DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

AN EXPERIMENTAL ANALYSIS OF THE TWO-ARMED BANDIT PROBLEM

Jeffrey Banks
University of Rochester

Mark Olson
University of Amsterdam

David Porter
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 892

October 1994

AN EXPERIMENTAL ANALYSIS OF THE TWO-ARMED BANDIT PROBLEM

Jeffrey Banks Mark Olson David Porter

Abstract

We investigate, in an experimental setting, the behavior of single decision makers who at discrete time intervals over an "infinite" horizon may choose one action from a set of possible actions where this set is constant over time, i.e. a bandit problem. Two bandit environments are examined, one in which the predicted behavior should always be myopic (the two-armed bandit) and the other in which the predicted behavior should never be myopic (the one-armed bandit). We also investigate the comparative static predictions as the underlying parameter of the bandit environments are changed. The aggregate results show that the cutpoint behavior in the two bandit environments are quantitatively different and in the direction of the theoretical predictions. Furthermore, while a a significant number of individual cutpoints exhibit nonstationarity (contrary to the theory), the most likely, i.e. maximum likelihood estimates, collection of decision rules that best explain overall behavior are those that are consistent with the underlying theory.

AN EXPERIMENTAL ANALYSIS OF THE TWO-ARMED BANDIT PROBLEM

Jeffrey Banks Mark Olson David Porter

1 Introduction

Models of search and learning have become quite pervasive in the field of microeconomics in the last two decades. For example, in the area of labor economics the "matching" models of Jovanovic (1979), Wilde (1979), and Viscusi (1979) all have, as their basic scenario, a worker who periodically receives information about her current job's true characteristics (and hence the wages she can expect), and has the opportunity to remain with her current employer or switch to a new job where again information about future wages may accrue.[1] In industrial organization, Rothschild (1974) models a monopolist attempting to learn the true state of consumer demand for its product: by setting different prices, the monopolist can gain differential information about demand. Ishikida (1992) models the decentralized assignment of a digital pipe as a search process where users randomly send packets and can cause delays to other users of a communication network.

Many of these models share an underlying canonical form of individual decision making under uncertainty, namely that of a *bandit problem.*[2] The structure of a typical bandit problem has a single decision maker, who at discrete time intervals over an infinite horizon may choose one action from a set of possible actions (i.e. the *arms* of the bandit), where this set is constant over time. Each arm, if employed in a period, generates a reward to the decision maker according to some time-invariant distribution, where for each arm there is a set of possible reward distributions known as the *types* of the arm. The decision maker begins with some prior belief about an arm's true type, and any additional information would be useful in ascertaining the best arm to play. Further, it is commonly assumed that the true types of different arms are independent; hence knowledge of an arm's type can only be generated by employing the arm for at least a single trial. Finally, the decision maker is interested only in maximizing the discounted sum of her expected rewards.

[1] Cf. Mortensen (1985) for an in-depth survey of these and other search models in labor economics.

[2] Cf. Banks and Sundaram (1992b).

Given the temporal stationarity of the above decision problem, there is an optimal strategy for the decision maker which depends only on her current belief (i.e. probability distribution over types) about the different arms. This stationarity can equivalently be thought of as a particular type of *path independence*, in that any two paths of reward realizations leading to the same updated belief should generate the same behavior by an optimizing individual. Indeed, a result by Gittins and Jones (1974) for finite-armed bandits shows that an *index* strategy is actually optimal: for each arm, and for each possible belief about that arm's true type, one can assign a number depending only on the characteristics of that arm such that an optimal strategy simply prescribes selecting in any trial the arm with the highest number.[3]

The presence of such an index, while highlighting the stationarity of the solution to the decision maker's problem, also renders certain types of comparative statics exercises more tractable. For instance, by changing one distribution associated with an arm, or more simply by changing the prior belief associated with an arm, one can deduce how this alters the index and consequently the index strategy. This stationarity of the optimal strategy, along with certain comparative statics hypotheses, provide the motivation for our experimental investigation.

This paper describes experiments in which individuals were faced with one of two relatively simple bandit problems similar to the one detailed above. In the first, labeled Bandit Problem I below, there are two arms, one of which pays a *certain reward*, while the other generates either a high or low reward, and where the probability of a high reward can take on one of two possible values: a "high" probability of generating a high payoff (which we refer to as the "good" type) and a "low" probability of generating a low payoff (the "bad" type). Thus, we have a two-armed bandit, with one arm having a single type while the other has two types. In Bandit Problem II there are again two arms, but now both arms generate uncertain payoffs. Each arm again produces either a high or low reward, and each can be either a "good" or "bad" type, where the true types are drawn independently. Therefore the only difference in the arms is the decision maker's current belief about each arm's type, i.e. probability she assesses to each arm being "good".

In both of these problems the predictions implied by the theory alluded to above are straightforward: in Bandit Problem I, once an individual has begun playing the arm producing the certain payoff, she should remain there forever, since no new information is being generated and hence the indices of the arms remain unchanged. Alternatively, if she employs the uncertain arm and a high payoff results she should remain with that arm, since such a high-payoff should lead her to increase the expected benefits from continual play of that arm; that is, the index on the uncertain arm increases. More generally, the theory predicts that an individual will have a *critical belief* about the uncertain arm, wherein if her current belief about the uncertain arm being "good" is above this level she should employ that arm in the current period, while if her belief is below this level she

[3]This result has been extended to the case of a countable infinity of arms by Banks and Sundaram (1992b).

should play the certain arm. Further, this critical belief will be a function of the specific reward probabilities, as well as an individual's discount factor and attitudes towards risk.

In Bandit Problem II, the theoretical prediction is considerably more stark: the optimal strategy for an individual, regardless of the reward probabilities, her discount factor or her attitudes towards risk, is *myopic*; that is, in any period the choice of arm consistent with maximizing the discounted sum of expected payoffs is simply the arm that generates the highest expected one-period reward; or equivalently the arm with the higher probability of being "good."[4] The intuition behind this result is readily apparent given that index strategies are optimal: when there are only two possible types, the index on an arm will be increasing in the probability the arm is the "good" type; and since in Bandit Problem II the arms are identical up to the current belief about type, the indices will be the same up to this belief as well. Therefore the arm with the higher probability of being "good" will have the higher index, and hence will constitute the optimal choice.

This predicted myopia in Bandit Problem II is in contrast to the behavior predicted in Bandit Problem I, where if an individual's belief is such that either the certain or the uncertain arm would generate the same expected one-period payoff, the optimal strategy always prescribes the uncertain arm. The reason for this is the "option value" or learning aspect inherent in the uncertain arm: even if playing this arm gives the same payoff as the certain arm today, it may, in addition, generate information about the arm which will be useful tomorrow, information which (by definition) is not generated by playing the certain arm. In this sense, then, behavior in Bandit Problem I should *never* be myopic, whereas behavior in Bandit Problem II should *always* be myopic.

The experimental literature related to "bandit type problems" has focused primarily on the Bayesian updating assumption. Search experiments (e.g. Cox and Oaxaca (1989, 1991), Harrison and Morgan (1990)) have an agent's information unaffected by an action {search, stop searching} unless the agents belief of the initial prior distribution of rewards is misspecified, (e.g. an agent has a prior with support (0,1) but observes an outcome of 2). But the models of search (and the experiments that follow them) suppose that an agent knows the correct distribution and does not update beliefs. Grether (1992) studied Bayesian updating and other heuristics that agents may use in making decisions under uncertainty. His results suggest that "in making judgements under uncertainty individuals use different decision rules in different decision situations" and what we want to discover as economists are the variables or factors in terms of which decision strategies are stable.[5] In the experiments we design, the focus is on the stationarity of strategies and the comparative static properties of the variables that theory predicts will influence behavior for relatively simple bandit problems.

The next section provides a more formal description of our Bandit problems along with the predictions and notation we will use in analyzing data from our experiments. We then present the experimental design and results.

[4]Cf. Banks and Sundaram (1992a)

[5]Recently there has been an emerging literature concerning Bayesian learning in games. However, this literature focuses on learning how to play a game against others.

2 Description of the 2-Armed Bandit Problems

A decision maker, referred to as the principal, selects an option or arm $a \in \{A, B\}$, available at time $t = 1, 2, 3, \ldots$. After choosing an arm a reward is obtained which is a realization from a distribution with a possibly unknown parameter.

In Bandit Problem I arm A pays a constant amount of 50 each time it is selected, whereas option B's reward is uncertain. Specifically, B can be one of two possible types, good (G) or not (N), where B's type is the same for each time t but is unknown ex ante by the principal; let p be the principal's prior belief that B is good. Type G generates a reward of 100 with probability g and 0 with probability $1 - g$, whereas a bad type generates 100 and 0 with probabilities n and $1 - n$, respectively, where $g > n$. Assume the principal has a per period utility function, updates her beliefs in a Bayesian fashion, discounts rewards by the factor $\delta \in (0, 1)$, and seeks to maximize discounted expected utility. An optimal strategy exists for this problem; indeed the following simple characterization describes the **stationary** optimal strategy:

> There exists a $p^* \in (0, 1)$ such that at time t if the updated prior p_t concerning arm B is greater than p^* the principal selects arm B; otherwise the principal selects arm A.

Bandit Problem II is the same as I, except that now A is structurally similar to B, *viz.* A can be one of two types, G and N (where these are the same possible types as for B), where the true types of the two arms are drawn independently. Let p_0^j be the prior belief $j \in \{A, B\}$ is a good type and p_t^j is the current belief about j being a good type. Then the optimal strategy is characterized as follows:

A (resp. B) is an optimal selection at time t if and only if $p_t^A \geq p_t^B$ (resp. $\leq$).

In both Bandit Problem I and Bandit Problem II the optimal decision rule depends on the principal's updated beliefs concerning whether the uncertain arm is good or bad, which given the setup here is a function of the number of 1's and 0's an arm has generated. To make this updating relatively transparent, in the experiments below, we assume that $g = 1 - n$; under this assumption, Bayesian updated beliefs will depend only on the *difference* between the 1's and 0's generated. With this in mind, define $\mathcal{C}(B)$ inductively as follows:

$$\mathcal{C}(B_o) = 0$$

$$\mathcal{C}(B_t) = \begin{cases} \mathcal{C}(B_{t-1}) + 1 & \text{if B selected at t-1 and high payoff observed} \\ \mathcal{C}(B_{t-1}) - 1 & \text{if B selected at t-1 and low payoff observed} \\ \mathcal{C}(B_{t-1}) & \text{if A selected at t-1} \end{cases}$$

For Bandit Problem II

$$\mathcal{C}(B_t) = \begin{cases} \mathcal{C}(B_{t-1}) - 1 & \text{if A selected at t-1 and high payoff observed } \underline{\text{or}} \\ & \text{B selected at t-1 and low payoff observed} \\ \mathcal{C}(B_{t-1}) + 1 & \text{if A selected at t-1 and low outcome } \underline{\text{or}} \\ & \text{B selected at t-1 and high outcome} \end{cases}$$

Using $\mathcal{C}(B_t)$ we can restate the implications provided in the previous subsection. Let $\sigma_t : \mathcal{C} \rightarrow \{A, B\}$ denote the strategy based on counts; then

(i) For Bandit Problem I there exists a **critical belief cut point** $c \in \{\ldots, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, \ldots\}$

$$\sigma_t = \begin{cases} B & \text{if } \mathcal{C}(B_{t-1}) > c \\ A & \text{if } \mathcal{C}(B_{t-1}) < c \\ A \text{ or } B & \text{if } \mathcal{C}(B_{t-1}) = c \end{cases}$$

(ii) For Bandit Problem II the optimal strategy is

$$\sigma_t = \begin{cases} B \text{ if} & \mathcal{C}(B_{t-1}) > 0 \\ A \text{ if} & \mathcal{C}(B_{t-1}) < 0 \\ A \text{ or} B \text{ if} & \mathcal{C}(B_{t-1}) = 0 \end{cases}$$

The comparative statics for this problem can be computed to find, that the cutpoint c will decrease as an individual's discount rate or level of risk aversion increases and c decreases as the probability of obtaining a high reward increases.

In our experiments, we look for *stationary strategies*, that is, strategies based on counts, "more" myopic behavior in Bandit Problem II than Bandit Problem I, and the *comparative statics properties* of cutpoints with respect to discount rates, reward probabilities, and risk aversion.

3 Experimental Procedures and Design

3.1 Risk Attitudes and Infinite Horizons

From the discussion above, we note that the theory implies that the decision rule depends on risk attitudes, discounting, priors, payoffs, etc. Parameters such as the probabilities and arm payoffs can be induced and/or controlled; controlling discounting and risk is more problematic.

The lottery procedure of Berg, Daley, Dickhaut, and O'Brian (1986) (hereafter BDDO) is sometimes used in experiments to induce specific risk attitudes on subjects. BDDO is a generalization of a procedure proposed by Roth and Malouf (1979). The BDDO procedure induces any prespecified risk preference. It is a two phase decision process. First, subjects choose actions that yield "points" which are stochastically related. Second, the number of points determine the probability of winning some dollar amount in a lottery. Walker, Smith and Cox (1986) concluded that the "lottery payoff"/procedure does not seem to work for the first price auction. Results by Reitz (1993) suggest that the procedure may work if carefully applied, however, Cox and Oaxaca (1993) demonstrate that Reitz's results are inconclusive at best.

For our experiments we use the elicitation technique of Becker, DeGroot, and Marschak (1964) (hereafter BDM) to obtain certainty equivalents to see if there is a correlation between responses from the BDM procedure and decisions made in the Bandit Problems part of our experiments.

To induce discounted infinite horizons we employ a probabilistic end rule (see Camerer and Weigelt (1993)). That is, after each decision, there is a fixed and known probability that the period will end. This procedure has some problems that cause some concern. The inability of a subject to understand the probabilistic nature of the end rule (or probability at all) is possible. Second, the (small) chance that an experiment will continue for more than several hours is not credible. Nonetheless, the ability to understand probability is an attribute of the population, as such we realize that only comparative statics may be valid with this procedure. Concerning the second issue, the beliefs that subjects may have on the actual ending rule, we simply announced a specific ending rule (a maximum time for the experiment or a maximum number of periods whichever happens first).

3.2 Environment Parameters (Treatments)

For all experiments there was a choice between one of two options (arms) called A and B. In Bandit Problem I there is one certain (A) and one uncertain (B) arm and in Bandit Problem II both arms are uncertain. If the arm is uncertain then it can be one of two types called good and bad, with the prior probability of an arm being good fixed at .50. The payoffs of the uncertain arm (high and low) is fixed for all experiments (high=100 tokens low=0 tokens). The payoff for the certain arm is 50 tokens. Ten tokens were equivalent to 5 cents.

We induce discounting in our "infinite" horizon models with the probabilistic end rules with either of two **probability of continuation** values δ=0.80 and 0.90. Each subject was advised of this fact along with the *average* period length before end. These values are provided in Table 1.

[Table 1 here]

These "discount rates" were selected because δ less than 0.8 will have an expected length less than 5, and we will not be able to observe long runs. Any δ greater than 0.9

has a high probability of having a large number of rounds and thus there will likely be a number of periods that will end from the time limit rule. For δ=0.95, there is an 8% chance that the number of trials is greater than 50, and there is almost a 2% chance that the number of rounds is greater than 100.

In addition to the "discount rate," we also vary g, the probability that the good type has a high payoff (recall that we always set $n = 1 - g$). We use g=.7 and .9. Given the values of g and δ we can calculate the optimal critical belief cut point c^* for a risk neutral player (see Table 2 for a summary).

3.3 Experimental Procedures

Subjects were recruited from the student population at the University of Amsterdam. An experimental session was constructed as follows (instructions were computerized; an abbreviated set of instructions is supplied in Appendix A):

(a) Each subject was seated at an individual personal computer, the experimental program is started with a predetermined set of parameters and instructions. Each subject proceeds through the instructions for the BDM process at their own pace, and can practice as much as they like without any time constraints. Each subject's experimental session is independent of any other subject's session.

(b) The subject is then asked to answer 4 different BDM questions in which they are paid based on their response/outcome. The specific BDM lottery questions are provided in the instructions in Appendix A.

(c) After the BDM procedure the subject proceeds through the instructions for the bandit problem.

(d) Each experimental session consists of periods and rounds in a period. At the beginning of each period the state of each uncertain arm is drawn according to the probabilities that are given to the subject at the beginning of the experiment. Information concerning payoffs and probabilities was provided to subjects.

(e) A sequence of rounds is then run in each period. At the beginning of a round the subject is asked to choose an arm. Given the subjects choice and the fixed and known probabilities, the subject's payoff is drawn. A random number is then drawn to determine if the period is to continue. We incorporate a computerized roulette wheel for the random draw of the stopping rule, this feature was used to reduce subject boredom and automatic responses. If the period continues a new round is run. If the period ends then a new period begins and new state variables are drawn (i.e. the types of the arm good or bad) the subject is informed that this is being done. The subject screen is provided below:

[Figure 1 here]

(f) Each subject plays the bandit problem for a maximum amount of time (60 minutes) or a maximum number of periods (5) whichever comes first.

(g) Each subject repeats each period with the same parameters given at the beginning of the experimental session.

3.4 Experimental Design Summary

Table 2 lists the treatments, the associate optimal risk neutral cut point (c^*), posterior probability prediction that arm B is a good type, and the expected increase in payoff (D) of using the optimal strategy over a myopic strategy, the continuation probability δ, and the probability of high payoff with the good arm. We label each cell with a two digit code which we will use for distinguishing the treatments (e.g. hL means high δ, and Low g).

[Table 2 here]

4 Experimental Results

The first part of this section will investigate the best cutpoint distribution for each Bandit problem and parameter set. The *best cutpoint* for a subject will be defined as the value c that is most consistent with a subject's observed choices. A strategy is most consistent if it has the smallest number of observed deviations. For example, consider a subject in Bandit Problem II who chooses B, A, B, A, when $\mathcal{C}(B)$ is 0, 0, 0, 1 respectively. A strategy of $c = 0$ would imply that a deviation was made at the 4th choice since the $c = 0$ strategy would be to choose B. A strategy of $c = 1$ would imply that deviations were made at the 1st and 3rd choices. So the strategy $c = 0$ is the most consistent of the two strategies $c = 0$ and $c = 1$. The second part will focus on uncovering behavioral strategies used by subjects.

4.1 Aggregate Cutpoint Behavior

We investigate the behavior of the cutpoints that best describe the period "behavior" of each participant. We take these best cutpoint estimates and then investigate their stationary within and across periods, their values across the Bandit Problems and how they vary with the environment parameters. For example, Figure 2 supplies decisions made by a subject in our experiments.

[Figure 2 about here]

In period 1 of the figure we find that any cut point less than 0 describes the choice pattern. For period 2 the best cutpoint is any number between -1 and -2, while for period 3 it is -1. We will use the midpoint of the period best cutpoint intervals for each subject to construct a cutpoint distribution to investigate "stationarity."

Observation 1 : The best cutpoints change across periods.

Support : The Hotelling T2 statistic to test the hypothesis that the cutpoints are the same for each period are provided below. We use this test since we have multiple observations per subject and this test does not require independence of the observations. The table below shows that we can safely reject the hypothesis of equal cutpoints across periods in 5 of the 8 parameter sets.

Bandit Problem I

Parameter Set	*df*	F-stat	Prob level
ℓL	(4,16)	4.28	0.02
ℓH	(4,15)	1.87	0.17
hL	(4,14)	10.12	0.00
hH	(4,15)	6.98	0.00

Bandit Problem II

Parameter Set	*df*	F-stat	Prob level
ℓL	(4,10)	1.68	0.23
ℓH	(4,8)	2.53	0.12
hL	(4,7)	21.80	0.00
hH	(4,9)	8.31	0.00

Next we consider the number of choices that are consistent with the best cutpoint for each subject. Specifically, for period τ and subject i we have a best cutpoint $c^*_{i\tau}$ described by the choices made by the subject. For all n choices made by i in τ, we determine the ratio $\Omega = \frac{n^*_{i\tau}}{n_{i\tau}}$ where $n_{i\tau}$ is the number of decision made by subject i in period τ and $n^*_{i\tau}$ number of choices that coincide with the best cutpoint $c^*_{i\tau}$.

Observation 2 : The single best cutpoint within a period, which is allowed to change across periods, selects a large percent of individual choices.

Support : The table below shows the aggregate (overall periods) mean, and standard deviation for each parameter set. The pooled results are used since there is not a significant difference in the percentage of consistent choices across periods for each parameter set.

Bandit Problem I

Parameter Set	$\overline{\Omega}$	Standard Deviation	n
ℓL	.95	.11	100
ℓH	.99	.05	95
hL	.90	.15	90
hH	.94	.12	95

Bandit Problem II

Parameter Set	$\overline{\Omega}$	Standard Deviation	n
ℓL	.93	.14	70
ℓH	.92	.14	60
hL	.80	.18	55
hH	.90	.15	65

. If we now select a single best cutpoint for each subject, i.e., we select a best cutpoint c_i^* for all the selections made by subject i, we can calculate $\Theta = \frac{n_i^*}{n_i}$ where n_i is the number of decisions made by subject i and n_i^* is the number of choices predicted by c_i^*. The table below shows the descriptive statistics for Θ over all periods and the last period (period 5).

Bandit Problem I

	Period 5		All Periods	
Parameter Set	$\overline{\Theta}$	Standard Deviation	$\overline{\Theta}$	Standard Deviation
ℓL	.93	.13	.88	.12
ℓH	1.0	...	.96	.05
hL	.89	.16	.85	.13
hH	.94	.15	.92	.11

Bandit Problem II

	Period 5		All Periods	
Parameter Set	$\overline{\Theta}$	Standard Deviation	$\overline{\Theta}$	Standard Deviation
ℓL	.89	.21	.85	.15
ℓH	.93	.08	.82	.11
hL	.82	.16	.71	.15
hH	.89	.15	.84	.16

To summarize, we find that the best cutpoints are not stationary in that they tend to change across periods, however within a period the single best cutpoint estimates describe individual choices extremely well.

Observation 3 : The best cutpoints within a period decreases as the g falls, or as δ increases. The best cutpoints select a larger percentage of consistent choices in Bandit I than Bandit II.

Support : See tables used above.

Observation 4 : The distribution of cutpoints are significantly different between Bandit Problems I and II. The cutpoint distribution is consistent with myopic behavior with Bandit Problem II but not for Bandit Problem I.

Support : Figures 3 and 4 show the distribution of cutpoints and Table 3 shows the descriptive statistics for these distributions for Bandit Problems I and II. The t-statistics for the mean shows that there is a difference in cutpoints across the Bandit Problems. In addition, we can reject the hypothesis that the cutpoints are zero for Bandit Problem I but not for Bandit Problem II.

[Table 3 about here]

[Figures 3 and 4 about here]

In terms of the comparative static results, recall that for Bandit Problem II none of the parametric conditions $(\delta, g, \ldots)$ should affect the optimal cutpoint which is 0. However, for Bandit Problem I the *higher* the probability that the period continues the *lower* the cutpoint; the *higher* the probability of the good type paying high the *lower* the cutpoint; and the *less* risk averse the subject the *lower* the cutpoint (all the statements are ceteris paribus of course).

The following ANOVA model was estimated

$$CV = \alpha + \beta * P9 + \gamma * D9 + \theta * P9D9 + \psi * x + \nu * r + \epsilon$$

where

CV = median cutpoint of each subject's period best cutpoint

$$\begin{aligned} P9 &= \begin{cases} 1 & \text{if probability that good type pays high payoff is .9.} \\ 0 & \text{otherwise} \end{cases} \\ D9 &= \begin{cases} 1 & \text{if probability period continues is .9} \\ 0 & \text{otherwise} \end{cases} \\ P9D9 &= \text{interaction term for both variables} \\ x &= \begin{cases} 1 & \text{if subject is experienced} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

r is a measure for risk aversion that we derive from the responses provided by subjects in the BDM portion of the experiments. Specifically, we use the following non-parametric estimate for a subject's risk aversion

$$r_i = \#(S_{ij} > S_{rj}) - \#(S_{ij} < S_{rj})$$

where

S_{ij} = subject i's selling price for BDM lottery j
S_{rj} = risk neutral selling price of lottery j
$\#(x)$ = number of instances where x is true

Thus $r_i \in [-4, 4]$ and $r_i = 0 \Rightarrow$ risk neutrality, and larger r_i show increasing risk aversion.[6]

Observation 5 : While the direction of the parameter estimates correspond with the predictions, none of the effects are significant.

Support : Table 4 shows the estimates of the ANOVA. Only the constant term for Bandit Problem I is significant with all the parametric effects accounted for we find that the cutoff for Bandit Problem II is not significantly different than 0, but not for Bandit Problem I.

[Table 4 about here]

To summarize, the aggregate results show:

- The best cutpoint choices select a large percent of the choices made by subjects.
- The percent of choices consistent with the best cutpoint is affected positively by g and negatively by δ in Bandit Problem I.
- Cutpoint choices in Bandit Problem II are consistent with myopic behavior, while they are not in Bandit Problem I.
- The underlying cutpoint distributions are not significantly affected by the g and δ. However, the direction of change is consistent with the comparative static predictions for Bandit Problem I.
- There is no significant correlation between the decision made in the BDM mechanism and the cutpoints in the Bandit Problems.
- Experienced subjects have a slightly lower cutpoint in Bandit Problem I but show no difference in Bandit Problem II.

The analysis provided in this section examined the best cutpoints per subject and then used these "estimates" to make aggregate statements about behavior. However, this method does not consider potential strategies that are employed by subjects in making their cutpoint decisions, since each estimate is subject specific. To determine which individual behavioral strategies best describe the data we estimate the likelihood that specific strategies describe the cutpoint decisions.

4.2 Behavioral Strategies

In the analysis above, allowing the cutpoint to change each period gave a better description of the data than not allowing the cutpoint to depend on the period. This should

[6]We calculate two other measures based on the BDM responses which we supply in Appendix B. In all cases the measures are consistent with subject risk aversion. Furthermore, none of the measures affect the results of our ANOVA estimates.

be expected since we are in essence adding more variables to the estimation procedure. But is the improvement in fit an adequate trade-off for the loss of parsimony or degrees of freedom? We address this question by utilizing a procedure developed by El-Gamal and Grether (1993). Their procedure finds the most likely collection of decision rules that best explain the behavior of experimental subjects. The procedure estimates the maximum likelihood collection of rules and chooses the best set based on an information criterion that penalizes the procedure for admitting more rules.

Following their process, we construct a class of potential decision rules that could be used by subjects in our experiment. We construct a cut point decision rule as an *initial cutpoint* x, and a new *cutpoint* y in period z. Thus, the decision rules we consider for a subject consists of three parts $x/y/z$: initial cutpoint x, new cutpoint y, and switching period z. For example the decision rule -2/1/4 means the subject starts with a cutpoint of -2 and then switches to a cutpoint of 1 at the start of period 4. The rules we will search over are where $x \in \{-3,-2,-1,0,1,2\}$, $y \in \{-3,-2,-1,0,1,2\}$ and $z = \{2,4,6\}$ ($6 \Rightarrow$ initial cutpoint does not switch). In addition to these decision rules, we allow for errors (ϵ) so that the probability of not using a specified rule is part of the model; we restrict $\epsilon \in \{0.01, 0.05, 0.10, \ldots, 0.50\}$. It should be emphasized that the notion of error in this procedure describes the inability of the decision model to describe the choice, and does not necessarily reflect the errors that individuals make.

Given the rules, environment parameters, and the allowable error rates, we calculate the likelihood function for each combination of k rules. That is, we restrict the search to finding the single "best" rule, the two best rules and so forth. We estimate the maximum likelihood collection of 1 rule; we then compute the Akaike information criterion (IC) that incorporates a penalty for the number of parameters in the model (the number of allowable rules). We then increase the number of rules by 1 and compute the IC, if the IC has decreased we stop; if it has increased we increase the number of rules by 1 and continue. The set of rules with the highest IC is the best collection of rules for a particular parameter set.

The estimates for the rules and errors for each experimental treatment are given in Appendix C. The tables report for each bandit problem and each parameter set the chosen rules, the likelihood value, and the IC for each k.

We now summarize the results from the estimates found in Appendix C:

- For Bandit Problem II the best single rule for each parameter set is a stationarity myopic strategy. Furthermore, as more rules are allowed, the stationary myopic rule is always selected.
- For Bandit Problem I the best single rule is non-myopic except for parameter set hH (a myopic strategy after period 1) .
- When we allow more rules to be selected, Bandit Problem I continually admits nonmyopic strategies, with the proportion of the myopic strategy selected falling. Furthermore, more decision rules are selected for Bandit Problem I (nonmyopic) than Bandit Problem II (myopic prediction).
- In Bandit Problem I, the proportion of nonmyopic strategies increases in the direction the comparative static predictions suggest.
- For Bandit Problem I the error rate decreases as more rules are allowed to be selected; for Bandit Problem II the error rate decreases slightly as more rules are allowed to be selected.

The basic summary of results from this investigation of individual decision-rules are contained in the following observations:

Observation 6 : Behavior in Bandit Problem I is consistent with non-myopic stationary strategies. However, this non-myopic behavior has error rates of approximately 10 to 30%.

Observation 7 : Behavior in Bandit Problem II is consistent with myopic stationary strategies. However, the selected strategies have error rates of approximately 30 to 50.

5 Conclusions

Our research program presented above was very simple: design experiments to investigate individual decision-making behavior in two diverse but relatively simple bandit problems. Then, determine the pattern of behavior in the two environments relative to the following theoretical predictions:

1. Cutpoint behavior should be consistent with stationary cutpoint strategies for both environments.
2. In one of the environments behavior should be consistent with myopic behavior while the other should exhibit non-myopic behavior.
3. For various parameter choices (discount rates and probabilities of good outcomes) the comparative static predictions should be neutral in the myopic case and important in the non-myopic environment.

The experimental results suggest that there are individual violations of stationarity in cutpoint behavior but that the most likely collection of decision rules that explains subject behavior are stationary strategies. For all the measures we used there is a clear distinct pattern in the two bandit environments: behavior is more myopic in the environment in which that form of behavior was predicted. Furthermore, the most likely collection of decision rules shows this distinctive pattern convincingly. The comparative static results are not as significant but the "signs" on the estimates are as predicted.

Appendix A: Instructions

Below is an abbreviated set of the computerized instructions we used.

Screen 1:

Instructions

You are about to participate in an individual decision making experiment. The decisions you will make during the experiment will result in Dutch guilder profits that will be yours to keep. Thus, you should follow the instructions carefully to understand how you can make profits. In this experiment all values will be stated in terms of tokens. Each token you earn can be redeemed into guilders at a rate of ____ tokens per Dutch guilder. The experiment will be broken-up into two different parts in which you make decisions and earn profits.

<Press any key to continue>

Screen 2:

Instructions for Part 1

In this portion of the experiment you will be asked a series of questions. Given the answer to the questions a spin of a computerized roulette wheel will be made. Given your answer and the outcome of the roulette spin you will earn profits and proceed to part 2 of the experiment.

We will now take you through an example of how decisions made during this part translates into profits. In questions 1 and 2 a roulette wheel containing the twenty numbers 5,10,15,...,100 will be spun (all numbers are equally likely to be selected). If the wheel selects a certain set of numbers you will receive nothing, if the wheel selects any number not in that set you will receive a fixed amount of tokens. You will then be asked how much you would be willing to sell this game for. The roulette wheel will then be spun and if the number the wheel selects is equal or greater than the price you asked, you will be paid that amount. If the number the wheel selects is less than the price you ask, you will play the game

We will now go through several sample questions of this type so that you can see how it works. The outcomes of these sample questions will not count toward your profits; this is only for practice.

<Press B to go back or any other key to continue>

Screen 3:

Practice Round

In this game we will spin the roulette wheel with the twenty numbers 5,10,15,...,100. If a number less than or equal to 50 is selected you win nothing. If the number is greater than 50 you will receive 100 tokens

How much are you willing to accept instead of playing this game?

Please enter a number between 0 and 100.

Given that you are asking ____, the roulette wheel will be spun and if it selects a number that is greater than the price you are asking you will be paid that amount. Otherwise you will play the game.

** Random spin of the Roulette Wheel **

The wheel landed on the number ____ which is less than your asking price of ____. Thus, you must play the game.

** Random spin of the Roulette Wheel **

The wheel landed on the number ____ which is less than or equal to ____ so you win 0 tokens.

<Press any key to continue>

//////////////////// **************************** ////////////////////

Bandit instructions

/***/

[Bandit problem 1]

Screen 1:

Instructions for Part 2

In part 2, the experiment will be broken up into ____ periods. Each period in turn will be divided into rounds in which you will make decisions and earn profits. At the beginning of a round, you will make a choice between two alternatives called A and B. The A alternative will pay you ____ tokens if you select it. The B alternative will be one of two possible types which we will call good and bad. If B is good, you will receive ____ tokens with a specified chance and ____ tokens with a specified chance. If B is bad, then your chance of obtaining ____ tokens will be lower and the chance of obtaining ____ tokens will be higher. After you select either alternative A or B you will be informed of your payoff for the round and the roulette wheel will be spun. If the wheel selects a certain set of numbers the period will go to the next round, otherwise the period will end and we will proceed to a new period.

Before you go on to the experiment you will go through some practice periods. The outcomes of your decisions in these practice periods will not count toward your profits; they are only for practice.

Screen 2:

In round 1, of each period of this experiment, the A alternative pays ____ tokens. The B alternative could be a good type that pays ____ tokens with a ____ percent chance and ____ tokens with a ____ percent chance. This means, if B is good, then over many rounds you could expect to earn, on average ____ tokens a round. On the other hand, B could be a bad type which pays ____ tokens with a ____ percent chance and or B you will be informed of your payoff for the round and the roulette wheel will be spun. If the wheel selects a certain set of numbers the period will go to the next round, otherwise the period will end and we will proceed to a new period. We will now take you through an example of how decisions made during this part translates into profits. The outcomes of your decisions in these instructions will not count toward your profits. This is only for practice. press any key to continue

Screen 2:

Suppose we are in period 1, round 1. The alternatives A and B could be a good type that pays ____ tokens with a ____ percent chance, and ____ tokens with a ____ percent chance. This means that if A or B , is good, then over many rounds you could expect to earn, on average ____ tokens a round if you always choose that alternative. On the other hand, A or B could be a bad type which pays ____ tokens with , a ____ percent chance and ____ tokens with a ____ percent chance. This means if A or B is bad, then over many rounds you could expect to earn, on average, ____ tokens a round. Remember that you do not always have to make the same choice, at each round you may either choose A or B. You will be given information concerning the chance that A or B is a good or a bad type. Lastly, the chance that a period will continue after the current round is ____ percent. That means that, on average, in any round in a period, you could expect the period to last ____ more rounds. Further information will be handed out to you.

<Press B to go back or any other key to continue>

Screen 3:

We now summarize the specific features of this practice period:

1. At the beginning of a period, the A alternative will be selected as either a good type or a bad type. It will remain that type for the entire period.

2. At the beginning of a period, the B alternative will be selected as either a good type or a bad type. It will remain that type for the entire period.

3. The chance that alternative A is good is ____ percent.

4. The chance that alternative B is good is ____ percent.

5. If A or B is good it will pay ____ tokens with a ____ percent chance and ____ tokens with a ____ percent chance.

6. The chance that the period continues at the end of the round is ____ percent.

A sheet with this information will be given to you.

<Press B to go back or any other key to continue>

Screen 4:

Before you begin the practice period, there are several features of the program that may be helpful. At any time you can press h to see the history of your choices and the outcomes. The screen always shows the payoff chances under a good type and the payoff chances for a bad type. If you understand the process and want to practice press enter, otherwise raise your hand and a monitor will answer your questions.

<Press B to go back or any other key to practice>

____ tokens with a ____ percent chance., This means if B is bad, then over many rounds you could expect to earn, on average, ____ tokens a round. You will be given information concerning the chance that B is a good or a bad type. Lastly, the chance that a period will continue after the current round is ____ percent. That means that, on average, in any round in a period,, you could expect the period to last ____ more rounds. Further information will be handed out to you.

<Press B to go back or any other key to continue>

Screen 3:

We now summarize the specific features of this experiment:

1. Alternative A pays ____ tokens.,

2. At the beginning of a period, the B alternative will be selected as either a good type or a bad type with a fixed chance. It will remain that type for the entire period.

3. The chance that B is good is ____ percent.

4. If B is good it will pay ____ tokens with a ____ percent chance and ____, tokens with a ____ percent chance in the first iteration of each, period.

5. If B is bad it will pay ____ tokens with a ____ percent chance and ____, tokens with a ____ percent chance.

6. The chance that the period continues at the end of the round is ____ percent.,

7. Your conversion rate is 1 Guilder for ____ tokens.

<Press B to go back or any other key to continue>

Screen 4:

Before you begin the practice period, there are several features of the program that may be helpful. At any time you can press 'H' to see the history of your choices and the outcomes. The screen always shows the payoff chances under a good type and the payoff chances for a bad type.

If you understand the process and want to practice press Enter, otherwise raise your hand and an experimenter dude will answer your questions.

There will be 2 practice periods.

<Press B to go back or any other key to practice>

/***********************************
***********************************/

Bandit Problem 2

Screen 1:

Instructions for Part 2

In part 2, the experiment will be broken up into periods. Each period in turn will be divided into rounds in which you will make decisions and earn profits. At the beginning of a round, you will make a choice between two alternatives called A and B. Alternatives A and B will be one of two possible types which we will call good and bad. A and B may be the same or different types. If your choice (A or B) is good, you will receive ____ tokens with a specified chance, and ____ tokens with a specified chance. If your choice (A or B), is bad, then your chance of obtaining ____ tokens will be lower and, the chance of obtaining ____ tokens will be higher. After you select either alternative A

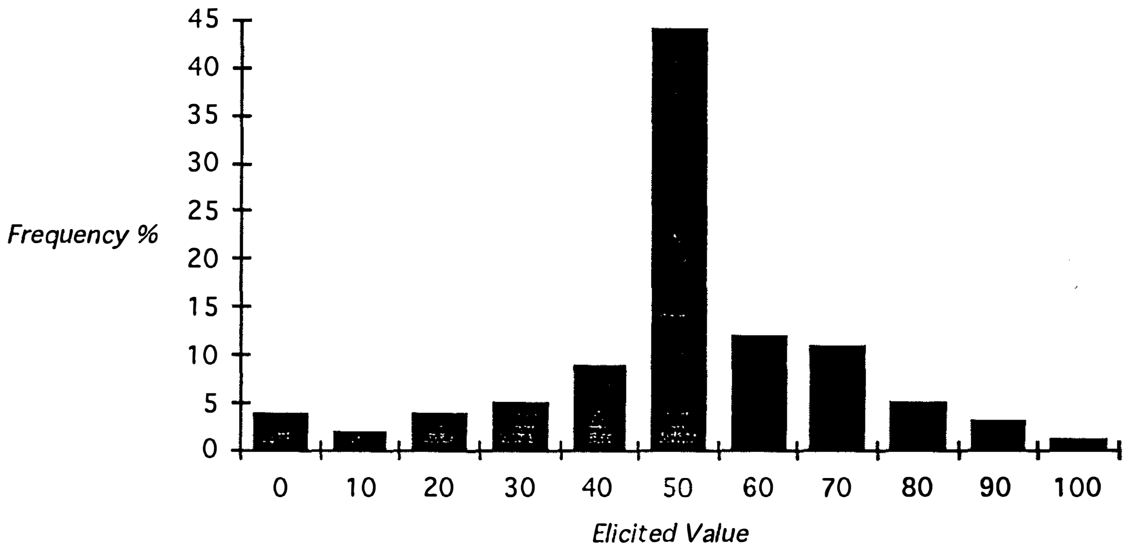
Appendix B

Three alternative measures of risk tolerance using data from the BDM procedure in our experiments are described below.

Let L_j (l_j, p_j, h_j) denote the lottery j where you receive the amount l_j with probability ($1-p_j$); you receive h_j with probability p_j .

The distribution of the selling price of the lottery L(0,.5, 100) we used in the experiment is provided below (risk neutral response = 50):

BDM Responses to L(0,.5,100)



Measures:

1. Non-parametric statistic:

For each subject i we calculate the statistic

$$r_i = \#(S_{ij} > S_{rj}) - \#(S_{ij} < S_{rj})$$

where S_{ij} = subject i's selling price for lottery j
S_{rj} = risk neutral selling price of lottery j
$\#(\cdot)$ = number of instances in which the condition $\cdot$ is true

$r_i = 0$: risk neutral behavior
$r_i > 0$: "risk loving" behavior
$r_i < 0$: "risk averse" behavior

2. Median ratio log statistic:

$r_i = \text{median}_j\,[\ln(U_{vj})/\ln(U_{sj})]$ -- {median over the j lotteries for subject i}

where

$U_{vj} = (S_{rj} - l_j)/(h_j - l_j)$ -- {normalized utility of risk neutral value of lottery j}

$U_{sj} = (S_j - l_j)/(h_j - l_j)$ -- {normalized utility of i's selling price for lottery j}

$r_i = 1$: risk neutral behavior
$r_i < 1$: risk loving behavior
$r_i > 1$: risk averse behavior

3. Median of selling prices

r_i = median of S_j selling prices for subject i

The median risk neutral selling price is 52.5. Thus $r_i < 52.5$ implies risk averse behavior and $r_i > 52.5$ shows more risk loving behavior.

The table below supplies the descriptive statistics for each of the measures described above.

Measure	mean	standard deviation	minimum	maximum	median	n
non-parametric	-.38	2.05	-4	4	-1	126
median ratio log	1.99	6.51	.13	59.9	.87	126
median of selling prices	49.9	17.92	0	100	50	126

Finally, in the table below we provide correlations between the three measures of risk attitude and the selling price of the L(0, .5, 100) lottery.

non-parametric	1.00	0.37	0.84	0.77
median ratio log	0.37	1.00	0.45	0.35
median of selling prices	0.84	0.45	1.00	0.90
L(0, .5, 100)	0.45	0.35	0.90	1.00

Appendix C

Estimates of Rule Classification Procedure

Legend

eps	=	error rate
rules	=	initial cutpoint/ delta cutpoint round switch
IC	=	information criteria

Bandit Problem I

parameter file: rate = .8 and g = .7
subjects = 20.0000

# rules	eps	rules	(#)	likelihood	IC
1	0.35	-2.0/ 0.0/ 4	20	5.77×10^{-94}	-95.36
2	0.25	0-0/ 0.0/ 6	10	3.36×10^{-78}	-81.71
		-3.0/-1.0/ 4	10		
3	0.25	0.0/ 0.0/ 6	7	5.65×10^{-74}	-79.61
		-3.0/-1.0/ 4	10		
		-3.0/ 0.0/ 2	3		
4	0.25	-3.0/-3.0/ 6	7	1.36×10^{-70}	-78.35
		0.0/ 0.0/ 6	6		
		-3.0/ 0.0/ 4	4		
		-3.0/ 0.0/ 2	3		
* * *5	0.20 1	-3.0/-3.0/ 6	5	5.53×10^{-67}	-76.86
		-2.0/-2.0/ 6	3		
		0.0/ 0.0/ 6	6		
		-3.0/-1.0/ 4	3		
		-3.0/ 0.0/ 2	3		

parameter file: rate = .8 and g = .9
subjects = 19

# rules	eps	rules	(#)	likelihood	IC
1	0.25	-1.0/-1.0/ 6	19	7.12 x 10^{-77}	-78.27
2	0.15	-3.0/-2.0/ 2	11	3.38 x 10^{-61}	-64.71
		0.0/-1.0/ 4	8		
3	0.15	-2.0/-2.0/ 6	8	1.47 x 10^{-53}	-59.19
		-1.0/-1.0/ 6	6		
		0.0/ 0.0/ 6	5		
* * *4	0.10 7	-3.0/-2.0/4	7	1.33 x 10^{-47}	-55.36
		-3.0/-1.0/ 2	6		
		-1.0/ 0.0/ 4	3		
		0.0/-1.0/ 4	3		
5	0.10	-3.0/-3.0/ 6	4	1.33 x 10^{-47}	-57.48
		-2.0/-2.0/ 6	4		
		-3.0/-1.0/ 2	5		
		-1.0/ 0.0/ 4	3		
		0.0/-1.0/ 4	3		

parameter file: rate = .9 and g = .7
subjects = 18

# rules	eps	rules	(#)	likelihood	IC
1	0.40	-1.0/ 0.0/ 4	18	1.22 x 10^{-161}	-163.04
2	0.30	-2.0/-1.0/ 2	14	3.74 x 10^{-138}	-141.67
		1.0/ 0.0/ 2	4		
3	0.30	-2.0/-1.0/ 4	13	7.03 x 10^{-133}	-138.52
		-1.0/ 0.0/ 4	2		
		1.0/ 0.0/ 2	3		
* * *4	0.30	-3.0/-2.0/4	9	7.24 x 10^{-130}	-137.62
		-2.0/-1.0/ 4	4		
		-1.0/ 0.0/ 4	2		
		1.0/ 0.0/ 2	3		
5	0.25	-3.0/-3.0/ 6	7	3.30 x 10^{-128}	-138.08
		-3.0/ 0.0/ 4	3		
		-2.0/-1.0/ 4	2		
		0.0/-1.0/ 2	3		
		1.0/ 0.0/ 2	3		

parameter file: rate = .9 and g = .9
#subjects = 19

# rules	eps	rules	(#)	likelihood	IC
1	0.25	-1.0/0.0/2	19	7.13×10^{-140}	-141.27
2	0.20	-1.0/-1.0/6	17	6.03×10^{-119}	-122.46
		-3.0/1.0/2	2		
3	0.20	-1.0/-1.0/6	13	2.33×10^{-110}	-115.99
		0.0/-3.0/4	4		
		1.0/0.0/4	2		
***4	0.20	-1.0/-1.0/6	12	1.38×10^{-105}	-113.34
		-3.0/0.0/2	3		
		-3.0/1.0/2	1		
		0.0/-3.0/4	3		
5	0.15	-1.0/-1.0/6	11	1.80×10^{-103}	-113.35
		-3.0/-1.0/4	2		
		-3.0/0.0/2	2		
		-3.0/1.0/2	1		
		0.0/-3.0/4	3		

Bandit Problem II

parameter file: rate = .8 and g = .7
#subjects = 14.0000

# rules	eps r	rules	(#)	likelihood	IC
1	0.40	0.0/0.0/6	14	7.33×10^{-68}	-69.26
2	0.35	0.0/0.0/6	13	7.96×10^{-64}	-67.34
		-1.0/-3.0/4	1		
* * *3	0.30	0.0/0.0/6	11	3.17×10^{-60}	-65.86
		-1.0/-3.0/4	1		
		2.0/-1.0/4	2		
4	0.30	0.0/0.0/6	10	1.02×10^{-58}	-66.47
		-2.0/0.0/4	1		
		-1.0/-3.0/4	1		
		2.0/-1.0/4	2		

parameter file: rate = .8 and g = .9
#subjects = 12

# rules	eps r	rules	(#)	likelihood	IC
1	0.40	0.0/0.0/6	12	7.23×10^{-64}	-65.26
* * *2	0.35	0.0/-3.0/4	2	7.00×10^{-59}	-62.40
		2.0/0.0/2	10		
3	0.35	-3.0/0.0/4	3	7.34×10^{-57}	-62.50
		0.0/-3.0/4	1		
		2.0/0.0/2	8		

parameter file: rate = .9 and g = .7
#subjects = 11

# rules	eps r	rules	(#)	likelihood	IC
1	0.50	0.0/0.0/6	11	1.98×10^{-144}	-145.82
$***2$	0.50	0.0/0.0/6	8	9.46×10^{-138}	-141.27
		2.0/-2.0/2	3		
3	0.50	0.0/0.0/6	7	2.55×10^{-136}	-141.95
		0.0/-3.0/4	2		
		2.0/-2.0/2	2		

parameter file: rate = .9 and g = .9
#subjects = 13

# rules	eps r	rules	(#)	likelihood	IC
1	0.40	2.0/0.0/2	13	2.93×10^{-112}	-113.65
$***2$	0.35	-3.0/0.0/2	5	5.02×10^{-104}	-107.54
		2.0/0.0/2	8		
3	0.35	0.0/0.0/6	8	2.37×10^{-103}	-108.99
		-3.0/0.0/2	2		
		2.0/0.0/2	3		

Table 1. Probability of Period End and Expected Number of Rounds

Chance Period Will Last	δ=.8	δ=.9
5 more rounds	32 out of 100	59 out of 100
10 more rounds	10 out of 100	34 out of 100
15 more rounds	3 out of 100	20 out of 100
30 more rounds	1 out of 100	4 out of 100
Expected # of rounds per period	5	10

Table 2. Parametric Treatment Conditions

parameter Set Name	δ	g	c	optimal Posterior	D
ℓL	.8	.7	0.5	.32	10
ℓH	.8	.9	0.5	.20	30
hL	.9	.7	-1	.23	40
hH	.9	.9	0.5	.12	60

Table 3. Best Cutpoint Estimates

Problem	Mean	Standard Deviation	Median	t-statistic
Bandit I	-.8	1.6	-.5	-5.2
Bandit II	0	1.5	0	-0.7

Table 4. ANOVA Estimates and Hypothesis Tests

Variable	Bandit Problem II	Bandit Problem II
constant	-.492*	0.009
P9	-.221	-.146
D9	-.082	.075
P9D9	.338	.110
Exper	-.212	.027
risk	.008	-.001
# of observations	76	50

* Significant at the .05 level of significance.

H_o: parameters have no effect

	Bandit Problem I	Bandit Problem II
F-Stat.	.970	.414
Prob.	.412	.744

Figure 1. Subject Screen Layout

Period : 1
Round : 3

Chance of period continuing
80%, less than 80 on wheel

A

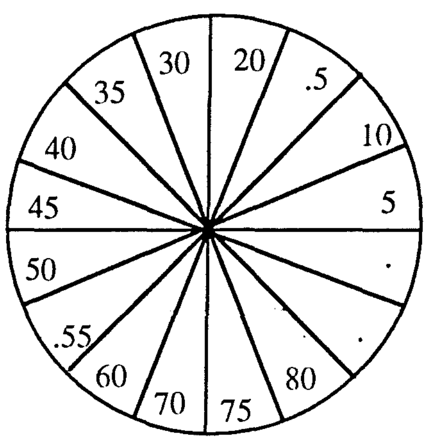
B

Payoff = 50

Type	High	Low	Chance of High
Good	100	0	60%
Bad	100	0	40%

Your selection: -
Payoff :
h for history

Chance of Being Good
50%



period lasts 5 more rounds
32 out of 100
period lasts 10 more rounds
10 out of 100
period lasts 15 more rounds
3 out of 100
period lasts 20 more rounds
1 out of 100

Figure 2:
Subject 3 Decisions
Bandit Problem I (*IH*)

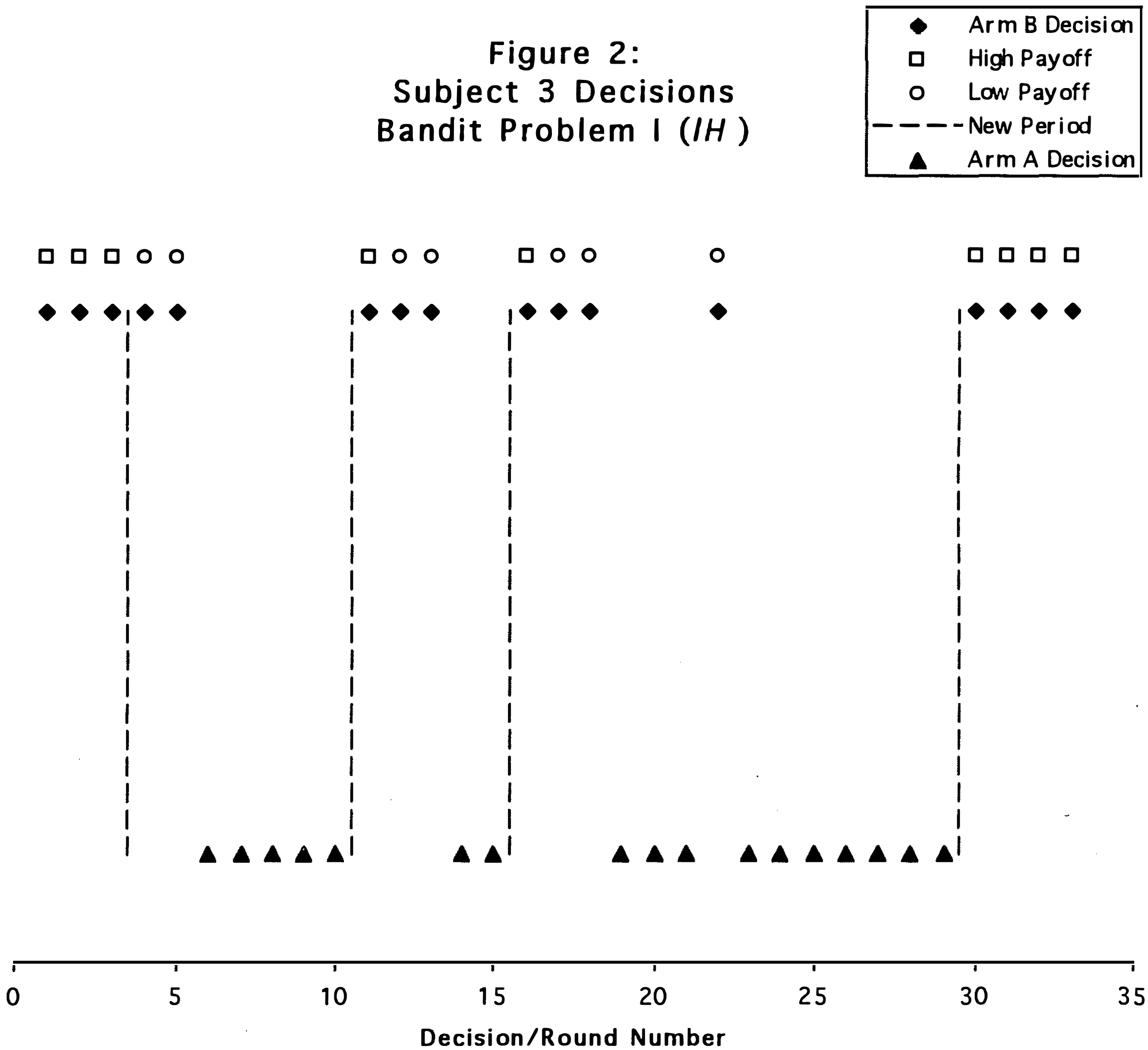


Figure 3: Distribution of Cutpoints-- Bandit Problem I

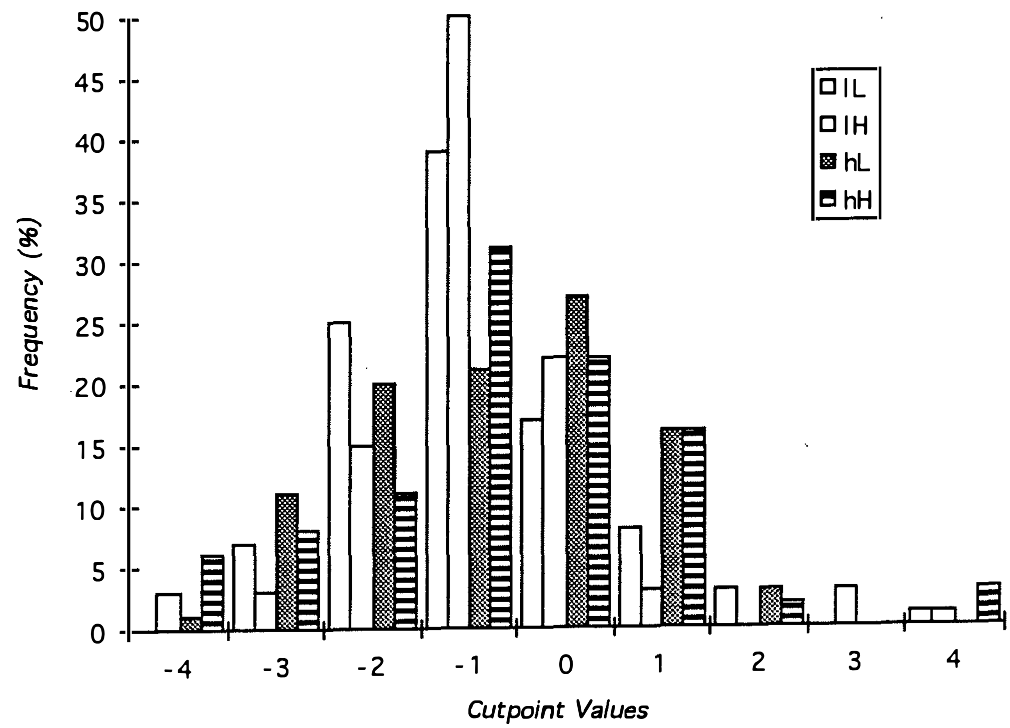
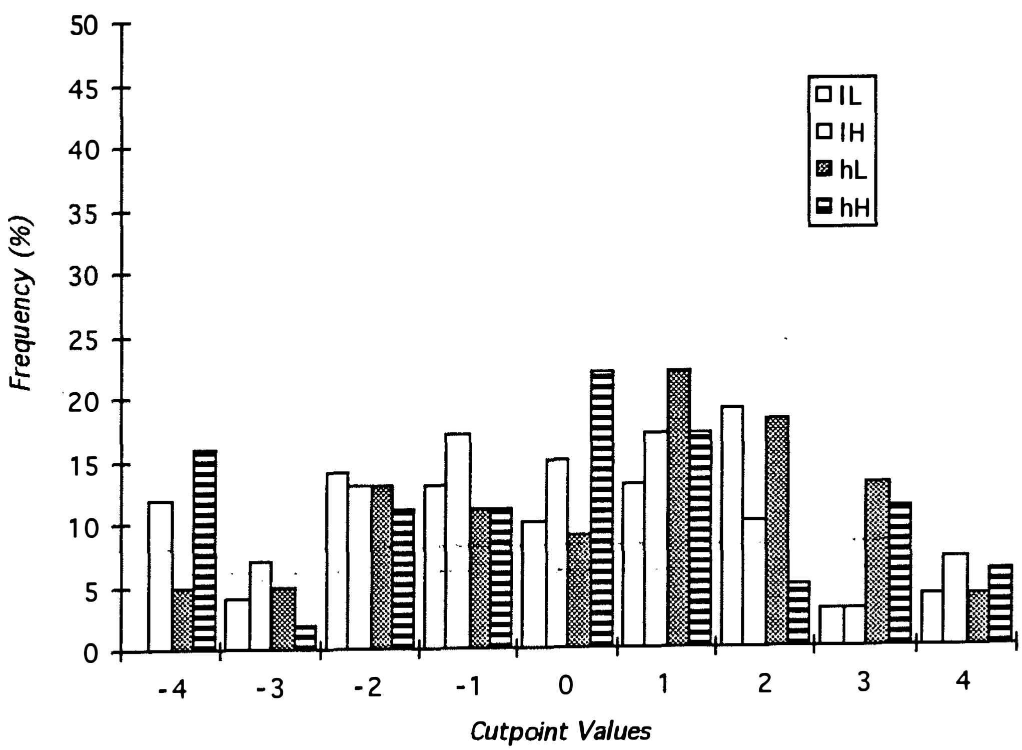


Figure 4: Distribution of Cutpoints-- Bandit Problem II



Bibliography

Banks, Jeffrey S. and Rangarajan K. Sundaram. 1992a. "A class of bandit problems yielding myopic optimal strategies," *Journal of Applied Probability*, 29:625-632.

Banks, Jeffrey S. and Rangarajan K. Sundaram. 1992b. "Denumerable-armed bandits," *Econometrica*, 60: 1071-1096.

Becker, G.M., M.H. DeGroot, and J. Marschak. 1964. "Measuring utility by a single-response sequential method," *Behavioral Science*, 9:226-232.

Berg, Joyce E., Lane A. Daley, John W. Dickhaut, and John R. O'Brian. 1986. "Controlling preferences for lotteries on units of experimental exchange," *Quarterly Journal of Economics*, 101:281-306.

Camerer, C. E. and K. Weigelt. 1993. "A test of a probabilistic mechanism for inducing stochastic horizons in experiments," in *Research in Experimental Economics*, V 6, R. M. Isaac (ed.).

Cooper, Russell W., Douglas V. DeJong, Robert Forsythe, and Thomas W. Ross. 1989. "Communication in Battle-of-the-Sexes games: some experimental results," *Rand Journal of Economics*, 20:568-587.

Cox, J. and R. Oaxaca. 1989. "Laboratory experiments with a finite horizon job-search model," *Journal of Risk and Uncertainty*, 2:301-329.

Cox, J. and R. Oaxaca. 1991. "Finite horizon search behavior with and without recall." Presented at the Economic Science Meetings, Tucson, AZ.

Cox, James C. and Ronald L. Oaxaca. 1993. "Introducing risk neutral preferences: further analysis of the data," Presented and the Economic Science October Meetings.

El-Gamal, Mahmoud and David M. Grether 1993. "Uncovering behavioral strategies: likelihood-based experimental data mining," SSWP 850, Caltech, Pasadena CA.

Gittins, J. and D. Jones 1974. "A dynamic allocation index for the sequential allocation of experiments," in *Progress in Statistics* (J. Gani et al., Eds). Amsterdam: North Holland, 241-266.

Grether, David M. 1992. "Testing Bayes rule and the representativeness heuristic: Some experimental evidence". *Journal of Economic Behavior and Organization*, 17:31-57.

Harrison, G.W. and P.B. Morgan. 1990. "Search intensity in experiments," *Economic Journal*, 100:478-486.

Ishikida, Takashi 1992. "Informational aspects of decentralized resource allocation," Technical Report, UC Berkeley, Interdisciplinary Group on Coordination Theory.

Jovanovic, B. 1979. "Job-search and the theory of turnover," *Journal of Political Economy*, 87:972-990.

Mortensen, D. 1985. "Job-search and labor market analysis," in *Handbook of Labor Economics*, Vol. II (O. Ashenfelter and R. Layard Eds.). Amsterdam, North Holland, 849-919.

Rietz, Thomas A. 1993. "Implementing and testing risk-preference-induction mechanisms in experimental sealed-bid auctions", *Journal of Risk and Uncertainty* , 7:199-213.

Roth, Alvin E., and Michael W. K. Malouf 1979. "Game-theoretic models and the role of information in bargaining," *Psychological Review*, 86:574-594.

Rothschild, M. 1974. "A two-armed bandit theory of market pricing," *Journal of Economic Theory*, 9:185-202.

Viscusi, W. 1979. "Job-hazards and worker quit rates: an analysis of adaptive worker behavior," *International Economic Review*, 20:29-58.

Walker, James M., Vernon L. Smith, and James C. Cox. 1990. "Inducing risk neutral preferences: an examination in a controlled market environment," *Journal of Risk and Uncertainty*, 3:5-24.

Wilde, L. 1979. "An information-theoretic approach to job quits," in *Studies in the Economics of Search* (S. Lippman and J. McCall, Eds.) New York: North Holland, 35-52.