approximation for the probabilities and the underlying concave utility representation, $u(z)$ for $z \in \{x, y\}$, over the outcomes. Leland's (1990) model also would predict the choice pattern of (A selected over B) and (D selected over C) given a violation occurs since the probabilities are similar and are likely to be evaluated on the same step level. A more general model of similarity based on moments of the distribution that has the same predictions for choice — selection of the less risky alternative A and selection of the more risky alternative D — was suggested in Buschena and is the basis for tests carried out in this paper.

B. Relationship Between Choice Effort and Similarity

The relationship between choice and decision effort have been the subject of research by psychologists and early efforts by economists. Findings from these studies shed light on the effects of similarity on risky choice for both (a) decision time and (b) patterns of risky choice. Some of the salient works are summarized briefly below. Questions remaining from these studies for the effects of similarity on risky choice further motivate the tests carried out in this paper.

Figure 1



Source: Azipurua et al.

Difficulty of Evaluation. Stone and Schkade (1991) and Edwards (1955, 1965) developed theoretical models of optimal information gathering efforts for general decision making. Dashiell carried out early empirical work on response times, where individuals were asked to cardinally rank their preferences over (but not to select between) pairs of colors. He found that individuals took the longest time in rating color pairs that had very small differences in preference ranks (i.e. similar pairs). This general relationship was supported in later work by Petrusic and Jamieson using questions over geometric shapes and over externally developed verbal quotes, suggesting the general relation (Figure 2) between time allocation in rating tasks and the differences between the relative preference ratings of the alternatives.

Note, however, that the goal for respondents in Dashiell's and in Petrusic and Jamieson's study was presumably to rank the relative desirability of a pair, but not to select their preferred color. This situation might be somewhat different from that for choice over risky alternatives where respondents are asked to select their preferred alternative. Respondents in Dashiell's and in Petrusic and Jamieson's study should be primarily concerned with the potential benefits of choice, not only in the ordinal rankings of the alternatives.

Benefits from evaluation effort have also been modeled through effects on decision accuracy using stochastic choice models. Computer simulations by Payne, Bettman, and Johnson (1993) indicate that agents may have substantial savings in effort[3] with only small losses in accuracy (using an imposed criterion of EV maximization) if they were to use simplified heuristics (rules-of-thumb) for evaluating and selecting between risky alternatives rather than using more thorough evaluation methods.

Wilcox (1993a) carried out statistical tests for the significance of decision complexity and the levels of benefits from choice on both evaluation time and the patterns of choice. Wilcox's (1993a) paper reported the significant effects of question complexity and payoff level on decision time and on choice patterns. These effects were consistent with patterns hypothesized under a decision cost model using similarity. Wilcox (1993b) found evidence supporting a decision cost argument for the occurrence of a type of behavior for lottery pricing that violates all other models for risky choice.

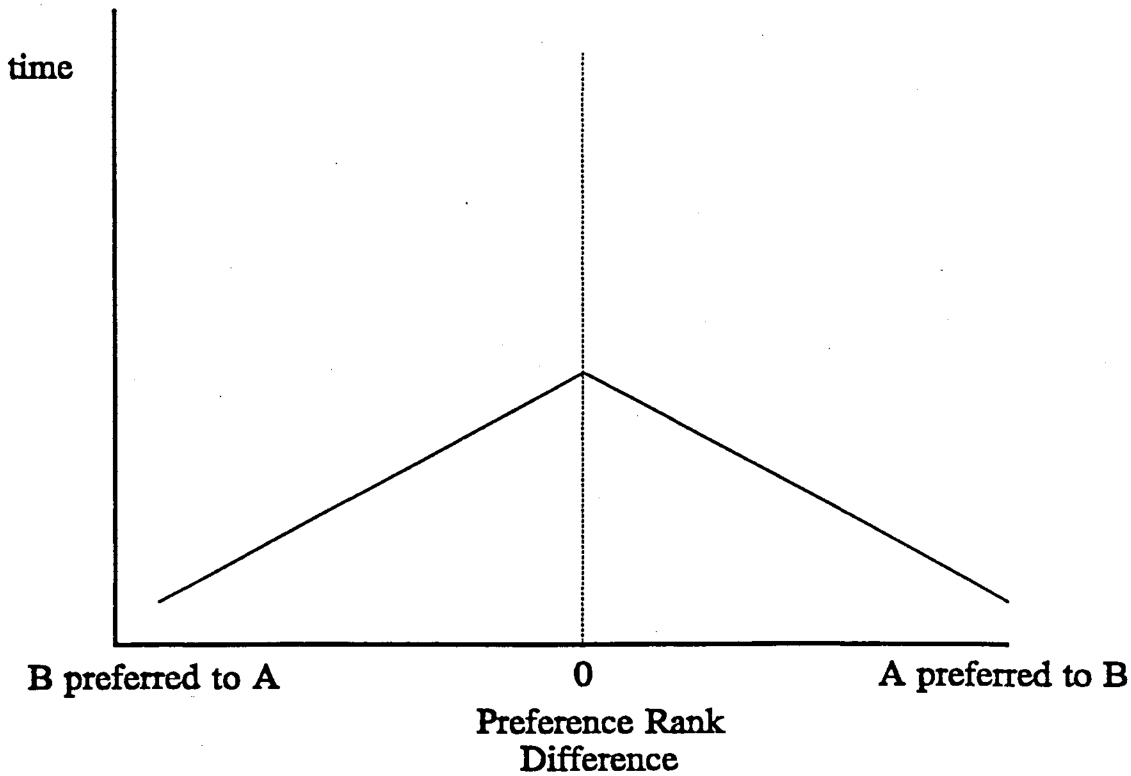
The current paper differs from Wilcox's work in that (a) similarity measures are used to explain decision complexity, (b) agents' evaluation time and choice patterns are evaluated with respect to a polynomial approximation of utility, (c) the range of the EV differences among the alternatives is larger that in Wilcox, (d) a much larger number of choice questions was asked per respondent, (e) generalized-least squares and discrete choice (probit) regression analysis methods used a continuous similarity measure, and finally (f) Wilcox included some probability of real payoffs to respondents while the present study does not.

C. Discussion: The Lack of Real Payoffs

As this study does not use real payoffs (a subject of an upcoming work), these results should be seen as somewhat preliminary. The inclusion of some level of real payoffs could be hypothesized affect both the time allocated and the patterns of choice as in Wilcox (1993). Further, the existence of real payoffs may affect the influence of similarity on choice. In light of the lack of real payoffs, the underlying assumption here is that the choice patterns from these gambles using strictly hypothetical payoffs do to some degree reflect the decision making process used for risky alternatives with real payoffs.

Figure 2

Hypothesized Relationship between Evaluation Time and Preference:



II. The Decision Maker's Problem

A. Motivation

Most individuals face a series of risky alternatives throughout their lifetime for which a common evaluation and choice strategy would be useful.[4] Decision makers are taken here to follow a pre-developed (ex-ante) process for selecting a decision method or algorithm. This ex-ante selection process can be periodically updated, but is as a rule not updated for every new risky choice question. Each algorithm has its own level of cost and of imperfection (error) in evaluation. Further, cost and error rate for these algorithms are taken to be inversely related. As a result, perfection in evaluation can be quite costly in terms of evaluation effort.

To examine the problem of ex-ante evaluation and decision algorithms in a more concrete and familiar manner, consider the development of an individual's problem-solving skills in general. Individuals receive both formal and informal training in decision making during their life through education and experience, giving rise to a number of problem solving strategies. Some reevaluation and refinement of these skills may also occur through introspection and the mulling over invented future or past (actual) decision scenarios. Throughout this learning and development process, decision criteria and methods are developed as guidelines for actual future choice. These decision rules prove useful since evaluation time is often limited when actual choice opportunities arise.

The following section develops a model for testing the level of effort (time) for the evaluation of risky choices.[5] In this model, the optimal level of effort is not reevaluated for each new risky problem, but rather a functional form defines the rule or algorithm to be utilized for a set of choices. The functional form predicting the selection of the evaluation effort level will be defined over characteristics of the risky alternatives, the individual's personal characteristics, and an error distribution. A useful summary of further psychological findings relevant for the perceptual and cognitive limitations addressed here can be found in Moyer and Dumais and in Busemeyer and Townsend.

B. Benefits of Choice

The benefits of evaluation effort for risky choice are defined here for the prospects (p, x) and (q, y). As discussed later, the risky alternatives in this study are defined so that (p, x) is more risky than (q, y) ($p < q$ and $x > y$) and neither alternative First-Degree Stochastically Dominates the other (Newbery and Stiglitz).

Assume that there is an underlying preference relation $\succsim$ between risky alternatives defined through the utility function $u(x, \alpha)$ on outcomes (x) and for the individual's characteristics as represented by the vector α. This function is taken to be consistent with EU (Jensen; Fishburn 1988) and allows evaluation of the alternatives through comparison of $p \cdot u(x, \alpha)$ and $q \cdot u(y, \alpha)$. The assumption of EU maximization is made to evaluate the potential for the similarity models to address EU violations. To make the point clearly, $u(\cdot)$ will be treated here as a representation that fulfills the EU axioms for *preferences*, but need not necessarily follow EU for *choices*.

The benefits of choice will be defined for the representative agent through (a) the EU differences between the alternatives and (b) the probabilities of incorrect choice.

Define the function

$$B(p, q, x, y) = [p \cdot u(x) - q \cdot u(y)] \tag{1}$$

as the underlying difference in the representative agent's underlying Expected Utility when (p, x) is selected over (q, y). When (p, x) $\succ$ (q, y), the choice of (p, x) over (q, y) would be correct and $B(\cdot) > 0$; when (q, y) $\succ$ (p, x), the choice of (p, x) would be incorrect and $B(\cdot) < 0$.[6] Given equation (1), we need to model choice error and discuss its relationship with similarity.

The individual would forego utility from incorrectly selecting a risky alternative [e.g., (q, y) when (p, x) is preferred] that differs from his/her optimal alternative. This loss would be small for those pairs of risky alternatives that have a nearly zero EU difference (small absolute level of $B(\cdot)$), but very large for other risky pairs whose EUs differ greatly. However, the EU differences defined in (1) are taken as not directly available to the decision maker. Only estimates of these EU differences are available, where these estimates are costly and are made with some error.

The choice of the estimation rule (or algorithm) to be used in the evaluation of risky alternatives is much like the problem faced in statistical or econometric analysis in selecting a method for estimating parameters of interest (Judge et al.). However, the evaluation problem here differs from that of selecting estimators in statistics in two ways and is parallel to problems of the optimal level of information gathering as discussed in Pratt, Raiffa and Schlaifer and in Stigler.

First, unlike the problems of estimator selection in statistics, benefit estimation and the choice between the risky alternatives are virtually simultaneous since the sign of the estimator for $B(\cdot)$ determines both the choice and the outcome of choice. This simultaneity is why the agents are taken to decide on an optimal ex-ante evaluation rule for the estimation of a series of upcoming risky choices, rather than recomputing the optimal rule for each new choice pair. Second, the loss function for the evaluation of prospects as defined in equation (1) is asymmetric. If the estimation is simply an overstatement of the true benefits $B(\cdot)$, there is no loss from subsequent choice. Only when there is a qualitative (sign) difference in the estimate will choice differ from the optimal, error free, selection.

Similarity for prospects is taken to depend on observable probability and outcome characteristics (p, q, x, y) of the alternatives and on the individual through α. Specifically, similarity is generalized to depend on absolute probability ($|p - q|$) and outcome differences ($|x - y|$) as in Rubinstein, and on an interaction term $x^* |p-q|$. The interaction term incorporates the dependence on the outcomes of the critical level of similarity for the probabilities from Azipurua et alia. This similarity is defined through the function:

$$s = s\left(p-q, x-y, x*(p-q), \alpha\right) \tag{2}$$

There may be other factors affecting the similarity and complexity of the choices, such as for the presentation of the gambles in a complex or in a simple manner as in Wilcox (1993a) and other qualitative characteristics of the problems as in Buschena (1992). Such dimensions of complexity have been added to the definition of and experiments for similarity and affect the qualitative measures influence on similarity. Similarity may also be influenced by the number of alternatives and the number of dimensions such as in Shugan (1980) and in Biggs et al. (1985) on choice.

Consider the effects of similarity for the case of $(p, x) \succ (q, y)$. Selection of (q, y) over (p, x) depends on the characteristics of the risky choice pair given by (p, q, x, y, s), the effort devoted to the evaluation of the alternatives defined by τ, characteristics of the individual given by α, and a random term from the distribution $g(\varepsilon_B)$. The probability of error from incorrectly selecting the risky alternative (q, y) over the truly preferred alternative (p, x) is defined by:

$$P[(q, y) \text{ selected over } (p, x) / (p, x) \succ (q, y)] \equiv \phi(B(\cdot), s, \tau, \alpha, g(\varepsilon_B)). \tag{3}$$

The benefits from choice for equation (1) and the likelihood of choice error in equation (3) give rise to the expected loss from choice error when (p, x) is preferred to (q, y):

$$\int_{\varepsilon} B(p,q,x,y,\alpha)\,\phi\left(B(\cdot),s,\alpha,\tau,\varepsilon_B\right)dg\left(\varepsilon_B\right) \tag{4}$$

To review, the terms in (4) are:

$B(\cdot)$ denotes the estimated benefits for the choice;
$\phi(\cdot)$ is the probability of choice error when (q, y) is incorrectly chosen over (p, x), x and y are the prospects' outcomes;
p and q are the probabilities for these prospects;
s is the similarity of the prospect pair, depending on p, q, x, y, and α;
τ is the effort level devoted to the evaluation (e.g., τ is evaluation time);
α is a vector of personal characteristics; and
ε_B is a draw from a random distribution that affects the error likelihood of the pair.

The decision maker is taken to minimize the following expression for the sum of expected choice error and decision cost when determining the optimal evaluation effort level τ:

$$\min_{\tau} \quad \psi \equiv \int_{\varepsilon} B(p,q,x,y,\alpha)\,\phi\,(B(\cdot),s,\alpha,\tau,\varepsilon_B)\,dg(\varepsilon_B) + c(\tau,\alpha) \tag{5}$$

Evaluation time is taken here to impose a cost in terms of utility $c(\tau, \alpha)$. This cost depends on the evaluation time and on the personal characteristics of the individual, stemming from either direct mental or opportunity costs of making decisions.

The model given here must be viewed very much in the "as if" predictive sense characterized by Friedman. There are two difficulties with a strict interpretation of (5) as an accurate *description* of behavior. First, if agents actually made decisions as called for in this equation, they would encounter an infinite loop of decision costs as discussed in Conlisk (1988) and in Lipman.[7] Second and more critically, if the EU difference function $B(\cdot)$ were known without error, no evaluation effort would be needed since selection is simultaneously determined with evaluation. Components of the risky alternatives (p, q, x, y, s) and the key individual characteristics (α) that underlie risky evaluation and decision are hypothesized to affect choice in prescribed ways discussed below. These hypothesized effects allow for tests of behavior in a repeated choice setting.

Under the conditions for the Implicit Function Theorem, the optimal selection of the effort level minimizing the expression in equation (15) could be implicitly defined by the function:[8]

$$\tau^*\left(B(\cdot), s, \alpha, g(\varepsilon_B)\right). \tag{6}$$

This optimal effort level depends on the components of the risky alternatives, on personal characteristics, and on the random error distribution $\underline{\varepsilon}_B$. Again, let the riskier alternative (p, x) be preferred to the less risky alternative (q, y). Alternative (p, x) is selected over (q, y) whenever $B(\cdot\,;\varepsilon_B) > 0$, occurring in two ways:

1. no choice error: $(p, x) \succ (q, y)$, $B(\cdot) > 0$, probability $[1 - \phi(\cdot\,;\varepsilon_B)_{B(\cdot)>0}]$,
2. choice error: $(q, y) \succ (p, x)$, $B(\cdot) < 0$, probability $\phi(\cdot\,;\varepsilon_B)_{B(\cdot)<0}$.

The expected likelihood of the riskier alternative (p, x) being selected over (q, y) is then:

$$\int_{\varepsilon_B} \left\{\left[1 - \phi(\cdot;\varepsilon_B)_{B(\cdot)>0}\right] + \phi(\cdot;\varepsilon_B)_{B(\cdot)<0}\right\} dg(\varepsilon_B) = \pi\left(B(\cdot), s, \alpha, g(\varepsilon_B)\right) \tag{7}$$

The minimization of (5) and the subsequent selection of $\tau^*(\cdot)$ in the choice hierarchy represented in (6) gives rise to the following function for the likelihood of the more risky alternative (p, x) over (q, y):

$$\pi^*\left(B(\cdot), s, \alpha, g(\varepsilon_B)\right), \tag{8}$$

estimated with some error throughout the range of $\underline{\varepsilon}_B$. Comparative statics results could in principle be carried out on $\tau^*(\cdot)$ and $\pi^*(\cdot)$ with respect to similarity from the first order condition. However, little is known about the signs of the second derivatives of $\phi(\cdot)$ and $c(\cdot)$ with respect to τ, and with respect to s. Instead, the next section presents testable hypotheses in a less formal manner.

C. Hypothesized Effects of Similarity on Evaluation Time and Choice

There are two observable measures affected by similarity as indicated implicitly by equations (6) and (8). The relationship between evaluation time and similarity is hypothesized to follow a cubic relation as in Figure 3. First, if the alternatives are already quite dissimilar (outside the set (μ_{pq}, μ_{qp})), evaluation time is hypothesized to decrease as they become more dissimilar as the likelihood of choice error $\phi(\cdot)$ is taken to be small. When the alternatives are quite similar (within the neighborhood bounded by σ_{pq} and σ_{qp}), the returns to evaluation time decrease as the alternatives become more similar since the potential losses $(B(\cdot))$ from choice error are quite small. Evaluation time should be maximized for intermediate levels of similarity, illustrated by $(-\eta_*, \eta_*)$ in Figure 3.

Another testable assumption regarding similarity is its influence on patterns of risky choice through $\pi^*(\cdot)$. When agents select the more risky alternative, they trade off safety for upside gain potential. The similarity models predict that the riskier alternative is more likely to be selected as similarity increases. These testable hypothesized effects of similarity are stated formally below.

> *Hypothesis 1*: Trade-offs between the costs and benefits of evaluation effort lead individuals to evaluate and select between risky alternatives through decision algorithms as represented by the degree of evaluation effort. The costs and benefits of evaluation can be represented through question similarity, giving the following testable properties for evaluation effort and choice error:
>
> H1a. The decision time $\tau^*(\cdot)$ is locally minimized for extremely similar alternatives, has two local maxima for moderately similar alternatives, and decreases significantly with decreases in similarity for extremely dissimilar alternatives (illustrated in Figure 3).
>
> b. For otherwise risk averse agents and where the outcomes are not similar, the likelihood of selection for the riskier alternative $\pi(\cdot)$ increases as the pairs become more similar.

These hypotheses are further developed in Section C for a specific utility representation. This utility representation uses a third order polynomial approximation for utility. Empirical support for the pattern of choice in Hypothesis 1b is given in a study by Mosteller and Nogee from their early effort to elicit risk attitudes.

D. A Specific Utility Representation for Choice Under the Effects of Similarity

This section defines a third-order polynomial representation for an EU preference representation over the outcomes, where the influence of each of these polynomial terms is hypothesized to be affected by the similarity of the alternatives. This third-order polynomial form is quite general and allows for tests of the effects of similarity on choice patterns and evaluation time.

Define the underlying utility representation for a representative agent as a cubic relation on the moments of the alternatives (see for example Newbery and Stiglitz):

$$u(x) = \theta_1 px + \theta_2 p * x^2 + \theta_3 p * x^3. \tag{9}$$

This cubic representation allows either risk averse, risk neutral, or risk preferring behavior. In addition, risk attitudes could either be qualitatively uniform or differing qualitatively among segments of the outcomes as in Friedman and Savage (Figure 4).

Figure 3
Hypothesized Relationship between Time and Eu Difference

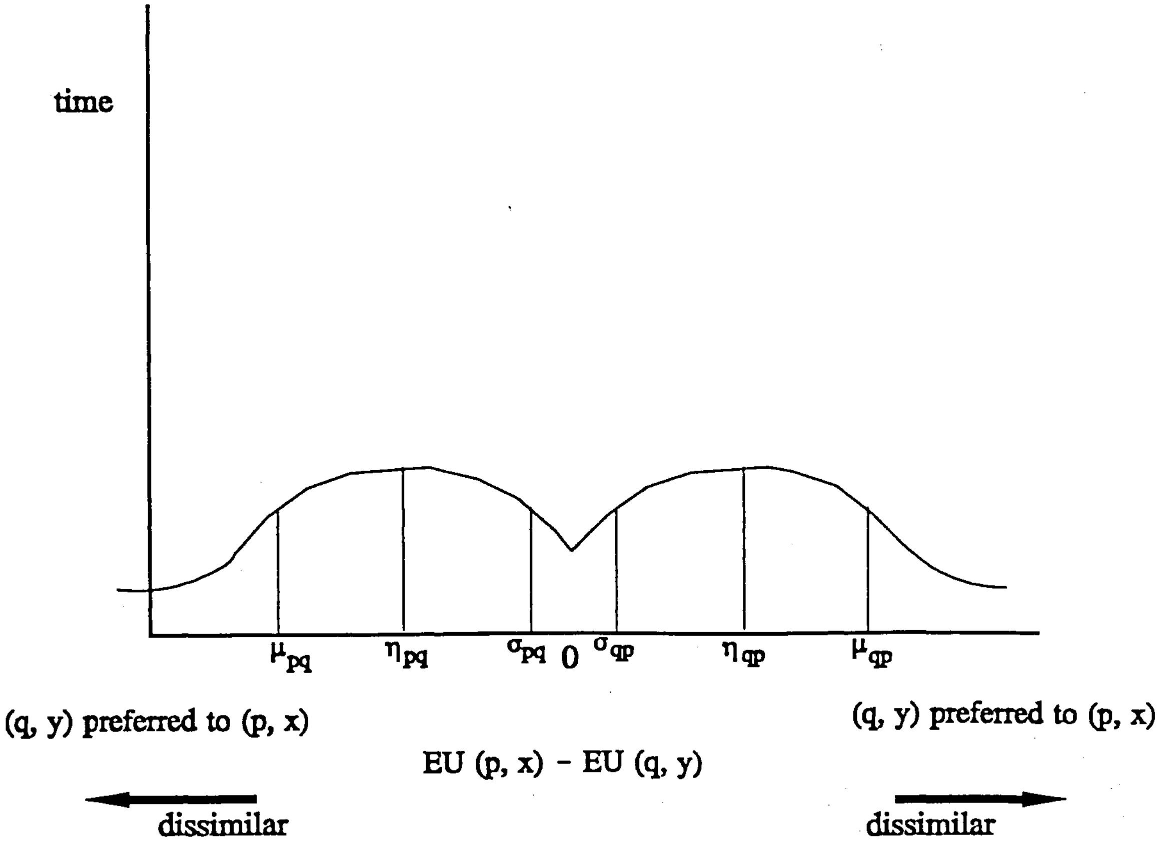
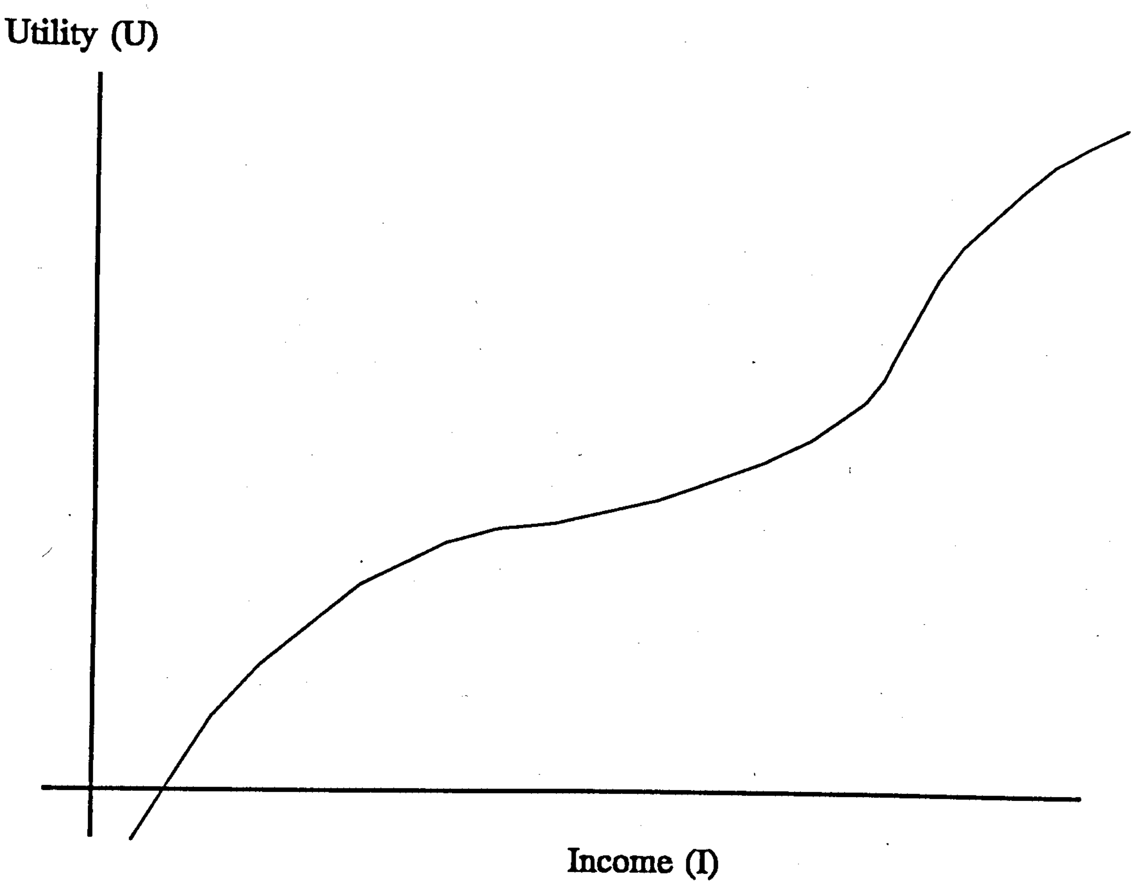


Figure 4
Illustration of Typical Shape of Utility Curve



Source: Friedman and Savage.

Similarity is hypothesized to affect choice patterns and evaluation time for a third moment utility representation in the following manner. For alternatives that are quite dissimilar, (e.g., those beyond the neighborhoods defined by (μ_{pq}, μ_{qp}) in Figure 3), the first order difference in EU determines the desirability of the alternatives and the evaluation effort since it overwhelms the higher order terms. For quite similar alternatives (e.g., those within the (σ_{pq}, σ_{qp}) neighborhood in Figure 3) the EU differences are so small that the evaluation of the alternatives should depend primarily on the upside potential reflected by their higher order differences.[9] For the risky pairs in this study with relatively low stakes, as similarity increases the influence of the higher order terms on evaluation effort and choice patterns is hypothesized to increase.

The actual benefit from choice (from the definition of B(·) in equation (1)) and under the utility representation in equation (9) then becomes:

$$
\begin{aligned}
B'(p,q,x,y,\alpha) &= pu(x) - qu(y) \\
&= p\left[\theta_1 x + \theta_2 x^2 + \theta_3 x^3\right] - q\left[\theta_1 y + \theta_2 y^2 + \theta_3 y^3\right] \\
&= \theta_1 z_1 + \theta_2 z_2 + \theta_3 z_3, \quad z_i = px^i - qy^i, \quad i \in \{1,2,3\}
\end{aligned}
\tag{10}
$$

Equation (10) states that the underlying EU differences can be separated into differences between products of the probabilities and polynomial outcome terms (moment differences).

Let the coefficient weights on the polynomial terms in (10) be affected by similarity through interaction terms z_i*s carrying the respective coefficients $\gamma_i(s)$ for all iε{1,2,3}. These hypothesized similarity effects can be tested through estimates of the $\gamma_i(s)$ coefficients' effects on choice defined by $\pi^*(\cdot)$ from equation (8) and by evaluation time $\tau^*(\cdot)$ from equation (6). This test of the similarity model further hypothesizes that, for the risky alternatives in this study, the higher order difference terms receive more weight in choice as similarity increases since the potential gain (reflected by the higher order moments for these questions) becomes large relative to the risk. The estimated benefit equation (10) is re-defined as:

$$
\begin{aligned}
&\beta^*(p,q,x,y,s,\alpha) \\
&\quad = B'(p,q,x,y,\alpha) + \gamma_1(s) * z_1 * s \\
&\quad + \gamma_2(s) * z_2 * s + \gamma_3(s) * z_3 * s
\end{aligned}
\tag{11}
$$

The estimated benefits in equation (11) will be taken as explanatory variables for the likelihood of the riskier choice in equation (8) and for the time devoted to the evaluation of the alternatives in equation (6) in a simultaneous manner, giving equations (12) and (13), respectively:

$$\begin{aligned} &\textit{optimal evaluation time} = \tau^*\left(\beta'(\cdot), s, \alpha, g\left(\varepsilon_B\right)\right), \\ &\textit{where } \beta'(\cdot) = \beta'(p,q,x,y,s,\alpha) \textit{ from } (11) \end{aligned} \tag{12}$$

$$\begin{aligned} &\textit{optimal likelihood of selecting } (p,x) \textit{ over } (q,y) \\ &= \pi^*\left(\beta'(\cdot), s, \alpha, g\left(\varepsilon_B\right)\right) \end{aligned} \tag{13}$$

The primary goal of the empirical analysis will be to estimate the coefficients $\theta_i(\alpha)$ and $\gamma_i(s)$, $i = \{1, 2, 3\}$ from equation (12) and equation (13). Tests will be conducted to test the hypothesis that the influence of the higher order moments on choice and evaluation effort increases as the alternatives become more similar.

Hypothesis H2: *For a cubic representation of utility over outcomes and for small gambles offering increased risk for upside gain, increased similarity will:*

a.) increase the effects of the second and third order polynomial terms on evaluation time.
b.) increase the effects of these second and third order polynomial terms on the likelihood of the riskier choice.

III. Empirical Testing and Results

A. Experimental Design

The experiment was designed to test for the effects of similarity on evaluation effort and on risky choice patterns. Subjects were shown pairs of risky alternatives through a visual display on a computer screen as is illustrated in Figure 5.[10] Respondents were instructed to make their choices as if they were selecting between risks with actual payoffs and to think carefully about their choices.

Design. Gambles were constructed from the 6 x 6 (amount by probability) factorial design well-known from many previous experiments (e.g., Mellers et al. 1991; Mellers et al. 1992). This design gives a very intensive treatment of choice over risky gambles that have low payoff levels. Only risky choice pairs that did not include a first-degree stochasically dominant alternative were used in this design, giving a total of 225 risky choice pairs as illustrated in Table 1. Each entry in the table indicates a pair of risky alternatives for which respondents were asked to select between the row gamble and the column gamble. For example, choice was elicited for the risky pair comparing the column gamble giving a .29 chance of $17.50 with a row gamble giving a .17 chance of $31.50. Note that, for a number of gambles, the riskier row gamble has a higher EV than the less risky column gamble.

Figure 5

Select the gamble (A or B) that you prefer.

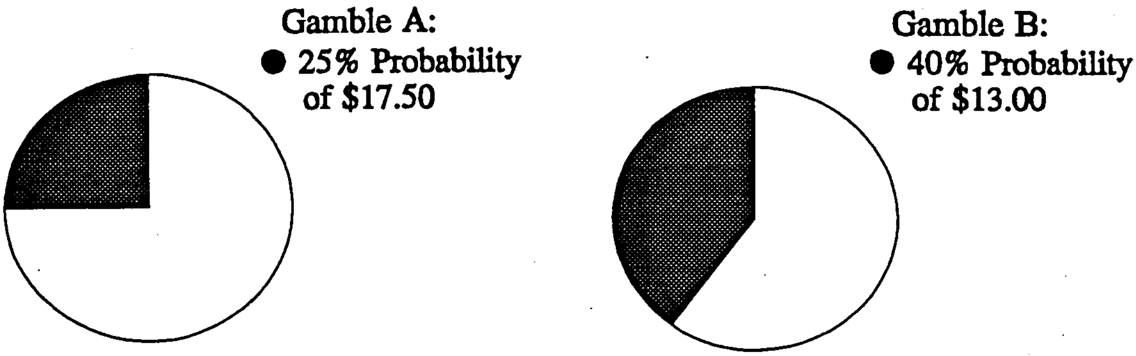


TABLE 1. Gambles Included in the Study

	.05					.09					.17					.29					.52					.94				
	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e
.05 b						■					■					■					■					■				
c						■	■				■	■				■	■				■	■				■	■			
d						■	■	■			■	■	■			■	■	■			■	■	■			■	■	■		
e						■	■	■	■		■	■	■	■		■	■	■	■		■	■	■	■		■	■	■	■	
f						■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
.09 b											■					■					■					■				
c											■	■				■	■				■	■				■	■			
d											■	■	■			■	■	■			■	■	■			■	■	■		
e											■	■	■	■		■	■	■	■		■	■	■	■		■	■	■	■	
f											■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
.17 b																■					■					■				
c																■	■				■	■				■	■			
d																■	■	■			■	■	■			■	■	■		
e																■	■	■	■		■	■	■	■		■	■	■	■	
f																■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
.29 b																					■					■				
c																					■	■				■	■			
d																					■	■	■			■	■	■		
e																					■	■	■	■		■	■	■	■	
f																					■	■	■	■	■	■	■	■	■	■
.52 b																										■				
c																										■	■			
d																										■	■	■		
e																										■	■	■	■	
f																										■	■	■	■	■
.94 b																														
c																														
d																														
e																														
f																														

Note: a = $3; b = $5.40; c = $9.70; d = $17.50; e = $31.50; and f = $56.70.

Forty (40) students participated in the experiment; of these students, 25 completed the entire set of 225 questions for risky choice selection and 10 repeated questions, while 15 partially completed the study due to time constraints.[11] These students were undergraduates at UC-Berkeley enrolled in an introductory psychology course and received course credit plus a small non-monetary payment for their participation. The results for each question from the 15 uncompleted responses in the sample were treated the same as each question was for those from the 25 completed experiments after comparisons suggested no significant differences in responses. There were a total of 8,461 risky choice responses from these 40 students.

After the visual display of the gambles, respondents were asked to complete three tasks; these tasks were given in the same order for each question. First, the respondent was asked to select his or her preferred alternative; i.e., to pick gamble A or gamble B in the pair. The pair itself and its display order (left or right in the display) was randomly selected for each respondent. Second, the respondent was asked to rank his or her perceptions of similarity for this risky pair on a scale from 1 to 100. This continuous scale was further described to the respondent through the following qualitative reference terms:

1 = Very Very Dissimilar
25 = Very Dissimilar
50 = Similar
75 = Very Similar
100 = Very Very Similar

Third, the respondent was asked to rank his or her strength of preference between the risky alternatives on a scale from 1 to 100. This continuous scale was also further described to the respondent through the following qualitative reference terms:

1 = Very Very Strong
25 = Strong
50 = Moderate
75 = Weak
100 = Very Weak

The experiment incorporated a clock to measure the time taken to select an alternative.[12] This clock began when the risky pair was displayed and stopped when the choice between the pair was made. This response time measure includes the time required to read the alternatives (information gathering), to evaluate their relative desirability, and to physically make the selection. As a result, the important issue is the relative differences in evaluation time between gambles. This time measure is was hypothesized to reflect the evaluation effort level as previously discussed in Section II.

The respondents were encouraged to take breaks throughout the experiment; the task of completing 235 risky choice evaluations, choices, and rankings was not easy. These rest periods occurred throughout the experiment for each subject at non-standard intervals, affecting the evaluation time recording. These breaks caused considerably larger time measures for some observations. Corrections for this effect for the analysis of decision time are addressed below in Subsection B.

B. Results: Initial Analysis

Table 2 gives the percentage of respondents selecting the less risky column alternative over the more risky row entry. As expected, for rightward movements within a row where the outcome and the probability payoff level of the column bet increases, the row gamble becomes relatively less attractive and the percentage selecting the column increases consistently.

The outcomes or the probabilities (or both) for the row gamble increase through downward movements within a column. Therefore, the column gamble becomes relatively less attractive and the percentage selecting the column gamble decreases consistently for downward movements within columns. For each 5x5 triangular sub-matrix defined by a column and a row probability, the lowest population proportion selecting the column over the row alternative is in the lower left, where the row gamble is most attractive. The population percentage selecting the column gamble is highest for elements on the diagonal of these matrices, where the column gamble is most attractive relative to the row.

Table 3 lists the median perceived similarity measures. The scale used was from 1 (Very Very Dissimilar) to 100 (Very Very Similar). The larger the number in this table, the more similar the alternatives were perceived to be. Consider the 5x5 triangular sub-matrices near the diagonal of the larger matrix (e.g., those for the column probability .17 and the row probability .29). The alternatives are generally perceived to be rather similar (large median numbers) in the upper left corners of these sub-matrices, but this perceived similarity decreases as one moves down a column. Likewise, starting from the lower left corner of each sub-matrix, these alternatives are perceived to be more similar for rightward movements. In general, the lowest numbers (most dissimilar alternatives) are for pairs in the upper right 5x5 triangular sub-matrices, where the row (e.g., where the row gambles carry the 5% probability level and the column gambles carry the 9% probability level) is quite low.

Table 4 reports the population median strength of preference measures from the survey. The range of these numbers is from -100 (Very Very Strong for the less risky column entry) to 100 (Very Very Strong for the more risky row entry. These median values very closely reflect the percentage choice values from Table 2. In total, the correlation between these two sets of numbers is almost perfect (99%). As expected, downward movements within a column (the row gamble becomes more attractive) lead to increased strength of preference medians. Alternatively, as one moves across a particular row, the column gambles become relatively more attractive and the median numbers decrease.

The median population response time in seconds are given in Table 5. These median time levels ranged from 2 to 6 seconds; selections were made quite quickly. Still, the median times for some gambles were more than double those for others. Recall from Table 3 that the upper rightmost 5x5 triangular matrices were judged to be the most dissimilar. Median decision times for these matrices are lower than those for more similar triangular matrices below them or to their left. In addition, triangular sub-matrices in the center of Table 5 (e.g., for columns with probabilities .17 or .29) showed larger median evaluation and choice times over the same row gambles than do the more similar leftmost 5x5 triangular sub-matrices with column probabilities .09.

TABLE 2. Proportion of Less Risky Choice (column over row)

	.05					.09					.17					.29					.52					.94				
	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e
.05b						.44					.76					.78					.89					.97				
c						.18	.46				.68	.74				.81	.94				.94	.95				.95	.85			
d						.08	.20	.28			.32	.53	.75			.55	.78	.88			.89	.92	.86			.87	.88	.87		
e						.10	.13	.14	.28		.28	.26	.30	.78		.51	.55	.78	.85		.73	.89	.90	.90		.87	.89	.92	.92	
f						.11	.15	.13	.12	.28	.19	.19	.31	.58	.76	.31	.49	.61	.78	.81	.62	.73	.77	.80	.86	.76	.80	.79	.89	.93
.09b											.60					.74					.95					.92				
c											.43	.58				.74	.81				.89	.92				.95	.92			
d											.21	.24	.37			.42	.53	.81			.76	.95	.89			.81	.97	.95		
e											.14	.15	.20	.43		.36	.49	.59	.79		.51	.78	.83	.93		.77	.88	.86	.92	
f											.16	.16	.16	.32	.64	.26	.32	.38	.60	.73	.49	.62	.75	.89	.92	.67	.76	.82	.92	.97
.17b																.63					.84					.92				
c																.42	.62				.73	.95				.90	.92			
d																.16	.35	.44			.55	.71	.92			.70	.92	.93		
e																.13	.19	.26	.58		.42	.54	.65	.88		.67	.68	.83	.92	
f																.10	.13	.10	.25	.53	.38	.53	.46	.79	.91	.49	.63	.69	.97	.93
.29b																					.73					.95				
c																					.63	.85				.78	.92			
d																					.27	.54	.75			.66	.82	.88		
e																					.28	.26	.32	.83		.53	.58	.74	.85	
f																					.18	.18	.35	.45	.66	.30	.42	.55	.82	.87
.52b																										.82				
c																										.47	.81			
d																										.29	.48	.62		
e																										.26	.26	.45	.63	
f																										.24	.18	.33	.43	.76
.94b																														
c																														
d																														
e																														
f																														

Note: a = \$3; b = \$5.40; c = \$9.70; d = \$17.50; e = \$31.50; and f = \$56.70.

TABLE 3. Median Similarity Judgements from 1 (very, very dissimilar) to 100 (very, very similar).

Note: a = $3; b = $5.40; c = $9.70; d = $17.50; e = $31.50; and f = $56.70.

TABLE 3. *Median Similarity Judgements from 1 (very, very dissimilar) to 100 (very, very similar).*

	.05					.09					.17					.29					.52					.94				
	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e
.05b						89					75					32					25					10				
c						75	78				60	50				45	40				30	31				10	20			
d						65	60	75			50	50	50			45	50	45			34	25	25			15	20	15		
e						50	35	50	75		50	45	68	52		44	50	45	40		34	30	30	25		25	25	15	15	
f						30	40	45	46	75	45	40	40	50	50	40	30	40	45	30	25	25	25	25	25	25	20	23	20	20
.09b											75					75					25					9				
c											60	75				60	50				25	25				25	10			
d											50	60	70			50	50	50			50	30	43			25	20	25		
e											40	45	50	60		45	50	50	50		30	34	35	40		20	25	25	20	
f											40	32	30	50	50	35	34	40	50	50	26	40	30	40	25	25	25	20	25	20
.17b																70					40					20				
c																59	75				40	40				20	25			
d																50	50	67			50	40	40			30	30	25		
e																40	50	50	75		30	40	40	50		25	25	30	25	
f																40	30	40	50	60	30	32	40	40	50	22	30	30	25	23
.29b																					50					25				
c																					50	50				25	20			
d																					46	50	50			34	29	25		
e																					35	44	50	60		25	30	35	30	
f																					40	30	42	50	60	25	25	25	40	25
.52b																										50				
c																										40	45			
d																										34	35	40		
e																										30	45	35	50	
f																										25	26	28	50	50
.94b																														
c																														
d																														
e																														
f																														

Note: a = $3; b = $5.40; c = $9.70; d = $17.50; e = $31.50; and f = $56.70.

TABLE 4. Strength of Preference: Range, -100 (strong for less risky [column]) to 100 (strong for more risky [row])

	.05					.09					.17					.29					.52					.94				
	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e
.05b						33					-63					-70					-80					-99				
c						50	30				-40	-50				-65	-74				-70	-75				-90	-90			
d						80	60	60			55	-1	-50			-10	-50	-60			-60	-70	-75			-75	-87	-87		
e						75	75	80	70		50	50	50	-50		-1	-25	-50	-64		-50	-54	-60	-75		-65	-78	-80	-90	
f						78	80	80	75	67	59	70	50	-25	-60	50	10	-25	-54	-7	-25	-50	-50	-70	-75	-60	-75	-80	-80	
.09b											-21					-50					-75					-90				
c											34	-25				-50	-60				-74	-80				-85	-90			
d											65	59	50			30	-5	-58			-50	-70	-75			-75	-85	-85		
e											67	70	60	25		50	20	-25	-56		-1	-50	-67	-67		-75	-75	-80	-85	
f											70	75	75	50	-35	60	50	50	-25	-6	15	-25	-50	-60	-79	-25	-60	-80	-80	
.17b																-45					-70					-95				
c																50	-40				-50	-75				-80	-90			
d																70	50	40			-25	-50	-65			-50	-76	-80		
e																65	70	56	-30		25	-25	-45	-65		-70	-54	-80	-86	
f																76	70	75	50	-2	50	-20	-33	-50	-70	25	-50	-50	-80	
.29b																					-50					-90				
c																					-50	-61				-75	-85			
d																					55	-1	-50			-75	-70	-85		
e																					60	60	50	-50		-5	-45	-65	-80	
f																					70	60	50	50	-50	50	45	-50	-75	
.52b																										-85				
c																										25	-75			
d																										60	10	-60		
e																										65	65	30	-65	
f																										75	75	70	56	
.94b																														
c																														
d																														
e																														
f																														

Note: a = \$3; b = \$5.40; c = \$9.70; d = \$17.50; e = \$31.50; and f = \$56.70.

TABLE 5. Median Response Time (seconds)

Note: a = $3; b = $5.40; c = $9.70; d = $17.50; e = $31.50; and f = $56.70.

TABLE 5. Median Response Time (seconds)

	.05					.09					.17					.29					.52					.94				
	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e
.05b						4					5					4					4					3				
c						5	4				4	4				4	3				3	4				3	2			
d						4	4	4			5	5	3			4	4	5			4	4	3			2	2	3		
e						3	4	3	4		4	4	6	4		4	6	4	5		4	3	3	3		3	2	3	3	
f						3	3	3	4	5	5	5	4	5	4	4	4	5	4	3	3	4	4	4	3	4	3	3	2	2
.09b											5					4					3					2				
c											4	5				4	4				3	3				2	2			
d											4	5	5			4	6	4			4	3	3			3	3	3		
e											4	4	5	6		5	5	6	5		4	5	3	3		3	3	3	2	
f											4	3	3	4	5	4	5	4	5	4	5	4	4	4	3	3	3	3	3	3
.17b																4					5					3				
c																4	4				4	3				3	3			
d																4	4	6			5	3	3			3	3	3		
e																4	4	4	5		4	4	4	4		4	4	3	3	
f																4	4	4	4	5	4	4	4	4	3	3	4	3	3	3
.29b																					4					2				
c																					5	4				3	3			
d																					3	4	4			4	3	2		
e																					3	3	5	4		3	3	4	3	
f																					3	3	4	4	3	5	3	3	3	3
.52b																										3				
c																										3	3			
d																										3	3	2		
e																										3	3	3	3	
f																										2	3	3	3	3
.94b																														
c																														
d																														
e																														
f																														

Note: a = $3; b = $5.40; c = $9.70; d = $17.50; e = $31.50; and f = $56.70.

The patterns of median evaluation time in Table 5 suggest that evaluation time is related to similarity in a non-linear (e.g., quadratic) fashion as in Hypothesis 1a (Figure 3). The median response patterns given in Tables 2-5 are further evaluated below.

C: Regression Analysis

The results of this experiment are further evaluated through the use of regression analysis. This method of analysis is of particular value for these data since it allows for tests of the significance of a continuous measure of similarity on choice and evaluation time while allowing for randomness in responses. This approach also allows the estimation of the significance and direction of the effects of differences between individuals. For a review of the regression approach, see the texts by Judge et alia. A list of the variables used in this analysis, plus their means and standard deviations, are given in Table 6.

Determinants of Perceived Similarity. A generalized least squares (GLS) regression model was carried out to fit respondents' similarity perceptions over the prospects. The dependent variable was the elicited similarity scale level that took values from 1 (very very dissimilar) to 100 (very very similar). The basis for using this GLS model was the expected heteroscedasticity between individuals in the population due to differences in individuals' ranges for similarity perceptions.

An estimated covariance matrix was constructed for a weighted least squares model (Judge et al.) to fit similarity perceptions, where the weights are the inverse of estimates for individual (i's) standard deviations from an OLS regression, given by $1/s_i$, where $s_i^2 = \frac{1}{T}\sum_{j=1}^{T} e_{ij}^2$.[13] In this estimation, e_{it}^2 is the square of the estimated error (from an ordinary least squares model) for the ith respondent's jth risky choice pair using the same explanatory (right-hand side) variables as in the GLS model defined below.

The coefficients in this GLS regression were estimated through a linear model based on equation (2). This equation includes variables that describe the risky choice pairs, personal characteristic variables and variables reflecting the survey itself.

The right hand side (explanatory) question characteristic variables of interest include: the difference in the outcomes $(x - y)$, the absolute difference in the pair's probabilities $(p - q)$ of winning, and an interaction term as the product of the outcome level and the absolute probability difference $[x * |p - q|]$. Personal indicator (dummy) variables were included to allow for individual location differences for each respondent (save one) and a learning effect variable λ that is the order of the risky pair in the respondent's survey.[14]

Similarity perceptions were well defined through objective measures in this GLS estimation. A partial listing of coefficients for this model are shown in Table 7; individual intercept coefficients variables are not included. Due to concerns regarding the choice of the critical level given large sample size, an alternative test criteria using a Bayesian approach (Poirier) to calculate the size of the critical t-statistic is made available for tests of the null hypotheses of zero coefficients. This Bayesian significance level calls for the estimated t to be slightly greater than 3.00, larger than the conventionally used critical t-values from standard tables.[15]

TABLE 6:
Means and Standard Deviations of Variables Used in the Study

Variable Name	Mean	Standard Deviation
Outcome Difference (x – y)	23.56	16.59
Probability Difference (p – q)	-0.39	0.28
Interaction Term: Outcome*Probability Difference [x*(p – q)]	12.68	12.78
Question Order	101.13	61.81
$z_1 = px - qy$	-0.50	6.04
$z_2 = px^2 - qy^2$	130.70	292.55
$z_3 = px^3 - qy^3$	8542.20	16726.00
Estimated Similarity	59.14	18.42
Estimated Similarity * z_1	-13.82	383.68
Estimated Similarity * z_2	7764.20	11429.00
Estimated Similarity * z_3	0.25E+06	0.35E+06
Perceived Similarity	66.80	24.80
Response Time	4.55	3.62
log[Time + 1]	1.52	0.64
Percentage Riskier Choice	0.63	0.48

TABLE 7:
Generalized Least Squares for Dissimilarity Perceptions

Variable Name	Estimated Coefficient	Standard Error	T-Ratio 8417 DF
Outcome Difference $(x - y)$	-0.02	0.02	-1.26
Probability Difference $(p - q)$	-23.23	1.26	-18.46
Interaction Term: Outcome*Probability Difference $[x*(p - q)]$	0.12	0.03	3.72
Question Order	-0.01	0.003	-1.94
Constant	2.22	0.04	56.87

$R^2 = 0.7061$ Adjusted $R^2 = 0.7046$

The range of the dependent variable was from 1 (very very dissimilar) to 100 (very, very similar).

Thirty-five (35) of the 39 personal indicator (dummy) variables had t-values above 3.00.

The coefficient on the relative outcome difference $[(x - y)]$ was not estimated significantly. The coefficient on the absolute probability difference ($|p - q|$) was significantly different from zero using either their conventional t-value or those from the Bayesian method; the pair was judged to be more similar as this difference increased. The coefficient on the interactive product term ($x * |p - q|$) was also significantly estimated; increases in this product term reduce the effects of the probability difference on similarity. The regression results support the dependence of the critical probability difference on the outcome level ($\varepsilon_{P(X)}$) as suggested by Azipurua et al. and illustrated in Figure 1.

The learning effect of question order on similarity judgments had a small but significant effect in this GLS regression. Similarity was not greatly exaggerated as the respondents worked through the survey.

Although they are not listed in Table 7, all but four of the 39 personal intercept terms were significantly different from zero using the Bayesian criteria. These findings indicate considerable interpersonal location differences for similarity perceptions. Model fit is indicated by an adjusted R-squared measure of .70.

Similarity and Choice Error. Previous authors (Smith and Walker) have hypothesized that the stakes involved in choice (as influenced by the existence of real vs. hypothetical choice here) affects choice error in repeated questions. As extended to similarity, the more similar pairs should have higher error rates for repeated questions. This hypothesis calls for error rates to increase with similarity and has been tested (Wilcox 1993) for its subsequent effect on variance. An alternative hypothesis is for equal choice error for repeated questions with different but consistent decision rules (algorithms) for similar and dissimilar risky choices. That is, the method of decision is hypothesized to be relatively constant (small error rate) for a particular pair of risky alternatives, but the methods of evaluation differ between pairs.

To test these two competing hypotheses, a probit regression (dependent variable is 1 for choice error) was run on the repeated questions in the survey. Error occurs if A (B) is selected in the initial pair, but B (A) was selected when the pair was repeated. The large number of questions made subject indignation due to the repeated question unlikely. The first five questions were repeated at the mid-point of the experiment as questions #101 to #105.[16]

The results of this probit regression for choice error are shown in Table 8. The explanatory variables included estimated similarity (s), variables $z_i = px^i - qy^i$, $i\varepsilon\{1, 2, 3)$ as in equation (10), interaction terms of estimated similarity with the z_i's ($s * z_i$), question order, and individual dummy variables (not shown). There was no significant effect of fitted similarity on the occurrence of choice error; a joint test on the coefficient estimates for similarity and the interaction terms with the z_i's showed no significant differences from zero. None of the variables of interest (save for individual differences) had a significant effect on the occurrence of choice error. It appears that, while individuals used different methods of evaluation for similar and dissimilar risky alternatives, they also did so consistently throughout the survey and choice error (reversals) are random. These findings are consistent with those in Wilcox (1993a and 1993b).

TABLE 8:
Probit Regression for the Occurrence of Error
in Repeated Choice: Dummy Intercepts INTERCEPTS

Variable Name	Estimated Coefficient	Standard Error	N=461
Estimated Similarity (S)	-0.033	0.022	-1.497
$Z_1 = px - qy$	0.232	0.214	1.086
$Z_2 = px^2 - qy^2$	-0.012	0.0126	-0.980
$Z_3 = px^3 - qy^3$	0.152E-03	0.167E-03	0.911
Interaction: $S*Z_1$	-0.441E-02	0.332E-02	-1.370
Interaction: $S*Z_2$	0.248E-03	0.191E-03	1.299
Interaction: $S*Z_3$	-0.319E-05	0.255E-05	-1.255
Constant	1.073	1.028	1.044

Log-likelihood (0) = -272.33 Log-likelihood function = -235.56

Likelihood Ratio Test = 73.527 with 46 D.F.

Dependent variable is 1 if a choice error (reversal) occurred.

Prediction Success Table

		ACTUAL	
		0	1
	0	315	103
Predicted	1	18	25

Percentage of Right Predictions = 0.74

Similarity and Evaluation Time: Simple Model. The hypothesis H_{1a} calling for the dependence of evaluation time on similarity was tested against the null hypothesis of a zero coefficient. For this hypothesized concave relationship previously discussed and illustrated in Figure 3, decision time should (a) decrease as the alternatives become more similar for a class of risky choices that are already quite similar, (b) reach a maximum for risky choice pairs of moderate similarity, and (c) decrease as the alternatives become more dissimilar for a class of relatively dissimilar risky choice pairs.

In this regression model, evaluation time was transformed into logs through the formula log (time + 1). This monotonic transformation offered better model fit than did regressions on time alone, and gives qualitatively the same results in coefficient estimation. Because of the occurrence of periodic rest periods affecting the time recording within the experiment as discussed previously, the time recorded for each subject was corrected for outliers. Outliers were removed for this time regression for relatively long response periods using the criteria of two times each respondent's estimated standard deviation as the limit. That is, choice times more than two times the individual's sample standard deviation from the population means of these times were excluded. This approach avoids the imposition of an absolute limit on between subject choice times, while removing those observations that were larger than the central 95% of each subject's sample.[17]

Log decision time was the dependent variable in this weighted GLS regression. The inverse of the respondent's average estimated standard errors $\bar{s}_i^2 \frac{1}{T} \sum_{i=1}^{T} \bar{e}_{ij}$ from an OLS model were used as the weights. This simple log-time regression equation uses a linear in parameters form, where the explanatory variables include (a) the cubic relation over the fitted similarity of the risky choice pair (from the GLS similarity regression reported in Table 7), (b) the question order, and (c) personal indicator (dummy) variables allowing individual differences in the intercepts. This model was run to test Hypothesis H1a.

The results of this GLS regression for log time are given in Table 9.[18] The cubic relation on the fitted similarity of the risky choice pair had a significant effect on the log of evaluation and decision time; an F-test on this cubic relationship showed significance at less than the .001 level. The relationship between similarity and time is maximized at a moderately dissimilar level of 40 and has local minima at both the extremely similar value of 100 and at the extremely dissimilar value of 0.[19] This estimation supports the hypothesized concave relationship of evaluation time to similarity discussed previously and illustrated in Figure 3. The model fit from this partial model is shown by an adjusted R^2 measure of .44.

Similarity and Evaluation Time: Advanced Model. The results in Table 9 allow for a simple test for the effects of similarity on time as hypothesized in H1a. A regression was run to determine more completely the nature of the effects of similarity as hypothesized in H2a. Namely, testing if the characteristics of the alternatives were treated differently in choice depending on the alternatives' similarity. The underlying model for this regression is equation (12).

TABLE 9:
Generalized Least Squares for Log Evaluation Time: Simple Model

Variable Name	Estimated Coefficient	Standard Error	T-Ratio 7994 DF
Estimated Similarity	0.074	0.414E−02	17.860
Estimated Similarity Squared	-0.124E−02	0.104E−03	-11.940
Estimated Similarity Cubed	0.524E−05	0.712E−06	7.371
Question Order	0.186E−02	0.912E−04	20.370
Constant	0.791	0.106	7.485

$R^2 = 0.4451$ Adjusted $R^2 = 0.4421$

Evaluation time recorded in seconds.

Thirty-five (35) of the 39 indicator (dummy) variables had t-values above 3.00.

Log decision time was again the dependent variable in this more advanced GLS regression model whose variables include (a) the estimated similarity, (b) the differences ($z_i = px^i - qy^i, i \varepsilon \{1, 2, 3\}$) over the polynomial outcome/probability products, (c) the interaction terms of these variables with dissimilarity (z_1*s through z_3 * s) from the outcome and probability terms, (d) the order of the question, and (e) a personal indicator (dummy) variable for the individuals. The weights for the estimated covariance matrix in this GLS model are again the inverse of the respondents' average estimated standard error $\ddot{s}_i^2 \frac{1}{T} \sum_{i=1}^{T} \ddot{e}_{ij}$ from a corresponding OLS model.

The results of this advanced GLS model are given in Table 10. In this table, the coefficient estimates of the individual indicators (dummy variables) are not reported, although 20 of the 39 variables had absolute t-values above 3.00. The coefficients on the outcome and probability difference terms not adjusted for similarity effects (z_1 through z_3) give some evidence that utility is a concave relationship. Joint tests of the coefficients on the variables z_1 through z_3 (differences of these coefficients from zero) show that their effect jointly differs significantly from zero at the .005 level.

Most of the influence of similarity on choice was through the estimated similarity term, with no significant (single or joint) effects of similarity on the effects of the moments as hypothesized in H2a. The significant coefficient estimates on the question order term indicates that evaluation time increased as the respondents moved through the survey; respondents took longer to choose as the survey progressed.[20] The overall fit of this regression was an Adjusted R^2 measure of .41.

A concern that arises in light of the variables included in the model is the potential for excluded variables to influence the coefficients on the similarity level and interactive terms. Specifically, if individual differences for the coefficients z_i, i = {1, 2, 3} terms defining the utility function u(x, α) were used, would the similarity effects still be statistically different from zero? A regression was run to evaluate these questions and showed through an F-test that the interaction terms (sz_i) with similarity were still significant at less than the .00001 level. Interestingly, there was only a very slight increase in model fit from that for the model in Table 10 as evidenced by the adjusted R^2 measure of .46.[21]

The Effects of Similarity on Risky Choice. Tests for the effects of similarity on choice were carried out in a discrete choice statistical model through a probit regression. This model used the fitted similarity judgements from the GLS regression discussed previously (Table 7). The general equation supporting this linear in parameters model for the dependent variable defined as the likelihood of selecting the riskier alternative from equation (13).

In this regression model, differences were allowed for the intercepts of the probit equations through indicator (dummy) variables for each respondent. This approach allows for differences in the overall level of risk preference for each individual.

TABLE 10:
Generalized Least Squares for Log Evaluation Time:
Advanced Model

Variable Name	Estimated Coefficient	Standard Error	T-Ratio 7990 DF
Estimated Similarity	0.008	0.103E−02	7.487
$Z_1 = px - qy$	0.011	0.426E-02	2.512
$Z_2 = px^2 - qy^2$	-0.440E−03	0.175E−03	-2.515
$Z_3 = px^3 - qy^2$	0.156E−04	0.471E−05	3.328
Interaction: SZ_1	0.993E−04	0.644E−04	1.541
Interaction: SZ_2	-0.256E−05	0.277E−05	-0.926
Interaction: SZ_3	0.266E−07	0.766E−07	-0.347
Question Order	0.166E−02	0.935E−04	17.780
Constant	2.211	0.087	25.540

$R^2 = 0.4175$ Adjusted $R^2 = 0.4141$

Evaluation time recorded in seconds.
Twenty (20) of the 39 indicator (dummy) variables had t-values above 3.00.

TABLE 11:
Probit Regression for Risky Choice, Dummy Intercepts

Variable Name	Estimated Coefficient	Standard Error	N: 8461
Estimated Similarity	0.03289	0.198E–02	16.623
$Z_1 = px - qy$	-0.20856	0.02041	-10.217
$Z_2 = px^2 - qy^2$	0.597E–02	0.693E–03	8.304
$Z_3 = px^3 - qy^3$	-0.956E–04	0.155E–04	-6.1674
Interaction: SZ1	0.669E–03	0.314E–03	2.1290
Interaction: SZ2	-0.589E–04	0.109E–04	-5.4145
Interaction: SZ3	0.629E–06	0.254E–06	2.4750
Question Order	0.309E–03	0.279E–03	1.1078
Constant	-0.7878	0.08896	-8.8558

Log-likelihood (0) = -5541.3 Log-likelihood function = –901.0

Likelihood Ratio Test = 3280.60 with 47 D.F.

PREDICTION SUCCESS TABLE

		ACTUAL	
		0	1
	0	2139	688
PREDICTED	1	1116	4518

PERCENTAGE OF RIGHT PREDICTIONS = 0.78

Dependent variable equals 1 if the riskier (row) alternative was selected.

Thirty-three (33) of the personal indicator variables had t-values above 3.00.

The primary results of this probit regression are given in Table 11 (no personal intercept dummy coefficients are reported). Utility is estimated significantly for a cubic relationship on the coefficients for $z_i = px^i - qy^i$, $i \varepsilon \{1, 2, 3\}$ through a joint test at less than the .001 level for a cubic relationship of utility on the outcomes. Fitted similarity also has a positive and significant effect on the coefficients through a joint test for the terms (sz_i) at less than the .001 level.

Hypothesis H2b predicts that the higher order (second and third) polynomial terms will have less effect on choice as the alternatives become more similar. To test this hypothesis, consider the estimates of the relative differences from similarity effects for the coefficient on the z_i measures and the sz_i interaction terms. Recall from equations (9) through (12) that the coefficients on the z_i terms are Θ_i and the coefficients on the sz_i terms are γ_i. The statistics used to test for the effects of similarity as hypothesized in H2a are: $m_{12}=[(\gamma_1 + \theta_1)/\theta_1 - (\gamma_2 + \theta_2)/\theta_2]$, $m_{13} = [(\gamma_1 + \theta_1)/\theta_1 - (\gamma_3 + \theta_3)/\theta_3]$. These estimates measure the change in the relative influence of the cubic terms as similarity increases.

Similarity significantly reduced the influence of the mean differences ($z_1 = px - qy$) on choice relative to that for the higher moments through the interactive terms sz_i. F-tests of the change in the influence of z_1 on choice relative to the influence z_2 and z_3, respectively, showed a significant (at less than the .001 level] decrease in the influence of z_1 relative to z_2 and z_3. The statistics used for these tests were again over the estimates: $m_{12} = [(\gamma_1 + \theta_1)/\theta_1 - (\gamma_2 + \theta_2)/\theta_2]$ and $m_{13} = [(\gamma_1 + \theta_1)/\theta_1 - (\gamma_3 + \theta_3)/\theta_3]$. These statistics had the respective estimates -.00094 (asymptotic z-value -7.20) and -.0015 (asymptotic z-value -2.16), jointly significant at less than the .001 level. These test results are consistent with Hypothesis H_{2b} versus a null of equal effects of similarity on these polynomial terms. Although omitted from the table, 33 of the 39 personal indicators (dummies) for the intercept had asymptotic t-values above 3.00.

This probit model had an overall fit of 78% correct predictions, with symmetric prediction success for both the riskier and the less risky choice populations.

A probit model for risky choice was also carried out allowing for individual differences in coefficients for the z_i terms. Tests showed the continued significance of the interaction terms (sz_i) of the risky choice characteristics with similarity. The number of correct predictions in this model increased somewhat to 84%.

VII. Implications and Forthcoming Analysis

This paper reports the results of a study designed to extend testing of the similarity model as developed in previous literature. In addition to testing for the effects of similarity on choice patterns, a new model and testing approach was used to assess the effects of question similarity on evaluation and decision effort, where this effort was measured through the proxy of the time taken for the decision.

This paper made use of a well-known experimental design from studies in psychology, where respondents were asked to select between and report judgements over a set of risky choice pairs. Perceptions of risky pair similarity were fit using a Generalized Least Squares (GLS) regression over characteristics of the pairs.

Fitted similarity judgments from the GLS regression were used to test hypotheses regarding (a) the effects of similarity on decision effort (time), (b) the effects of similarity on choice error, and (c) the effects of similarity on choice patterns. The results of these tests indicate that fitted similarity had a significant effect on decision time in a manner consistent with optimal effort allocation. The importance of the first order difference term (px – qy) relative to higher order terms for decision time did not depend significantly on similarity.

Fitted similarity was used to evaluate choice behavior in a probit model. More similar pairs of risky alternatives carry an increased likelihood of selection for the riskier alternative. As the alternatives become more similar, the effects of the higher order moments of the alternatives on choice increase relative to the mean.

The significance of pair similarity on decision time and choice patterns give further evidence of the dependence of observed behavior on the nature of the experiment itself. Further, this paper shows the promise of methods that "correct for" these perceptual effects, either through construction of survey designs that avoid them, or through statistical regression analysis over the factors that determine them.

The findings in this paper support the view that an EU representation through a cubic utility representation may well be an accurate and useful framework for *preferences* since the evidence against EU from *choices* is seen only in combination with the respondents' perceptions and opportunity costs of the decision. This paper's results also point to the importance of further research for the effects of elicitation method and structure when analyzing behavioral patterns of choice.

Endnotes

1. Rubinstein further explores another intra-dimensional similarity definition, using relative differences (p/q and x/y) to define similar risky alternatives. This definition is not further explored as it cannot explain well-known patterns of EU violations such as Kahneman and Tversky's Certainty Effect example.

2. The similarity model is also consistent with choice patterns where no EU violation occurs (e.g., for choice patterns (A, C) or (B, D)).

3. Defined through the number of calculations required to evaluate the alternatives.

4. David Zilberman suggested this model for introducing and modeling evaluation strategies.

5. This approach abstracts from an important argument by Payne, Bettman, and Johnson that the quality of decision effort, not just the quantity, is important.

6. The error-free expression B(·) could take the properties of Fishburn's (1988) $\phi(p, q)$ function in the SSB model or Loomes and Sudgen's regret measure instead of the stronger assumption of the EU difference used here.

7. Lipman's title itself ("How to Decide How to Decide How to Decide ...") itself reflects the non-separable nature of the decision effort problem.

8. The Implicit Function Theorem would require differentiability throughout the range of the parameters for $\phi(\cdot)$ and with $\int_{\varepsilon_B} \beta(\cdot)\frac{\delta^2(\cdot)}{\delta\tau^2}\,d\varepsilon_B + \frac{\partial^2 c(\cdot)}{\partial\tau^2} \neq 0.$

9. The risky alternatives in this study have quite small payoffs (maximum outcome is $56.70), while the maximum EV level is $30.62. Also, for some pairs, one alternative has a higher variance and larger third moment with lower EV than does the other risky alternative. For other pairs, one alternative has both a higher EV and larger higher order moments.

10. As illustrated in Figure 4, the probabilities were given both visually and numerically.

11. Mid-way through the experiment, permission to conduct 2-hour rather than only 1-hour experiments was obtained, allowing the subjects to complete the experiment.

12. There were three clocks set up in the fortran program used in this experiment. The first began when the risky pair was displayed and stopped when the choice between the pair was made. The second began after the first ended and stopped after the similarity ranking was made. The third began after the second ended and stopped after the strength of preference ranking was made. Only the results from the first clock will be discussed in this paper.

13. The divisor (T) was not corrected to give an unbiased estimate of the standard deviation since T (the number of questions (usually 225) faced by each respondent) was quite large.

14. The survey order was completely random and each respondent received a different pattern for their survey.

15. This statistic is given by $t > [(T - K)(T^{1/T} - 1)]^{1/2}$, where $T = 8417$ and $K = 44$.

16. Questions #6 to #10 were also repeated as the final questions (#231-235) in the study. Since many respondents did not complete the study, choice error for these questions was not analyzed.

17. Models run on 99% and 100% of the sample showed qualitatively similar results for the coefficient estimates as those reported here. Overall model fit (measured by R^2) was lower than for the 95% model.

18. No degrees of freedom adjustment was made for the estimation of the covariance matrix.

19. The maximizing level was determined by using the equation for the roots of a quadratic equation to solve the first order condition.

20. However, some of this effect might be as a result of not completely removing the effects of rest periods by the respondents.

21. This GLS regression allowed for individual differences in the intercepts of the regressions, for the influence of the probability and outcome difference terms as in the previous regressions, and for the interactive terms of these outcome or probability differences with estimated similarity. The R^2 measure did not increase measurably and F-tests of model fit showed no significant difference from the model in Table 10. Again, the coefficients of the similarity measures (cubic relationship) and of the interactive terms were significant for time at better than the .005 level. Additional details for this regression are available from the author upon request.

References

Azipurua, J. M., T. Ishiishi, J. Nieto, and J. R. Uriarte. 1993. "Similarity and Preferences in the Space of Simple Lotteries." *Journal of Risk and Uncertainty* 6: 289-297.

Biggs, S. F., J. C. Bedard, B. G. Gaber, and T. J. Linsmeier. 1985. "The Effects of Task Size and Similarity on the Decision Behavior of Bank Loan Officers." *Management Science* 31: 970-987.

Buschena, D. E. 1992. "The Effects of Alternative Similarity on Choice Under Risk: Toward a Plausible Explanation of Independence Violations of the Expected Utility Model." Ph.D. dissertation, University of California at Berkeley.

Busemeyer, J. R., and J. T. Townsend. 1993. "Decision Field Theory: A Dynamic-Cognitive Approach to Decision Making in an Uncertain Environment." *Psychological Review* (forthcoming).

Conlisk, John. 1988. "Optimization Costs." *Journal of Behavior and Organization* 9: 213-228.

Dashiell, J. F. 1937. "Affective Value-Distances as a Determinant of Aesthetic Judgement-Times." *American Journal of Psychology* 50: 57-67.

Edwards, W. 1954. "The Theory of Decision Making." *Psychological Bulletin* 51: 380-417.

Edwards, W. 1961. "Behavioral Decision Theory." *Annual Review of Psychology* 12: 473-498.

Encarnaciôn, J. 1987. "Preferences and Lexicographic Choice." *Journal of Economic Behavior and Organization* 8: 231-248.

Fishburn, Peter C. 1988. *Nonlinear Preference and Utility Theory*. Baltimore: John's Hopkins University.

Friedman, D., and L. J. Savage. 1989. "The S-Shaped Value Function as a Constrained Optimum." *American Economic Review* 79: 1243-1248.

Heiner, Ronald A. 1988. "Imperfect Decisions in Organization: Toward a Theory of Internal Structure." *Journal of Economic Behavior and Organization* 9: 25-44.

Jensen, N. E. 1967. "An Introduction to Bernoullian Utility Theory: I. Utility Functions." *Swedish Journal of Economics* 69: 163-183.

Judge, G. A., W. E. Griffiths, R. C. Hill, H. Lütkepohl, and T. C. Lee. 1985. *The Theory and Practice of Econometrics*. 2d ed. New York: John Wiley and Sons.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica* 47: 263-291.

Leland, J. W. 1992a. "Choice Paradoxes as Decision Errors." Working Paper, Department of Social and Decision Sciences, Carnegie Mellon University. Pittsburgh, Pennsylvania.

Leland, J. W. 1992b. "Generalized Similarity Judgments." Working Paper, Department of Social and Decision Sciences, Carnegie Mellon University. Pittsburgh, Pennsylvania.

Leland, J. W. 1992c. "Similarity Judgments, Intransitivities and the Preference Reversal Phenomenon." Working Paper, Department of Social and Decision Sciences, Carnegie Mellon University. Pittsburgh, Pennsylvania.

Leland, Jonathan W. 1990. "A Theory of Approximate Expected Utility Maximization." Working Paper, Department of Social and Decision Sciences, Carnegie Mellon University.

Lipman, Barton L. 1991. "How to Decide How to Decide How to...: Modeling Limited Rationality." *Econometrica* 59: 1105-1125.

Luce, R. Duncan. 1956. "Semi-orders and a Theory of Utility Discrimination." *Econometrica* 24: 178-191.

March, James G. 1978. "Bounded Rationality, Ambiguity and the Engineering of Choice." *Bell Journal of Economics* 9: 587-608.

Mellers, Barbara, Shi-jie Chang, Michael H. Birnbaum, and Lisa D. Ordôñez. 1991. *Preferences, Prices and Ratings in Risky Decision Making*. Draft, Department of Psychology Working Paper, U.C. Berkeley.

Mellers, Barbara, Lisa D. Ordôñez, and Michael H. Birnbaum. 1992. "A Change of Process Theory for Contextual Effects and Preference Reversals in Risky Decision Making." *Organizational Behavior and Human Decision Processes* 18: 347-361.

Mosteller, F., and P. Nogee. 1951. "An Experimental Measurement of Utility." *Journal of Political Economy* 59: 371-404.

Moyer, R. S., and S. T. Dumais. 1978. *Mental Comparison*. In The Psychology of Learning and Motivation, ed. G. H. Bower, 117-155. New York: Academic Press.

Newbery, D. M., and J. E. Stiglitz. 1985. *The Theory of Commodity Price Stabilization: A Study in the Economics of Risk.* 2d ed. Oxford: Oxford University Press.

Ng, Yew-Kwang. 1975. "Bentham or Bergson? Finite Sensitivity, Utility Functions and Social Welfare Functions." *Review of Economic Studies* 42: 545-569.

Payne, J. W., J. R. Bettman, and E. J. Johnson. 1993. *The Adaptive Decision Maker.* Cambridge: Cambridge University Press.

Petrusic W. M., and D. G. Jamieson. 1978. "Relation Between Probability of Preferential Choice and Time to Choose Changes with Practice." *Journal of Experimental Psychology: Human Perception and Performance* 4: 471-82.

Poirier, Dale. "Intermediate Statistics and Econometrics: A Comparative Approach." Forthcoming text.

Pratt, J. W., H. Raiffa, and R. Schlaifer. 1965. *Introduction to Statistical Decision Theory*. New York: McGraw-Hill.

Radner, R. 1975. "Satisficing." *Journal of Mathematical Economics* 2: 253-262.

Rubenstein, Ariel. 1988. "Similarity and Decision Making Under Risk: Is There a Utility Theory Resolution to the Allais Paradox?" *Journal of Economic Theory* 46: 145-153.

Shugan, S. M. 1980. "The Cost of Thinking." *Journal of Consumer Research* 7: 99-111.

Simon, Herbert A. 1960. *The New Science of Management Decision*. New York: Harper and Brothers.

Smith, V. L., and James M. Walker. 1993. "Monetary Rewards and Decision Cost in Experimental Economics." *Economic Inquiry* 31: 245-261.

Stigler, G. 1970. *The Theory of Price*. New York: MacMillan.

Stone, D. N., and D. A. Schkade. 1991. "Numeric and Linguistic Information Representation in Multiattribute Choice." *Organizational Behavior and Human Decision Processes* 49: 42-59.

von Neumann, J., and O. Morganstern. 1953. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 3rd ed.

Wilcox, Nathaniel. 1992. *Lottery Pricing: Incentives, Complexity, and Decision Time*. Working Paper, Department of Economics, University of Houston.

Wilcox, Nathaniel. 1992. *On the Lottery Pricing Anomaly: Time Tells the Tale*. Working Paper, Department of Economics, University of Houston.