- PRELIMINARY AND INCOMPLETE -
PLEASE DO NOT QUOTE WITHOUT PERMISSION

BEHAVIOR IN A DYNAMIC DECISION PROBLEM: EVIDENCE FROM THE LABORATORY*

BY

DANIEL HOUSER**
DEPARTMENT OF ECONOMICS AND ECONOMIC SCIENCE LABORATORY
UNIVERSITY OF ARIZONA

MICHAEL KEANE
DEPARTMENT OF ECONOMICS
YALE UNIVERSITY
NEW YORK UNIVERSITY

KEVIN MCCABE
INTERDISCIPLINARY CENTER FOR ECONOMIC SCIENCE
GEORGE MASON UNIVERSITY

December 5, 2001

* We gratefully acknowledge that this research was supported by grants from the Russell Sage Foundation, the CV-STARR center, and the Office of the Vice President for Research at the University of Arizona. The first author thanks the International Foundation for Research in Experimental Economics for a fellowship that provided additional support for this research.
** Correspondent: dhouser@econlab.arizona.edu.

I. Introduction

The question "how do people behave when placed in environments where optimal decision making would require solution of complex optimization problems?" is of fundamental importance to economic analysis. Predictions of individual and market behavior can differ dramatically between models where: (1) people are able to solve complex optimization problems, (2) people are "boundedly rational" and adopt rule of thumb behaviors, or (3) people are "confused" or "irrational" and adopt blatantly sub-optimal decision rules. Recognizing the importance of this issue, there have been a large number of experimental studies that analyze the behavior of people confronted with complex decision problems in laboratory settings. Most of this literature adopts an "either/or" approach, asking whether subjects make optimal decisions or not, or asking what fraction of subjects behave optimally.

More recently, a new literature has emerged in which investigators adopt an exploratory approach, asking what types of decision rules people actually use, rather than simply asking whether or not people adopt the optimal rule. It has been long recognized that different people may use different strategies, or decision rules, when playing games or dealing with other complex decision problems. Nevertheless, statistical procedures aimed at rigorously characterizing the nature and number of such strategies have only recently started to emerge, with El-Gamal and Grether (1995) representing perhaps the earliest paper in this literature.

Our work represents both a substantive and a methodological contribution to this emerging literature. Methodologically, we provide a new Bayesian procedure for drawing inferences about the nature and number of decision rules that are present in a population of agents, where each agent is confronted with a dynamic decision problem. Our approach takes the exploratory nature of the analysis a step further than do earlier approaches to decision rule classification. Specifically, the El-Gamal and Grether (1995) approach requires the investigator to specify a priori both the number and form of all decision rules used in the population. In contrast, our approach allows the data to determine both.[1] Like El-Gamal and Grether, our approach generates the posterior

[1] We do not wish to claim that the added flexibility our procedure is always an advantage. In fact, in many contexts it may be particularly efficacious to use our procedure and the El-Gamal and Grether procedure in

probability that each agent is following each decision rule, conditional on his/her observed behavior. Thus, one can classify individual subjects into types a posteriori.

Substantively, we apply our new classification procedure to analyze the actual behavior of a people who are confronted with a particular dynamic optimization problem. Four characteristics of the problem are important to understand. First, it is a game against nature, where a computer generates stochastic payoffs according to rules that are explained to the subjects prior to the experiment. Second, it is a sequential decision problem, in which agents chose between two discrete alternatives ("1" and "2") in each of 15 time periods, and where the optimal choice between the two options changes over time as new information is revealed. Third, the problem is inherently dynamic, because current choices affect the distributions of future payoffs. Fourth, the problem is very difficult to solve optimally. Analytic solution is not possible, and we must solve for the optimal decision rule numerically. In our experimental design, participants are allowed to practice playing the game several times before they play for money.

Our prior was that few if any participants could determine the optimal decision rule in this problem, even given time for practice. Our interest was in characterizing the rules of thumb that subjects would adopt when confronted with a problem that was too difficult to solve optimally. We expected to see a great deal of heterogeneity in the rules of thumb that people adopted, and were concerned that the classification procedure might perform poorly when confronted with a large number of potential types.

Currently we have run the experiment on 139 subjects, and we find the results to be surprising. Our classification procedure produces a very clear assignment of the population into only <u>three</u> distinct types. Statistical tests overwhelmingly reject the hypothesis that there are more, or that there are fewer. Visual inspection of the choice sequences of the subjects, as well as examination of exit interviews in which subjects were asked about their strategies, provide external validation that there clearly are three distinct types. Posterior probabilities of type assignments are almost always close to one.

---

concert. For instance, in some contexts, a preliminary application of our approach could determine the number and form of the candidate rules, and subsequent application of the E-Gamal and Grether approach might then provide a more efficient assignment among these candidate rules. Also, merging of the two approaches may be useful in contexts where particular candidate rules are of particular interest (e.g., the optimal rule) and the investigator wants to test these against alternatives that are not specified a priori.

That is, conditional on his/her behavior, there is rarely much ambiguity about which of the three types a subject is.

Roughly one third of the subjects adopted a decision rule that is nearly identical to the optimal rule. Payoff loses for subjects following this rule, relative to what could be earned by following the exactly optimal rule, average only about 2 percent. Slightly more than one third of the subjects adopted a rule of thumb that is fairly simple to characterize, and that performs rather well in this problem. Payoff loses from following this rule rather than the optimal rule are about 5% on average. Finally, about a third of subjects follow a more clearly sub-optimal decision rule, indicating they have a more severe misunderstanding of the nature of the game. Payoff loses from following this rule average about 16%. To give a baseline, completely myopic play, in which a subject ignores the dynamic aspect of the game and simply chooses the options with the highest current payoff in each period, would lead to payoff loses of 25% on average.

In summary, we find that our type classification procedure performed very well at identifying types in the population and grouping people into types with a high degree of confidence. We were surprised to find that, in a very difficult dynamic decision problem, nearly a third of people were able to closely approximate the optimal decision rule, while more than another third were able to uncover a very good rule of thumb.

The success of most participants in our experiment was particularly surprising given that the monetary rewards for optimal performance were small. A person who used the optimal decision rule could only expect to earn a few dollars more than one who behaved completely myopically.[2] In future work, we plan to examine how various experimental design features, such as (1) size of payoffs, (2) complexity of the decision problem, (3) amount of time for practice, and (3) amount of information given to the participants, effect the types of decision rules people use. Our (very) long term goal is to be able to provide some characterization of the types of situations in which people do and do not behave close to optimally, and to characterize the types of decision rules people use in the latter cases.

[2] Experimental work that finds rejections of rational or optimal behavior is often criticized on the grounds that subjects had little incentive to behave optimally. We don't find this a compelling criticism of our findings here, since the interesting outcome was that so many people indeed behaved close to optimally.

The outline of this paper is as follows: In section 2 we give some background on the literature on experimental analysis of behavior in complex decision problems. In section 3 we present our experimental design. In section 4 we describe the Bayesian algorithm for classifying decision rules. Section 5 presents results and section 6 concludes.

II. Background and Literature Review

Behavioral heterogeneity plays an important role in both theoretical and applied economic research. For example, there is a long theory literature that examines the effects of various "heuristic" decision rules on equilibrium outcomes (see, e.g., Cyert and Degroot (1974), Radner (1975), Akerlof and Yellen (1985), Haltiwanger and Waldman (1985), Ellison and Fudenberg (1993), Krusell and Smith (1995) and Lettau and Uhlig (1999)).

There is a great deal of work in experimental economics that investigates whether or not people play a particular equilibrium or follow a particular decision rule when confronted with particular games or decision problems. For example, stopping behavior in dynamic environments has been analyzed extensively within the context of the wage search model (see Mortenson (1970) for the canonical model, and for experimental tests see, e.g., Braunstein and Schotter (1982), Cox and Oaxaca (1989, 1992), Harrison and Morgan (1990) and Hey (1987)).[3] Although the design specifics differ, this literature's idea is to compare formally observed with optimal search durations in order to determine whether subjects make "optimal" stopping decisions.[4] A typical finding is that only a proper subset of subjects behaves optimally, and little effort is made to describe the rules followed by subjects who play otherwise.

More recently, investigators have become interested in classifying subjects into types based on the strategies or decision rules they use. El-Gamal and Grether (1995) suggest an approach to type classification that requires the investigator to pre-specify the superset of candidate decision rules (or strategies). They present a statistical procedure that chooses a "best" subset of rules from the superset and assigns each subject to one of

[3] Subject typing has also been attempted in many other experimental environments. This includes the public goods game (see, e.g., Gunthorsdottir, et. al., (2000)) and the extensive form "trust and reciprocity" games (see, e.g., McCabe et. al., (2001)).

those rules. An advantage of their procedure is that it is quite simple to implement. However, unless it is feasible to include all of the rules that subjects can possibly use within the prespecified superset (see, e.g., Schachat and Walker (2000) for this case), a potential drawback of this method is that the "right" rules might not be included in the candidate set (see, e.g., Houser and Winter (2001) for this case). This sort of misspecification could lead to misclassification of subjects, and mask underlying commonalities in subjects' play.

The approach we pursue in this paper makes endogenous the number of decision rules used in the population, the nature of each decision rule, and the assignment of subjects to decision rules. We accomplish this by performing Bayesian inference for a model of decision rules that generalizes the mixture of normals probit model considered by Geweke and Keane, (1997, 2001). In our generalization, each of the elements of the mixture represents a unique decision rule. Each decision rule is modeled flexibly, following the suggestion of Geweke and Keane (1999) (see also Geweke, Houser and Keane (2001)), with its form inferred from the experimental data. The number of decision rules in the population is decided through standard Bayesian model choice procedures, which require calculation of the marginal likelihood. These procedures are discussed in Gelfand and Day (1994), Geweke (1997), Geweke and Keane (2001) and Rossi (2001).

III. Experimental Design

We wished to design a dynamic decision problem with the following features:

1) We wanted a problem that was very difficult to solve. The entire point of our experiment is to examine how people behave when confronted with a decision problem that is too difficult for anyone to solve exactly.[5]

2) We nevertheless needed a problem whose structure was easy to explain to participants in the experiment.

[4] In addition to observing search durations, Cox and Oaxaca (1992) also elicit the reservation wages of their subjects directly.

[5] We are particularly interested in examining heterogeneity in decision making behavior, and individual differences in behavior can be masked when the optimal solution is transparent.

3) We felt it was desirable to have a problem whose structure was in fundamental ways similar to dynamic decision problems that people actually confront in real world situations, particularly situations that economists are actually interested in.
4) We wanted to design a problem where there was some noticeable advantage to playing near optimally, so that sub-optimal behavior would be easy to detect statistically.
5) We needed a game that could be played quickly, so that subjects would be willing to participate, so that we could collect a reasonably large amount of data, and so that subjects would be able to practice. This also prevents subjects from getting bored.

We settled on a stochastic sequential discrete choice problem, that has features similar to a human capital investment or occupational choice problem. Each subject in our experiment makes 15 decisions sequentially. The problem has a nontrivial intertemporal component in the sense that early decisions influence the distributions of payoffs for later decisions. In addition, there are stochastic components to payoffs that make solving the problem optimally very difficult (because the optimal decision rule is a function of the realization of the stochastic variables). We set up the problem so that alternative "1" can be thought of similar to "school" or "white collar" work, in that this option tends to have low initial payoffs that increase later if the subject builds up sufficient experience in "1." Option "2" has a higher mean payoff, but does not have any such "investment" component.

A more precise description of the game is as follows: At the start of each of the 15 periods, the subjects receive a draw for the payoffs in alternatives "1" or "2." The stochastic payoff to "2" is 4000 points plus the realization of a random variable that is uniformly distributed over the interval [-5000, 5000], subject to the restriction that the payoff be non-negative. For example, if the realization of the random variable was 5000, then the immediate payoff to "2" would be 9000 points, while if the realization was negative 4000 (or less) then the payoff would be zero.[6] Note that the realizations of the random variables in period t occurs before the decision at t is made, but the realizations of period t+1's random variables occurs after the decision at t.

The payoff to alternative "1" was 3000 points plus the realization of an independently distributed random variable distributed uniformly over the interval [-5000,

[6] Thus, there is a .10 probability of a zero payoff in option 2.

5000], plus a "bonus" and a "cost" whose values depended on the history of the subject's choices. The bonus was 7500 points, and was available whenever alternative "1" had been chosen six, seven, eight or nine previous times (not necessarily in succession). A "switching" cost of 5000 points was incurred whenever the subject chose "1" after choosing "2" in the previous period. The subject's total payoff for the decision problem was the sum of the rewards they earned over the 15 periods.

There are several reasons that we chose this design to study dynamic decision-making behavior. First, it is a sophisticated dynamic investment problem that is nevertheless straightforward to describe. Because subject confusion seems ubiquitous in experimental environments (see, e.g., Andreoni (1995) and Houser and Kurzban (2001)), it is important that the nature of the decision problem be clear and unambiguous. However, while out problem is easy to describe, it turns out that it is quite hard to determine the expected wealth maximizing strategy. In fact, the dynamic optimization problem cannot be solved analytically – it must be solved numerically on a computer. Finally, our design also addresses the so-called "flat-maximum" problem. This arises when the optimal strategy has payoffs that are not very different than those associated with seemingly naïve strategy. Under the parameterization we chose, the optimal solution earns about 27% more, on average, than the myopic strategy that simply chooses the highest payoff each period. Finally, we found that if the game lasted only 15 periods then subjects could play the game quite quickly.[7]

This decision problem was coded in Visual Basic and subjects made decisions independently at a private computer terminal.[8] Figure A1 provides a screen shot of the

[7] We originally attempted to pattern the game quite closely on a white collar vs. blue collar occupational choice problem, using estimates of returns to experience and transition costs calibrated from Keane and Wolpin (1997). This meant that returns to experience were about 5% per period higher in Alternative "1" than in alternative "2." We found, however, that percentage payoff loses from following sub-optimal decision rules (like myopia) were quite small in this context. This led us to the more "dramatic" form of experience return that we actually specified for alternative "1," in which the mean payoff jumps drastically after 6 periods of experience are accumulated. We speculate that with only 15 periods, there is not enough time for experience to accumulate, so the loss from myopic behavior is not large unless the return to experience is very large. Note that in Keane and Wolpin (1997) there were 40 periods. This issue illustrates the tradeoff among goals 3, 4 and 5 for design of the game: If the game had been longer, we could have made it more "realistic," but then subjects would probably have gotten bored. Observation of the subjects suggested that the large majority found the game rather fun to play. Further evidence of this is that most spent a lot of time practicing, even though the monetary gain from optimal play was just a few dollars.

[8] The visual basic software used to implement this experiment is available from the authors on request.

interface seen by the subjects as they made decisions. The screen provides information on the current payoffs to both alternatives, the current round, the history of the subjects choices and payoffs, their current aggregate earnings and a summary of the decision problem's payoff structure. Moreover, subjects were provided with written instructions that provided additional information on the payoff structure and the rules of the experiment. An exact transcript of their instructions is provided in the appendix.

We report here data from the first 139 subjects who completed our experiment. We are still collecting additional data at this time. All experiments were conducted at the Economic Science Laboratory at the University of Arizona. Subjects were recruited from the general student population using the ESL's standard recruiting procedures. In an effort to ensure subjects were familiar with the task when they played it for money, subjects were recruited for two laboratory sessions. On arrival for the first session they were provided the written instructions and seated privately at a computer terminal. They were allowed to practice the decision problem as long as they liked (up to about 90 minutes) but did not play the decision problem for money. A typical subject practiced for about 30 minutes and played the decision problem about 45 times.

The second laboratory session was held two days after the first (for example, if the first was on Wednesday the second would be on Friday.) Upon arrival, subjects were again provided with the written instructions and told that they could practice for as long as they liked (up to about 90 minutes.) Practice was generally shorter during the second session, with many subjects practicing only a few times. When they were ready, subjects solved the decision problem one time for money. Their total earnings were the sum of their earnings in the decision problem and two show-up fees. Subjects earned about $9 on average during the decision problem, earned two show-up fees, and spent about 75 minutes on average in the laboratory.

IV. The Bayesian Classification Procedure

Our Bayesian approach to analyzing decision rule heterogeneity enables us to draw inferences about the number and nature of the population's decision rules, as well as the probability with which each subject uses each rule. The approach generalizes methods for Bayesian inference in mixture of normals probit models discussed in Geweke and Keane (1997, 2001)). In this section we first describe the way we model the decision rules used by the subjects in our experiment, and then derive the posterior distribution on which our inferences are based.

IV.A. A flexible model of decision rules

It is useful to start by considering the optimal decision rule in a dynamic stochastic discrete choice problem. Apply Bellman's (1957) principle, the value to subject n of choosing alternative j at time t is:

(1) $V_{njt}(I_{nt}) = w_{njt} + EV(I_{n,t+1} \mid I_{nt}, j)$ $\quad t=1,..,T$

where:

$I_{n,t+1} = H(I_{nt}, j)$,

Here w_{njt} is the current period payoff, meaning the monetary reward won by the subject in round t of the game. I_{nt} is the state of the subject at time t (i.e., the subject's information set). This might include, for example, their choice and payoff history. $EV(I_{n,t+1} \mid I_{nt}, j)$ is the "future component" of the value function which captures the expected value of the subject's state next period given his/her current state and choice,[9] and $H(\cdot)$ is the (possibly stochastic) law of motion for the state variables.

If subjects form expectations rationally, and know $H(\cdot)$, then E in equation (1) is the mathematical expectation operator, and the $EV(I_{n,t+1} \mid I_{nt}, j)$ at all state points can be solved for (possibly numerically) via dynamic programming. If, as in our experiment, there are two alternatives then the optimal decision rule is:

Chose alternative 1 iff $V_{n1t}(I_{nt}) > V_{n2t}(I_{nt})$.

We wish to generalize this framework by allowing for the possibility that subjects do not use the optimal decision rule,[10] and for the possibility that there is heterogeneity in the decision rules that exist in a population of subjects.

[9] If payoffs are received over time, the future component would be pre-multiplied by a discount factor, but in the game we consider here all payoffs are received at the same time (at the end of the game), so there is no discounting.

[10] This could occur for a number of reasons. Among these are: 1) subjects do not have rational expectations, so that E is not the mathematical expectation operator, 2) subjects are unable to do the

Rather than assume E is the mathematical expectation operator, we model the future component of each alternative's value as a flexible parametric function (i.e., polynomial) in the elements of the subject's information set I_{nt} and choice j. We also allow for the parameters of this function to differ across subjects of different type, denoted k. And finally, we allow for the possibility of optimization errors. Then, we can write the future component for type *k* as:

$$F(I_{nt}, j, \pi_k) + \varsigma_{njt}$$

Here $F(\cdot)$ represents the future component polynomial. It is characterized by a finite vector of parameters π_k. Note that the type k specific future component parameters π_k capture differences in the way subjects assign values to alternatives. The random variable ς_{njt} that accounts for idiosyncratic errors made when attempting to implement decision rule *k*. We allow the distribution of the idiosyncratic errors to vary with the decision rule, so that stochastic optimization error may be more important for some types than others. Let σ_k denote the standard deviation of the optimization error for type k.

The value that subject *n*, who is type *k*, assigns to alternative $j \in \{1,2\}$ in period t, can then be written

$$(2) \qquad V_{njt}(I_{nt} \mid k) = w_{njt} + F(I_{nt}, j, \pi_k) + \varsigma_{njt}$$

This specification can approximate well the "correct" value function given by (1) if the function F is sufficiently flexible, the π_k are chosen appropriately, and the variance of the optimization error ς_{njt} is set close to zero. In practice, many authors have found that value functions in problems of interest to economists can be approximated extremely accurately using low order polynomials (see, e.g., Krusell and Smith (1995), Geweke and Keane (1999), Keane and Wolpin (1994)). Equation (2) nests myopic behavior if the π_k are set equal to zero and the variance of the optimization error is set equal to zero.

calculations necessary to solve the DP problem and compute the optimal decision rule, 3) subjects do not understand the structure of $H(\cdot)$, 4) subjects solve a simpler problem to save on thinking costs.

Equation (2) generates random behavior in which each option is chosen with equal probability if the π_k are set equal to zero and σ_k is sent to infinity.

It is important to note what aspects of this specification are and are not restrictive. Our specification does not require that subjects understand what variables are and are not relevant state variables in the problem. We have not made this explicit to conserve on notation, but we could also allow for the possibility that the subjects consider superfluous information, and denote the expanded information set by I^{+}_{nt}. The specification may appear to impose additive separability, but in fact it does not, since current and past payoffs, and interactions between them, may be included as elements of I_{nt}. Observe that non-stationarity arises if time is an element of I_{nt}.

The three important restrictive aspects of the specification are that we will have to make a distributional assumption on the optimization errors (we assume they are normal), we will have to specify the elements of I_{nt} that enter into the F polynomial, and we will have to specify the order of the polynomial. It is important to note however, that the same Bayesian decision theoretic procedures that we use to determine number of types may also be used to choose among these aspects of alternative specifications.

We denote the choice in period *t* of subject *n* following decision rule *k* with information I_{nt} by:

$$d_k(I_{nt}) = \begin{cases} \text{"1" if } Z_{nt}(I_{nt} \mid k) > 0 \\ \text{"2" otherwise} \end{cases} \qquad \forall k \in K$$

where $Z_{nt}(I_{nt} \mid k) = V_{nAt}(I_{nt} \mid k) - V_{nBt}(I_{nt} \mid k)$.

Note that choices only depend on the value function differences Z_{nt}. The data observed by the investigator will be a sequence of current payoff realizations w_{njt} for j=1,2, t=1,T and choices d_{nt} for t=1,T for each subject n. But the latent indices Z_{nt} that determine choices are unobserved. Thus, the model we are considering is formally a mixture of probit models, with the added information that an additive part of the latent index, $w_{1nt} - w_{2nt}$ is observed. In contrast to the probit, this sets the scale for the other parameters, so further normalization is not needed (i.e., the scale of the π_k and σ_k are both identified).

Given the observed data on current payoffs and choices, our goal is to draw inferences about the parameters π_k, σ_k for each type k=1,K, as well as about the number of types K, the population proportions of each type θ_k. Furthermore, we also want to construct the posterior probability $p_n(k)$ that each subject used each decision rule, conditional on his/her observed decisions. We do this using the Bayesian inference procedure described below.

IV.B. Functional forms

At this point we describe the specific application of our procedure to the experiment described in section III. Here, w_{njt} is the known immediate reward for alternative *j*. For instance, w_{nBt} is the (possibly censored) sum of 3000 and period *t*'s realization of the uniformly distributed random variable.

In our experiment, the relevant state variables for forecasting values of future states are the number of times a person has chosen each alternative, which we denote by X_{n1t} and X_{n2t} for alternatives one and two, respectively, the time remaining until the last period (since it is a finite horizon problem), and an indicator for whether the current choice is "1" or "2." The current choice matters for future payoffs because of the cost of a transition from "2" to "1." The prior "experience" in "1" and "2," as well as the time left in the game, matter because of that fact that the mean payoff in "1" jumps substantially when one reaches 6 periods of experience in "1." This "bonus phase" only lasts until the person chooses "1" four additional times, and rational subjects should try to get through this bonus phase before the end of the game.

Note that current and lagged payoffs, while they are elements of the information set I_{nt}, are not useful for forecasting future payoffs, because payoff draws are iid over time in our experiment. Thus, the "rational expectations" (RE) future component $EV(I_{n,t+1} \mid I_{nt}, j)$ does not vary over state points (I_{nt}, j) that only differ in terms of the realizations of the w_{njt}. Therefore, a polynomial approximation to RE future component would not depend on the current and/or lagged w_{njt}. So far, we have only estimated specifications in which the future component is as a polynomial in the relevant state variables for forecasting future payoffs. In future work we plan to allow the future component to include superfluous state variables (like the payoffs) to investigate whether subjects use these erroneously.

In our work to date, we assume that the future component F for subjects of type k is a third order polynomial that takes the following form (where we have suppressed the type subscript *k*):

$$
\begin{aligned}
F(H(X_{n1t}, X_{n2t}, j) \mid \pi) = {} & \pi_0 + \pi_1 (X_{n1t} + 1(j=1)) + \pi_2 (X_{n2t} + 1(j=2)) \\
& \pi_3 (X_{n1t} + 1(j=1))^2 + \pi_4 (X_{n2t} + 1(j=2))^2 + \pi_5 (X_{n1t} + 1(j=1))(X_{n2t} + 1(j=2)) + \\
& \pi_6 (X_{n1t} + 1(j=1))^3 + \pi_7 (X_{n2t} + 1(j=2))^3 + \pi_8 (X_{n1t} + 1(j=1))^2 X_{n2t} + \pi_9 (X_{n2t} + 1(j=2))^2 X_{n1t} + \\
& \pi_{10} 1(j=1) + \pi_{11} 1(j=1)(X_{n1t} + 1(j=1)) + \pi_{12} 1(j=1)(X_{n2t} + 1(j=2)) + \\
& \pi_{13} 1(j=1)(X_{n1t} + 1(j=1))^2 + \pi_{14} 1(j=1)(X_{n2t} + 1(j=2))^2 + \\
& \pi_{15} 1(j=1)(X_{n1t} + 1(j=1))(X_{n2t} + 1(j=2)) + \\
& \pi_{16} 1(j=2) + \pi_{17} 1(j=2)(X_{n1t} + 1(j=1)) + \pi_{18} 1(j=2)(X_{n2t} + 1(j=2)) + \\
& \pi_{19} 1(j=2)(X_{n1t} + 1(j=1))^2 + \pi_{20} 1(j=2)(X_{n2t} + 1(j=2))^2 + \\
& \pi_{21} 1(j=2)(X_{n1t} + 1(j=1))(X_{n2t} + 1(j=2))..
\end{aligned}
$$

Note that time remaining in the game is linearly dependent on the number of choices in the two alternatives, so we omit it from the polynomial. Note that we have implicitly defined the law of motion as

$$H(X_1, X_2, j) = (X_1 + 1(j=1), X_2 + 1(j=2)).$$

Since choices depend only on relative alternative valuations it is clear that the future component is not identified in levels. We achieve identification by differencing the future component as follows. Let

$$
\begin{aligned}
f(X_{n1t}, X_{n2t} \mid \pi) = {} & F(H(X_{n1t}, X_{n2t}, 1)) - F(H(X_{n1t}, X_{n2t}, 2)) \\
& = \pi_1 - \pi_2 + \pi_3(2X_{n1t} + 1) - \pi_4(2X_{n2t} + 1) \\
& + \pi_5(X_{n2t} - X_{n1t}) + \pi_6(3X_{n1t}^2 + 3X_{n1t} + 1) \\
& - \pi_7(3X_{n2t}^2 + 3X_{n2t} + 1) + \pi_8(-X_{n1t}^2 + 2X_{n1t}X_{n2t} + X_{n2t}) \\
& - \pi_9(X_{n2t}^2 - 2X_{n1t}X_{n2t} - X_{n1t}) \\
& + \pi_{10} + \pi_{11}(X_{n1t} + 1) + \pi_{12}X_{n2t} \\
& + \pi_{13}(X_{n1t} + 1)^2 + \pi_{14}X_{n2t}^2 \\
& + \pi_{15}(X_{n1t} + 1)X_{n2t} \\
& - \pi_{16} - \pi_{17}(X_{n2t} + 1) - \pi_{18}X_{n1t} \\
& - \pi_{19}(X_{n2t} + 1)^2 - \pi_{20}X_{n1t}^2 \\
& - \pi_{21}(X_{n2t} + 1)X_{n1t}.
\end{aligned}
$$

Note that not all of the parameters that enter the future component are identified. For example, π_1 and π_2 are not separately identified, π_5 is not separately identified from π_3 and π_4, and π_9 is not separately identified from $\pi_3, \pi_4, \pi_6, \pi_7$, and π_8. Our analysis is based on the following identified relative future component.

$$
\begin{aligned}
f(X_{n1t}, X_{n2t} \mid \pi^*) = {} & F(H(X_{n1t}, X_{n2t}, 1)) - F(H(X_{n1t}, X_{n2t}, 2)) \\
& = \pi_1^* + \pi_2^*(2X_{n1t} + 1) + \pi_3^*(-2X_{n2t} - 1) \\
& + \pi_4^*(3X_{n1t}^2 + 3X_{n1t} + 1) + \pi_5^*(-3X_{n2t}^2 - 3X_{n2t} - 1) \\
& + \pi_6^*(-X_{n1t}^2 + 2X_{n1t}X_{n2t} + X_{n2t}).
\end{aligned}
\tag{3}
$$

where the π_i^* are derived from the π_i in the obvious way.

Note that the decision rule for subject n of type k at time t is:

Choose “1” iff $Z_{ntk} = w_{n1t} - w_{n2t} + f(X_{n1t}, X_{n2t} \mid \pi_k) + \eta_{nt} > 0$.

It is therefore intuitive to think of $f(X_{n1t}, X_{n2t} \mid \pi_k)$ as a reservation payoff differential that the subject requires in order to chose “2” over “1,” (subject to the added noise induced by the mean zero optimization error η_{nt}). In our game, the optimal value of this reservation differential varies in a complex way with the state variables X_{n1t} and X_{n2t}. This is what makes it very difficult to play the game optimally. Our algorithm will allow us to infer the actual reservation payoff differential function $f(X_{n1t}, X_{n2t} \mid \pi_k)$ used by each type k. We can then compare these to the optimal *f* in order to characterize the manner in which play of each type of subject deviates from optimality.

IV.C. Likelihood function and joint posterior distribution.

The probability that subject *n* using decision rule k chooses alternative one at period *t* is as follows.

$$P(d_k(I_{nt})=1)=P(V_{n1t}(I_{nt})>V_{n2t}(I_{nt}))=P(w_{n1t}-w_{n2t}+f(X_{n1t},X_{n2t}\,|\,\pi_k)+\eta_{nt}>0)$$

where $\eta_{nt}=\varsigma_{n1t}-\varsigma_{n2t}$. We assume $\eta_{nt}\sim IIDN(0,\sigma_\lambda^2)$.

If all subjects' types are known, then the likelihood function can be written down easily as follows.

$$L[(\pi_k,\sigma_k^{-2})_{k\in K}\,|\,\{I_{nt},d_k(I_{nt})\}_{n,t}]=\prod_n\prod_k\{\theta_k\prod_t P(d_k(I_{nt}))\}^{1(n\text{ uses rule }k)}.$$

Here, θ_k is the probability that a person chosen at random from the population follows decision rule k. However, since we do not actually know subjects' types, we use a Bayesian Markov Chain – Monte Carlo algorithm to generate inferences about the subjects' types along with π_k and σ_k.

The model is closed by specification of prior distributions for the model's parameters. We assume proper priors for all of the model's parameters. Our priors are centered on myopia, in the sense that under prior means the future component is zero, and have standard conjugate specifications as follows.

$$\begin{aligned}&\pi_k\sim N(0,\underline{\Sigma}),\text{ where }\underline{\Sigma}\text{ is a 6 x 6 diagonal matrix with entries }\underline{\Sigma}(1,1)=20{,}000^2,\\&\underline{\Sigma}(2,2)=\underline{\Sigma}(3,3)=1{,}000^2,\underline{\Sigma}(4,4)=\underline{\Sigma}(5,5)=100^2\text{ and }\underline{\Sigma}(6,6)=100{,}000.\\&\sigma_k^{-2}\sim\chi^2(1).\\&\{\theta_k\}_{k\in K}\sim Di(\{2.0\}_{\|K\|}),\theta_1>\theta_2>\cdots\theta_K.\end{aligned}\tag{4}$$

where $\|K\|$ denotes the number of mixtures in the population, and *Di* is the Dirichlet distribution, or multivariate Beta. Note that the restriction on the population probabilities is an identifying restriction that prevents interchanging the components of the mixture. There are several such restrictions that can work for this purpose, see Geweke and Keane (2001).

At this point it is straightforward to adopt a Gibbs sampling algorithm to approximate the marginal posteriors of the model's parameters as well as to draw inferences about the probability with which each subject follows each decision rule. We present the details on how to implement this algorithm in Appendix C.

V. Empirical Results

In the experiment, we recorded both the choices made by subjects and the idiosyncratic random variable realizations that they faced. Because we also know the payoff structure, this enabled us to construct the choices that would have been made by each subject had they been playing the optimal decision rule - which we call the "rational expectations" (RE) decision rule.[11] The fraction of actual and RE subjects who choose alternative "1" each period is provided in Figure 1. The choice paths are quite similar over the first six rounds. From round seven to nine, actual subjects choose alternative one more frequently than their RE counterparts, and after round ten they choose "1" much less frequently. On average, RE subjects choose "1" 10.7 times over the course of the decision problem, while our actual subjects chose that alternative 10.0 times. Actual subjects earned on average 11.1% less than the RE subjects.

However, it would be highly misleading to only look at "average" play, because there is substantial variation in play around these averages. Of our 139 subjects, 11 earned exactly the same as their RE counterpart (because they made perfectly identical decisions), while 16 earned more. Also, while on average our subjects chose "1" less often, 82 of them chose "1" at least as often as their RE counterpart, and of these 38 chose "1" more often. The fact that there are between subject differences in play relative to the RE baseline is consistent with subjects using different decision rules.[12]

[11] To be precise, we used standard numerical procedures to solve the appropriate dynamic programming problem and construct the decision rule that would be used by rational, expected wealth maximizing agents.

[12] The possibility that multiple decision rules are at use in our subject population is supported by inspection of individual decisions. While we noted above that 11 subjects played as though they were expected wealth maximizers, many subjects did not seem to use such easily described rules. For example, one subject chose the higher payoff in each of the first 13 rounds, and then chose the lower payoff in each of the last two periods. Decision rule differences are also suggested by answers to a questionnaire completed by some subjects after they finished the experiment. One subject described his/her strategy as, "choose the greater one," while another's response, "The first 10 times just choose 1, and then for the last 5 times just choose the biggest number each time," suggests an awareness of alternative one's investment value.

We use the Bayesian type classification procedure described above to organize this heterogeneity. We wrote Gibbs sampling software to implement the procedure, and to draw inferences about the marginal posterior distributions of the model's parameters.[13] We ran the algorithm for 12,000 cycles and discarded the first 2,000. Inspection of the draw sequence convinced us that the chain had converged by the 2,000th cycle. The discussion below is based on a model that allows for three decision rules. The marginal likelihoods for models including one, two, three and four mixtures are, respectively, –1950.07, -1607.15, -1410.11 and –1613.42. Since the marginal likelihood for the model with three types is over 200 points greater than that for the models with either two or four types, the likelihood of the data is vastly greater under the three type specification. If we assign equal prior probabilities to models with each number of types, then the posterior odds ratios overwhelmingly point to the three mixture model as the most appropriate specification.[14]

Table 1 reports the posterior means and standard deviations for the parameters of each decision rule, as well as the number of subjects that are assigned to each rule. Casual inspection of the parameters suggests that the rules are different, and we show in simulations below that they do have very different implications for behavior. Note that the posterior standard deviations of the model parameters are in all instances quite small relative to the prior standard deviations, indicating that the data is very informative about the parameters.

In order to characterize the behavior of each type, and to further assess the fit of the three type model, we assigned each person to a type based on his/her posterior type probabilities. Each subject was assigned to the rule that he/she followed with greatest probability. It turned out the vast majority of subjects can be assigned to one type very clearly, because the posterior type probabilities were almost always close to one. This means that a subject's choice behavior is usually highly informative about which type that subject is.

First, we assess the fit of the estimated decision rules to the actual play of the subjects who we classified as following that rule. To do this, we simulated the

[13] Our FORTRAN 77 code, which makes extensive use of IMSL subroutines, is available on request.
[14] Comparison of models that use polynomials of different orders will appear in future drafts.

hypothetical decisions that each subject would have made under his/her assigned decision rule, given the realizations of the random variables that he/she actually experienced. We compared these hypothetical choices to the choices the subjects actually made. Figure 2 describes the fit of the model averaged across all subjects. The fit seems reasonably close in all rounds, with the broad features of actual decisions, such as the peak that occurs in the ninth round, well matched by decisions under the estimated rules.[15] Figure 3 describes the fit of each decision rule to the play of subjects of that type. Again, the main features of each type's play seem to be reasonably well matched by the simulated choices.

V.A. Characterization of the Decision Rules

Next, we attempt to characterize the nature of decision rule used by each of the three types. Table 2 compares the play of subjects within each type with that of hypothetical RE counterparts. In each case, we simulated the optimal choices that hypothetical RE players would have made if confronted with the same random draws as the actual subjects. Type 2 subjects perform least well, earning about 19% less on average than their hypothetical RE counterparts. Type 3 subjects do best, earning on average only about 2.4% less. The type 1 subjects earn about 11% less than their RE counterparts.[16] At the outset, it appears that type 3 subjects are using something close to the optimal decision rule, while type 2 subjects are behaving clearly sub-optimally, and type 1 subjects are somewhere in between.

Next, we simulated the play of hypothetical subjects who use each of the three rules. In the first simulation all subjects use the first decision rule, in the second simulation all subjects use the second decision rule, and so on. In each simulation we construct 139 hypothetical choice histories, setting the realizations of the random variables to the values that the subjects in the experiment actually experienced. In this way, each decision rule is confronted with a common set of draws, so differences in choice behavior are due only to the differences in the rules. For comparison purposes, we also conducted a fourth simulation in which the hypothetical subjects use the RE rule,

[15] Average earnings under the estimated decision rules is 85,443 experimental points, which is slightly higher than the 82,292 experimental points actually earned on average by the experimental subjects. Each 10,000 experimental points was converted to one U.S. dollar.

and a fifth simulation in which the hypothetical subjects are myopic (they simply chose the alternative with the highest payoff in each period).

Figure 4 describes the results of our simulation exercise. Rules 1 and 3 track the RE rule closely through round six. This is not the case for Rule 2, which by the sixth round finds only half of the subjects in alternative one, while the other rules lead about 75% of the subjects to make this choice. Rule 2 continues to imply behavior that diverges substantially from the other rules until the last two rounds, when it and Rule 1 begin to converge. Under Rule 1, play begins to diverge from Rule 3 and RE at about round 11, at which point there is a sharp and continuing drop in the number of subjects that choose alternative one. One feature that both the Rule 3 and RE profile share is a peak in round 11, and a monotonic decline thereafter. The Rule 1 profile peaks at round 9, and Rule 2 peaks in round 1. Average experimental points earned under each decision rule is (in thousands) 86.2, 76.3, 89.9 and 91.4, for Rule 1, Rule 2, Rule 3 and RE, respectively. Note that myopic play differs substantially from the other rules. Average points under this rule are about 68.6 K. The percentage loses relative to the RE rule are 1.6 percent for rule 3, 5.7 percent for rule 1, 16.5 percent for rule 2, and 24.9 percent for myopia.

In addition to studying the behavioral implications of the different decision rules, one can compare the estimated future components directly. Figure 5 compares the estimated differenced future components, for each decision rule type, with the RE differenced future component at the initial round and at round five's possible state variables. In each case the vertical axis denotes the value of the differenced future component, so that larger values indicate that choosing alternative one is expected to generated greater future benefits in relation to alternative two. The horizontal axis indicates the number of times alternative one was previously chosen. Figure 6 provides analogous information for rounds nine and 13 for each behavioral type.

Consider first Type 1 subjects. Figure 5 indicates that these subjects value alternative one in about the same way as RE subjects in the first round of the game. This is consistent with Figure 4, which shows that the two types choose alternative one in the

---

[16] Moreover, while only two of 34 Type 2 subjects earn more than their RE counterpart, that number is 10 of 49 for Type 3. All of the 11 subjects who earned exactly the RE amount are classified as Type 3.

first round equally often. However, by round five it is apparent that their differenced future component differs from the RE values in both level and shape. While the RE future component tends to assign less value to experience in alternative one as experience increases, subject 2's seem to do the opposite. Since most Type 1 subjects have either three or four units of experience in alternative one by the fifth period, most overvalue alternative one, relative to the RE values, at this point in the decision problem. However, Figure 5B shows that by the 13th round Type 1s are generally placing less value on alternative on than their RE counterparts, and at some points in the state-space sharply less. This is consistent with the behavior exhibited in Figure 4, which shows that RE subjects choose alternative one more often in the latter rounds of the decision problem.

Figure 5 indicates that Type 2 and RE subjects have about the same future component value in the first round. The values assigned at round five are on average lower than the RE values, although both sets vary with experience in about the same way. Still, this is consistent with the relatively few type 2 subjects that choose alternative 1 even early in the decision problem. Figure 5B shows that the differences are quite apparent by the ninth round, and dramatic by the 13th. Again, it seems clear that type 2 agents are, at most points in the state space, substantially undervaluing alternative one in relation to two. This is implied by their choice paths as described in Figure 4, and is consistent with the fact that, on average, they choose alternative one about 3.8 fewer times than their RE counterparts.

Finally, Figures 5 and 6 provide evidence that Type 3 subjects value alternatives in a way that is similar to the RE values. This is true in both the earlier and later rounds of the decision problem, and this is consistent with the behavioral patterns described in Figure 4.

Yet another way to characterize differences between decision rules that our subjects used is to compare the answers provided to a questionnaire distributed to some subjects after the experiment. Responses were voluntary, and we obtained responses from only about a third of our subjects. Subjects were asked whether they followed a particular strategy and, if so, to describe it. Casual inspection of the survey results does not suggest a relationship between whether subjects claimed to be following a certain strategy and their type classification: there are Type 3 subjects who claimed not to be following any specific strategy, and Type 2 subjects who claimed they were. However, conditional on reporting a strategy, the classification often seems reasonable. Most subjects who reported a strategy and are classified as Type 3 indicated some awareness of the investment value of choosing alternative one early in the decision problem. For example, one Type 3 subject responded, "Choose 1 unless 2 is very high." Another Type 3 described his/her strategy as, "I tried to choose 1 as much as I could, because as I did so, I could have a possibility to make more money at the next round."

Some subjects who were classified as Type 2 seemed not to understand the decision problem. One Type 2 subject explained that he/she attempted "to keep choosing 2, or amount of 5-7 dollars in average. Keep choosing 1 for more profit, but maybe less profit too." Another Type 2 indicated that he/she made an effort to "choose option 2 [the] first couple of times and then 1 the rest, looking forward to the bonus of 7500." Yet another said that his goal was to, "choose 1 at least 9 times in a row." Casual inspection indicates that this sort of confusion is much more apparent in Type 2 subjects than either Type 3 or Type 1.

Subjects who were classified as Type 1 often indicated an awareness of the investment value of alternative one, but also indicated that they behaved "myopically" once the bonus had been obtained. One Type 1 subject indicated that his/her strategy was to "choose 1 at least 10 times. Choose higher of a or b the others." Another Type 1 said, "I chose alternative 1 nearly every time, because alternative 2 would not pay out a high yield in the long run. Near the last few periods, I chose alt. 2. I knew I was not going to choose alt 1 again (i.e.: did not want to incur the costs.) If the payoff was greater for alternative 2 near the end, I took that amount." Yet another wrote that he/she wanted to "just make sure that I had enough 1's to get the bonus."

Taken together, these findings suggest that our subjects' solve our dynamic decision problem in one of three ways. Type 3 subjects look very much like RE subjects, both in average play and in the way they assign values to alternatives. Hence, we might label these subjects as using a "sophisticated" decision rule. This rule performed best among the three rules in our population. Subjects following this rule earned only about 2 percent less, on average, than they would have earned by playing the exactly optimal decision rule. We were surprised that about one third of our subjects were able to "learn" to play nearly optimally in a very difficult dynamic problem after about a half hour of practice (one average).

Type 1 subjects followed a decision rule that leads to payoffs about 6% less, on average, than one would earn by playing the exactly optimal decision rule. Based on analysis of Figure 5 and visual inspection of the choice sequences for these subjects, their behavior seems to be well described by the following: These subjects understand that there is an investment value to choosing option "1." And they get the reservation payoff differential just about right in period 1. But type 1's don't understand how the reservation payoff differential should vary with experience in alternative X_1. In fact (see Fig. 5A), they think that choosing "1" becomes more valuable if they have chosen it a lot in the past, when in reality it is less valuable.[17] Thus, type 1's could be described as "pessimists." They start the game thinking that getting to the bonus phase is unlikely. But if they get "lucky" and get close to the 6 periods of experience necessary to attain the bonus, they decide that maybe they really can get the bonus. So they start to chose "1" with great urgency. We will call the type 1's the "rule-of-thumb" type.

Type 2 subjects performed relatively poorly. They follow a decision rule that leads to payoffs about 16% less, on average, than one would earn by playing the exactly optimal decision rule. It seems that these subjects behaved and assigned values similarly to the RE rule in the first couple of rounds, but then began to assign too little value to alternative "1" in subsequent rounds. In other words, like type 1's they recognize that there is an investment value to alternative "1," but they make even greater errors in figuring out how the reservation payoff differential should vary with experience in alternative X_1. Answers to survey questions suggested that some of the type 2 subjects were confused in the sense that they did not understand the rules or payoff structure of the problem. Thus, it seems reasonable to label this as the "confused" decision rule. Nevertheless, even the type 2's did better than if they had followed a myopic decision rule.

[17] If you choose "1" a few times in the first few rounds, it reduces the urgency of choosing "1." That is, you will probably get to the bonus phase before the end of the game even if you choose "2" the next few rounds. But type 1's don't think this way. They think it becomes more urgent to chose "1" if they have already chosen it a few times in the first few rounds.

VI. Conclusion

Our results suggest that subjects followed three decision rules, in roughly equal proportion, in our experimental decision problem. One group followed a rule very close to the rational expectations (expected wealth maximizing) rule. Another third of our subjects played according to a "rule-of-thumb" that performed about 5 percent less well than the optimal rule, on average. Hence, after practice, about two-thirds of our subjects were able to solve the problem near optimally, as judged by payoff losses. About one-third of our subjects performed substantially less well, following a rule that would earn about 16 percent less than the optimal strategy on average. We label this the "confused" type. Evidence of confusion in other experimental environments is well documented (see, e.g., Andreoni (1995) or Houser and Kurzban (forthcoming)), and our results suggest that it exists in dynamic decision problems as well.

The procedure described in this paper can be applied to field as well as experimental data. In either case, the econometrician learns about the way in which people use their information to make decisions. Because differences in observed decision rules could stem from differences in preferences, one cannot determine from the analysis of field data whether differences in decision rules are due to this or differences in heuristics. Moreover, there might also be differences in the nature of the problem people solve in the field, and these differences might also lead to observed differences in decision rules. In the experimental laboratory the researcher is able to exercise control over both period return functions and the nature of the problem that subjects are supposed to solve.

Because this is a laboratory investigation, our results need not have specific external validity. That is, the particular decision rules that we uncover might be valid only within the narrow context of our laboratory environment. The broader message of this paper is that there seems to be a multiplicity of decision rules at use in the population, and we believe it is reasonable to suspect that this heterogeneity also exists in the naturally occurring world. We believe that developing empirical strategies to investigate and determine the implications of such heterogeneity for policy and welfare analysis provides an important research agenda.[18]

[18] In applied work using field data it has become common to allow for heterogeneity in subject's decision rules. But, if we examine work in structural econometrics, the heterogeneity always enters through agents' preferences or constraints. For instance, Keane and Wolpin(1997) allow for heterogeneity in skill endowments, tastes for leisure and tastes for school attendance. Structural econometricians typically invoke dogmatic behavioral assumptions (e.g., all agents have rational expectations), so the possibility of heterogeneity in how agents solve the decision problems they face is not admitted. One advantage to learning more about how people actually behave in games and decision problems in the laboratory is that this information might eventually be useful as a guide to theoretical and econometric specifications.

References

Akerlof, G. and J. Yellen. 1985: The macroeconomic consequences of near-rational rule-of-thumb behavior. *Quarterly Journal of Economics*.

Andreoni, J. 1995: Cooperation in public goods experiments: kindness or confusion? *American Economic Review.*

Braunstein, Y. and A. Schotter. 1982: Labor market search: an experimental study. *Economic Inquiry*. *20*, 133-44.

Cox, J. and R. Oaxaca. 1989: Laboratory experiments with a finite horizon job search model. *Journal of Risk and Uncertainty*. *2*, 301-29.

Cox, J. and R. Oaxaca. 1992: Direct tests of the reservation wage property. *Economic Journal*. *102*, 1423-32.

Cyert, R. and M. DeGroot. 1974: Rational expectations and Bayesian analysis. *Journal of Political Economy*. 82, 521-36.

El-Gamal, M. A. and D. M. Grether. 1995: Are people Bayesian? Uncovering behavioral strategies. *Journal of the American Statistical Association*, 90, 1137–1145.

Ellison, G. and D. Fudenberg. 1993: Rules of thumb for social learning. *Journal of Political Economy*. 101, 612-43.

Gelfand, A.E. and D.K. Dey. 1994: Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society Series B*, 56, 501-514.

Geweke, J., D. Houser and M. Keane. 2001: Simulation based inference for dynamic multinomial choice models. In B. Baltagi (ed.), *Companion to Theoretical Econometrics*. Blackwell.

Geweke, J. and M. Keane. 1997. Mixture of normals probit models. Federal Reserve Bank of Minneapolis staff report #237.

Geweke, J. and M. Keane. 1999: Bayesian inference for dynamic discrete choice models without the need for dynamic programming. In Mariano, Schuermann and Weeks, (eds), *Simulation Based Inference and Econometrics: Methods and Applications*. Cambridge University Press. (Also available as Federal Reserve Bank of Minneapolis working paper #564, January, 1996.)

Geweke, J. and M. Keane. 2001: Computationally intensive methods for integration in econometrics. In J. Heckman and E. Leamer (eds), *Handbook of Econometrics*, Vol. 5. North Holland.

Gunthorsdottir, A., D. Houser, K. McCabe and H. Ameden. 2000. Excluding free riders improves reciprocity and promotes the private provision of public goods. Manuscript. University of Arizona.
Haltiwanger, J. and M. Waldman. 1985: Rational expectations and the limits of rationality: An analysis of heterogeneity. *American Economic Review*. *75*, 3, 326-340.

Harrison, G. and P. Morgan. 1990: Search intensity in experiments. *Economic Journal. 100*. 478-86.

Hey, J. 1987: Still searching. *Journal of Economic Behavior and Organization. 8*, 137-44.

Houser, D. and R. Kurzban. Forthcoming. Revisiting kindness and confusion in public goods experiments. *American Economic Review*.

Houser, D. and J. Winter. 2001: Time preference and heuristics in a price search experiment. Manuscript, University of Arizona.

Keane, M. and K. Wolpin 1997: The career decisions of young men. *Journal of Political Economy, 105*, 473-522.

Krusell, P. and A. A. Smith. 1995: Rules of thumb in macroeconomic equilibrium: a quantitative analysis. *Journal of Economic Dynamics and Control,* 20, 527-58.

Lettau, M. and H. Uhlig. 1999: Rules of thumb versus dynamic programming. *American Economic Review*. *89*, 148-74.

McCabe, K., D. Houser, L. Ryan, T. Trouard and V. Smith. 2001: A functional imaging study of cooperation in two person reciprocal exchange. Manuscript. University of Arizona.

Mortensen, D. 1970: Job search, the duration of unemployment and the Phillips curve. *American Economic Review*. *60*, 847-62.

Radner, R. 1975: Satisficing. *Journal of Mathematical Economics*. *2*, 256-62.

Schachat, J. and M. Walker. 2000:

Appendix A. An exact transcript of the written instructions provided to subjects.

Instructions

Thank you for coming today. This is a study of individual decision making, for which you will earn cash. The amount of money you earn depends on your decisions, so it is important to read and understand these instructions. All the money that you earn will be awarded to you in cash and paid to you privately at the end of the experiment. The funding for this experiment has come from a private research foundation.

The experiment lasts for 15 periods. Each period you will choose between two alternatives, which will be called '1' and '2'. Each alternative has a payoff which is shown on the left-hand side of the screen. If you choose '1' you earn the payoff associated with '1', and if you choose '2' you earn the payoff associated with '2'. *The payoff for each alternative will be shown to you before you make your choice.* At the end of the experiment, you will be awarded an amount of cash equal to the sum of your 15 chosen payoffs. Your choices are private: do not discuss them with anyone else in the room.

The future payoffs offered for alternative '1' depend on the previous choices that you made. The future payoffs offered for alternative '2' do not depend on any of your previous choices. No payoff will ever be less than zero. The specific structure of payoffs is as follows:

Payoff per period for alternative '1':

Base Pay:		3,000
	0	if you have chosen "1" 0, 1, 2, 3, 4, or 5 previous times
Bonus:	2500	if you have chosen "1" 6, 7, 8 or 9 previous times
	0	if you have chosen "1" 10, 11, 12, 13 or 14 previous times
Costs:		A cost of 5000 will be incurred if you chose '2' the previous period, otherwise none.
Lottery:		Random draw that takes value between -5000 and 5000 with equal chance.
Total payoff:		(Base Pay + Bonus - Costs +/- Lottery), or 0, whichever is bigger.

Payoff per period for alternative '2':

Base Pay:	4,000
Bonus:	None
Costs:	None
Lottery:	Random draw that takes value between -5000 and 5000 with equal chance.
Total payoff:	(Base Pay + Bonus - Costs +/- Lottery), or 0, whichever is bigger.

The payoff structure will be shown to you on the screen for easy reference. Your screen will also include a green window called 'Summary', which will show you the total number of periods in the experiment (15), the current period, your accumulated payoffs, the number of times you have chosen '1', the number of times you have chosen '2', and the choice you made the previous period.

The right hand section of the screen details the history of the payoffs of each alternative, and the choice you made, by period. Finally, you will see in the bottom left hand side of the screen a red window which describes the current period's payoff choices.

You will be paid $5 for attending the first day, another $5 for attending the second day, plus any earnings from the decisions you made on the second day. You will receive all of your payments at the end of the second day.

The first day you can practice as much as you like. The second day, when you are ready, you may play one time for money by pressing the "Play for Money" button in the bottom left hand side of the screen (you will only see this button on the second day). If you have a question raise your hand and an experimenter will come to answer. We cannot tell you which decision is 'best' for you. Your decisions are entirely up to you.

Appendix B. Questions asked in the post-experiment questionnaire. Answers to these questions were voluntary, and subjects were not paid to complete this questionnaire.

1. How old are you?
2. What is your academic classification (i.e., freshman, sophomore, junior, senior, grad student, special student, etc…)
3. Are you male or female?
4. What is your major?
5. What is your grade point average?
6. What is your ethnicity?
7. Approximately how many minutes did you practice last time?
8. Approximately how many minutes did you practice today?
9. Did you use a particular strategy when playing for money?
10. If you had a strategy, briefly describe it.

Appendix C. The Gibbs Sampling Algorithm

Following the notation developed in the body of the paper, suppose there are $n = 1, N$ subjects, $t = 1, T$ periods and $k = 1, K$ decision rules. The Gibbs sampler derives from the product of the priors and the complete data likelihood function. The latter is the likelihood we could write down if all latent utilities and types were observed. Under these conditions, the complete data likelihood function is:

$$f(\{Z_{nt}\}_{n,t}, \{\tau_n\}_{n=1,N} \mid \{\theta_k, \pi_k^*, \sigma_k\}_{k=1,K}) \propto$$
$$\prod_{k=1,K} \prod_{n:\tau_n=k} \left[\theta_k \prod_{t=1,T} \frac{1}{\sigma^k} \exp\left\{ -\frac{(Z_{nt} - (w_{n1t} - w_{n2t} + Y_{nt}^{'}\pi_k^*))^2}{2\sigma_k^2} \right\} I(Z_{nt}, d_{nt}) \right]. \qquad \text{(CDL)}$$

Here, the θ_k, π_k^* and σ_k are as defined for equation (1), and the $Y_{nt}^{'}$ is the vector of state variables conformable with π_k^* also as defined by equation (1). The indicator function $I(Z_{nt}, d_{nt}) = 1$ if $Z_{nt} > 0$ and $d_{nt} = 1$, or if $Z_{nt} < 0$ and $d_{nt} = 2$, but is zero otherwise.

The product of the complete data likelihood and the prior structure defined by equation (2) define the joint posterior used to construct the Gibbs sampler. The sampler includes the following steps.

(1) Draw latent utility values Z_{nt}.
(2) Draw decision rule coefficients π_k^* for all $k = 1, K$.
(3) Draw standard deviation of optimization error $\sigma_k, k = 1, K$.
(4) Draw population type probabilities $\theta_k, k = 1, K$.
(5) Draw individual types $\tau_n, n = 1, N$.

These draws are implemented as follows.

(1) Draw latent utility values Z_{nt}.

Conditional on everything else being known, it is clear from (CDL) that a given Z_{nt} follows a truncated normal distribution, with mean $(w_{n1t} - w_{n2t} + Y_{nt}^{'}\pi_k^*)$ and variance σ_k^2. The truncation is from below at zero if $d_{nt} = 1$, and is from above at zero otherwise. It is trivial to draw from this distribution using standard inverse CDF procedures.

(2) Draw decision rule coefficients π_k^* for all $k = 1, K$.

Conditional on everything else being known, each vector π_k^* can be drawn using simple rejection methods. From (CDL), the source distribution can be chosen to be normal with a mean of $(Y_k^{'}Y_k)^{-1}Y_k^{'}W_k$, and variance $\sigma_k^2(Y_k^{'}Y)^{-1}$. Here, Y_k denotes the stacked array $(Y_{nt})_{n,t}$ for those subjects who are type k. The vector W_k is created by stacking the numbers $Z_{nt} - w_{n1t} + w_{n2t}$ conformably, and again for subjects who are type k. The prior is evaluated at each candidate draw from the source distribution, and that evaluation compared against a random number drawn from a uniform [0,1] distribution. If the uniform random variable is greater than the value of the prior when evaluated at the candidate draw, then the draw is rejected. Otherwise it is accepted.

(3) Draw the standard deviation of the optimization error σ_k.

It is easy to show that, conditional on everything else being known, (CDL) and the prior structure together imply

$$\frac{1+\sum_{\substack{t=1,T\\ n=1,N}} (Z_{nt} - (w_{n1t} - w_{n2t} + Y_{nt}^{'}\pi_k^*))^2 I(\tau_n = k)}{\sigma_k^2} \sim \chi^2(N_k \cdot T),$$

where N_k is the number of people who are type k. It is easy to draw from this distribution using standard software.

(4) Draw the population type probabilities θ_k.

Because we assume $Di(\{2\}_K)$, it is trivial to verify that the posterior conditional on everything else being known is $Di(\{2+N_k\}_{k=1,K})$. We draw from this distribution using the procedure suggested by Anderson (1984, p. 284). In addition, for identification purposes, we impose the order restriction that $\theta_1 > \theta_2 > \cdots > \theta_K$.

(5) Draw individual types $\tau_n \in \{1, 2, ..., K\}$.

Let $f_k(n)$ denote the value of (CDL) when evaluated only for individual n and under the assumption that they use decision rule k and that everything else is known. Then, the conditional posterior distribution of τ_n is

$$\Pr(\tau_n = k') = \frac{f_{k'}(n)}{\sum_{k=1,K} f_k(n)}.$$

It is easy to draw from this distribution using standard software.

	Prior Distribution		Type 1: N=56 "Rule-of-Thumb"		Type 2: N=34 "Confused"		Type 3: N=49 "Rational"	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
π_0^* :Intercept	0.0	$2x10^4$	3565.45	201.94	3641.57	449.01	4167.73	190.68
π_1^* : X1	0.0	10^3	410.02	40.82	-419.48	79.57	-7.65	20.79
π_2^* : X2	0.0	10^3	298.80	51.80	-455.26	120.37	84.49	50.48
π_3^* : X1^2	0.0	10^2	-32.17	2.95	-17.40	5.63	-29.26	1.58
π_4^* : X2^2	0.0	10^2	-4.39	2.56	-2.67	6.99	-73.77	5.23
π_5^* : X1*X2	0.0	$10^{5/2}$	-.27	4.95	-103.47	11.35	-87.49	3.59
σ_η : opt error	Not Def	Not Def	854.23	30.30	2254.99	80.70	218.22	76.50
θ_k :population type probability	0.33	0.18	0.42	0.05	0.31	0.03	0.27	0.03

Table 1. Prior and posterior means and standard deviations of future component parameters. "X1" Denotes experience in alternative "1," and X2 denotes experience in alternative "2."

	Type 1	Type 2	Type 3
Number/Percent of Subjects	56 / 40%	34 / 24%	49 / 35%
Mean Earnings (Points) vs. RE	81589 / 91499	75268 / 93167	87969 / 90108
SE of Mean vs. RE	1765 / 1166	2494 / 1789	1381 / 1508
Number who earn at least as much as RE subjects	4	2	21
Number who earn exactly RE	0	0	11
Number who complete all bonus rounds	40 (71.4%)	12 (35.2%)	46 (93.9%)

Table 2. Mean earnings under the rational expectations decision rule and actual mean earnings for each type.

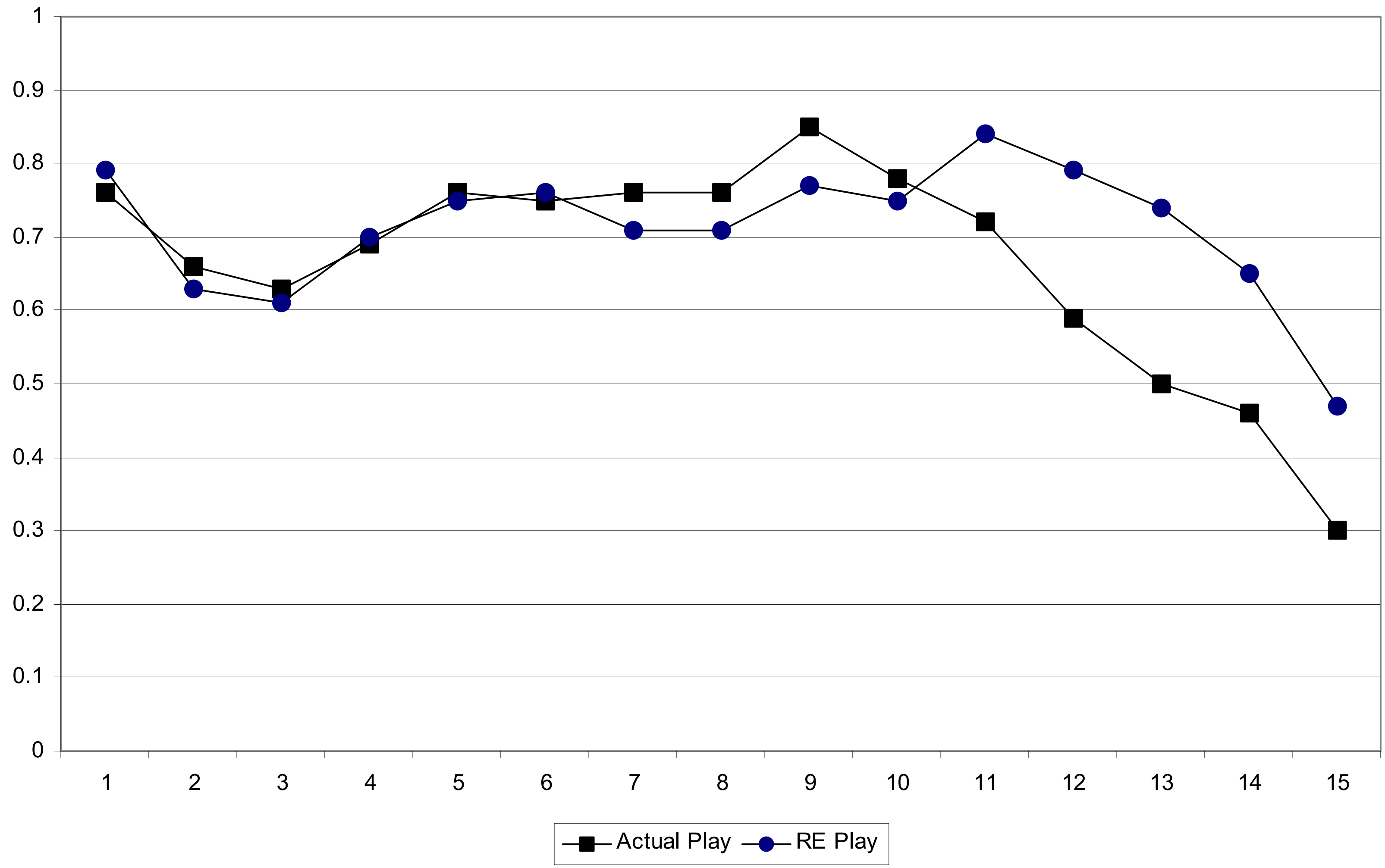

Figure 1. Fraction of actual and rational expectations subjects who choose alternative one by round.

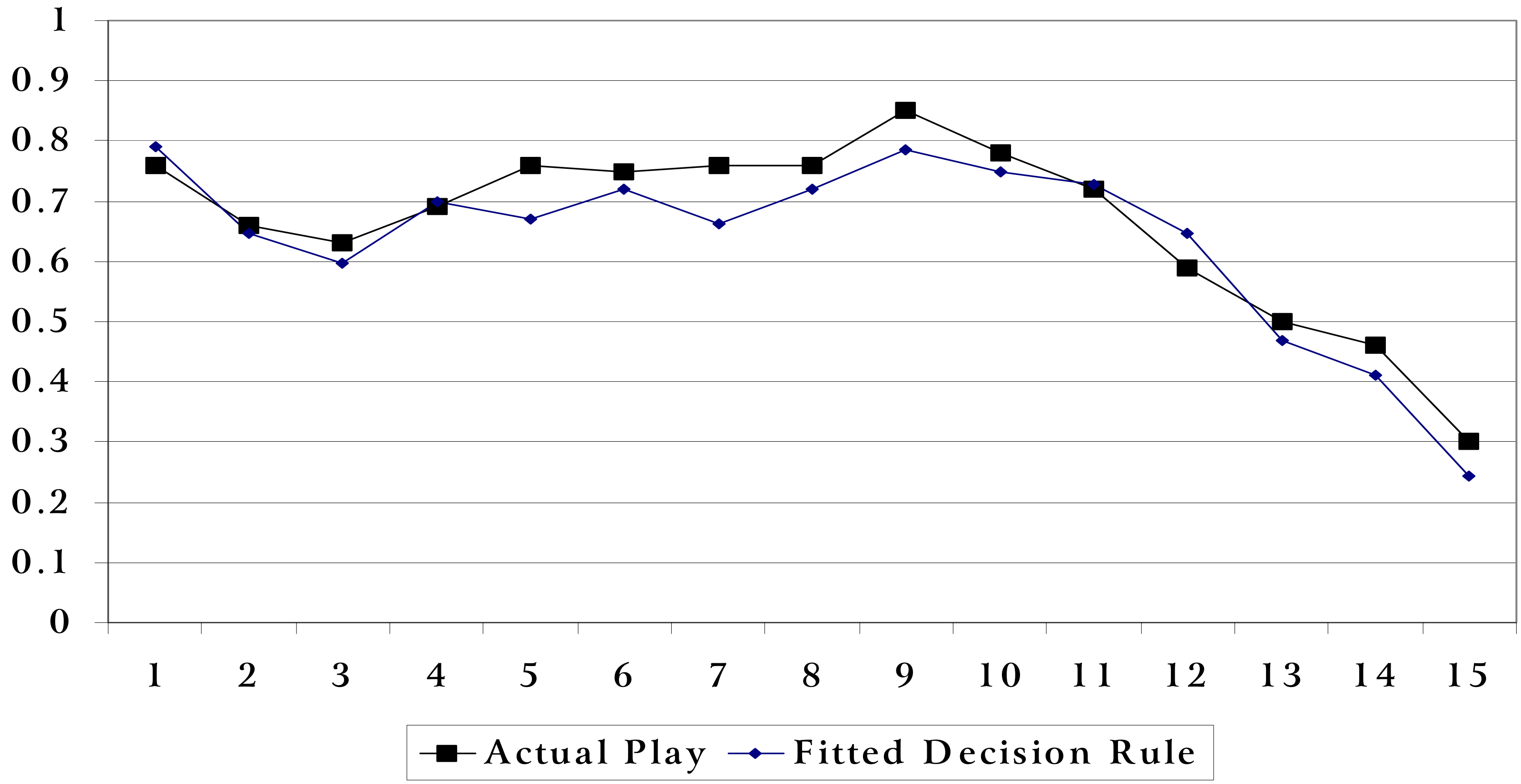

Figure 2. Fraction of people who choose alternative 1 under estimated decision rules and actual play.

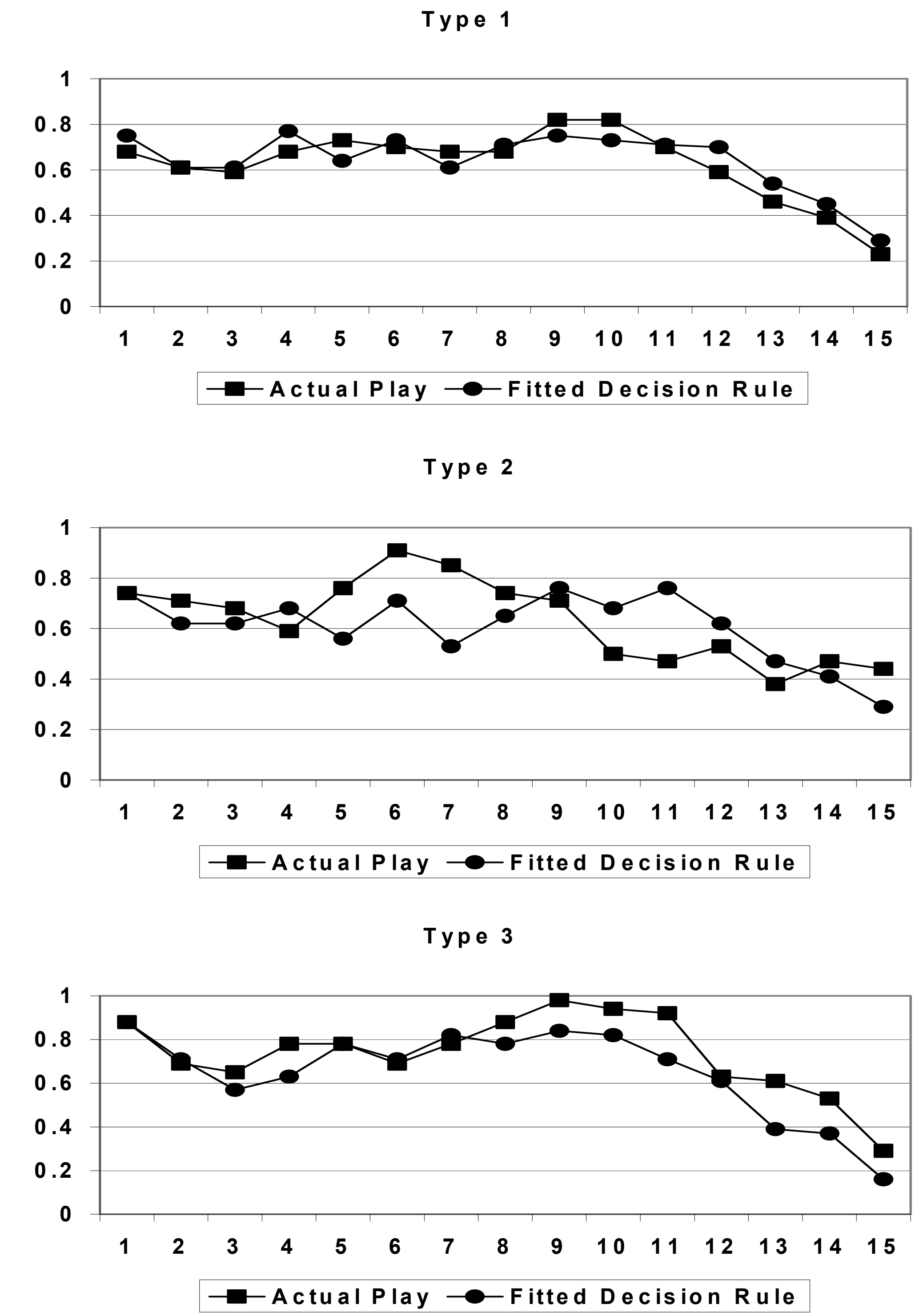

Figure 3. Decisions under fitted decision rule and actual decisions by type.

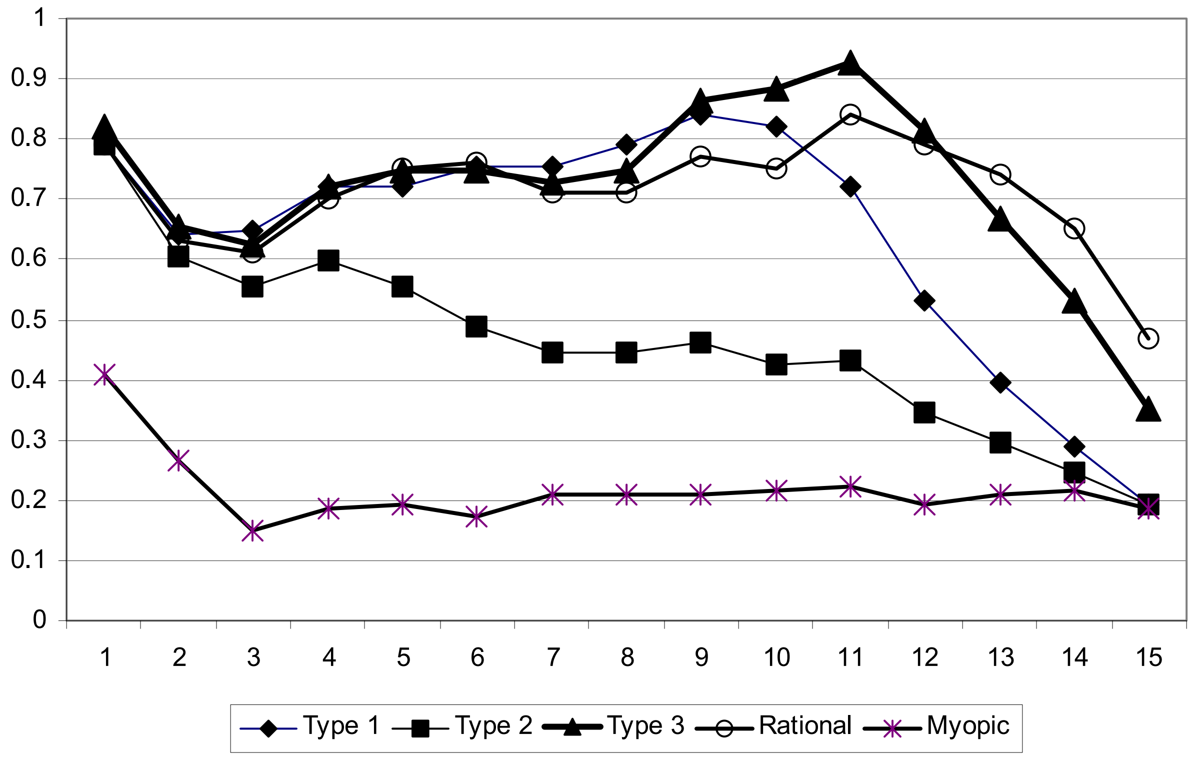

Figure 4. Choice behavior under counterfactual conditions that all subjects follow the Type1, Type2, Type3, RE and myopic decision rules

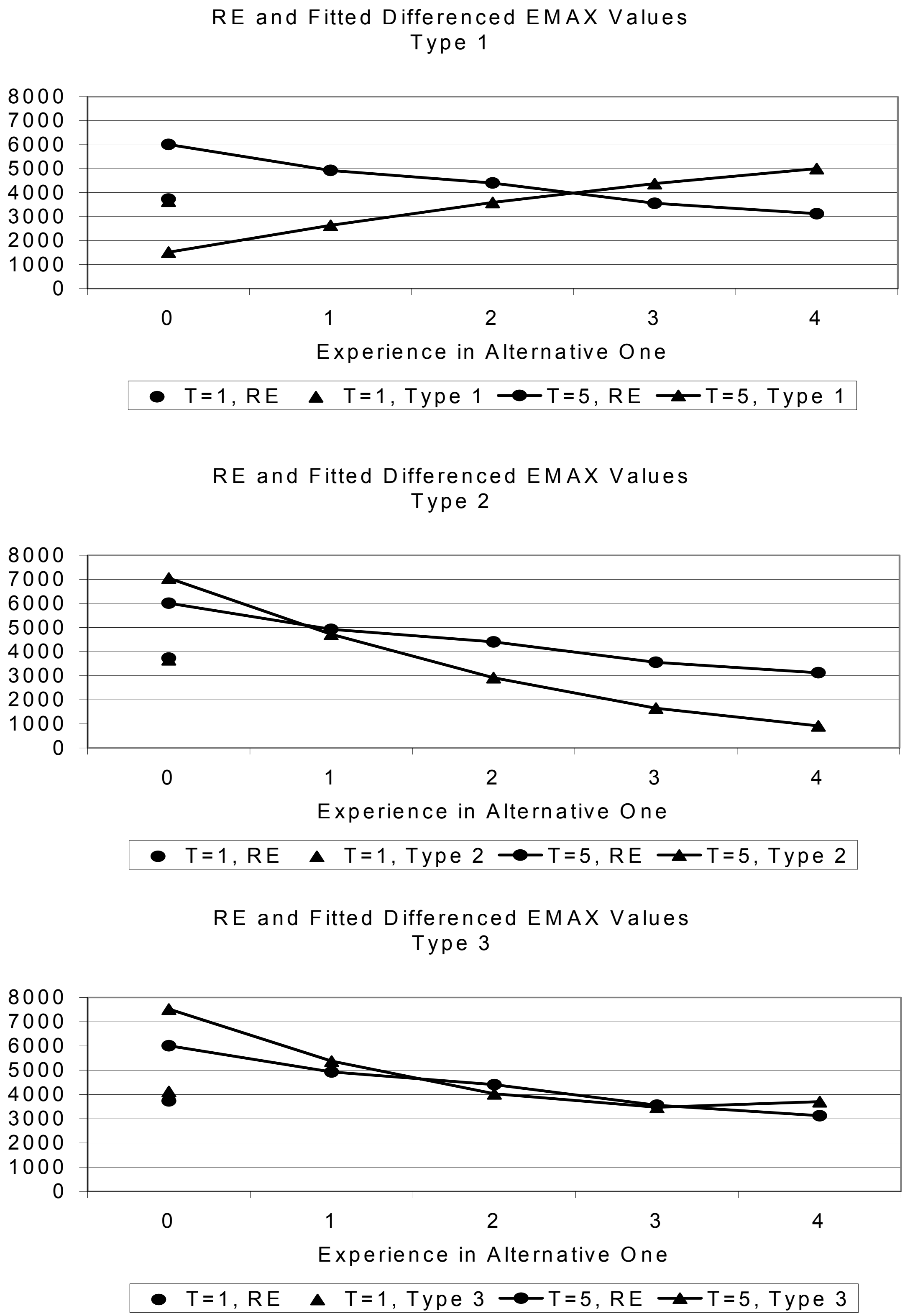

Figure 5A. Fitted and actual differenced EMAX values by type at various state-vectors.

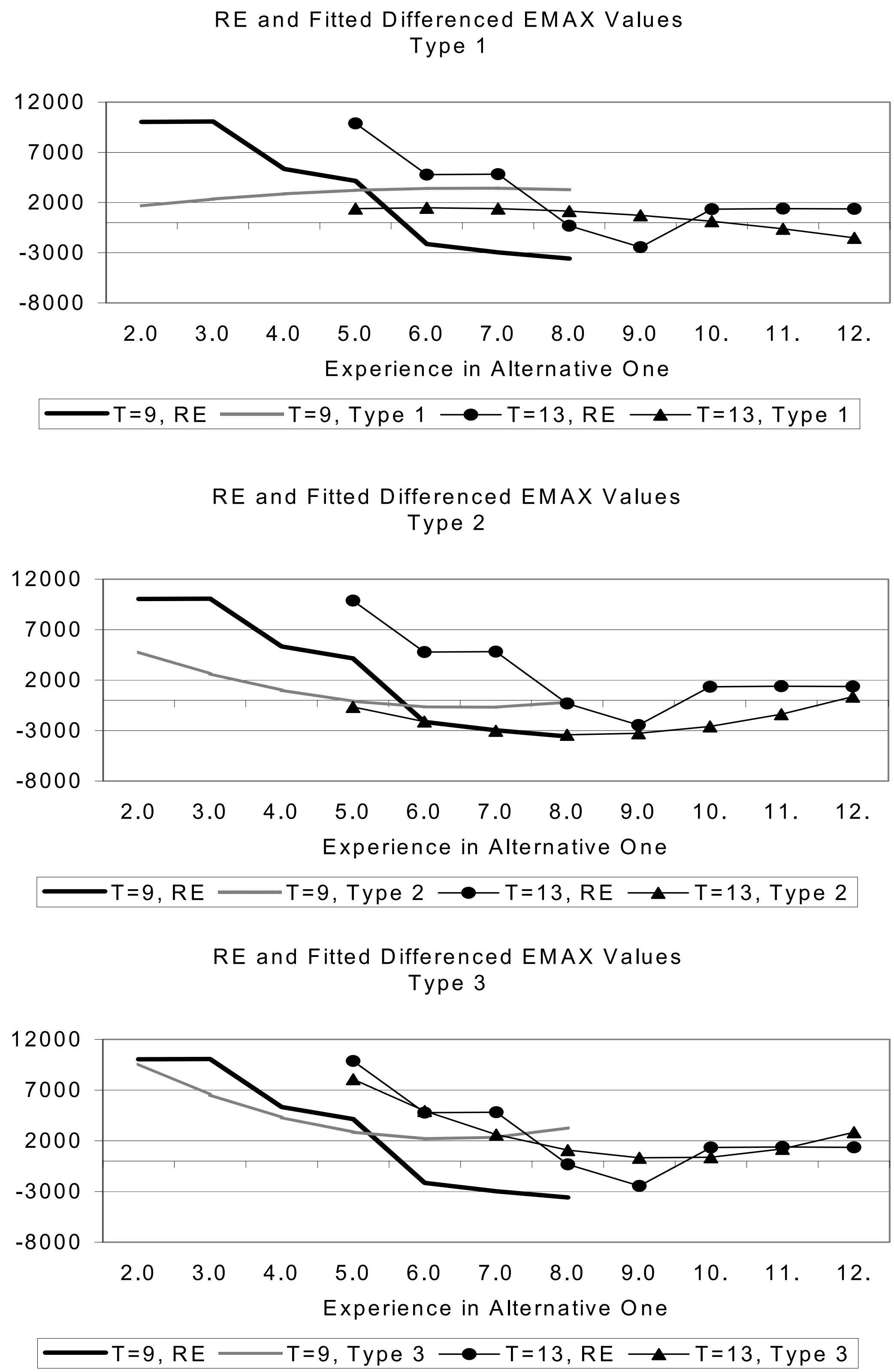

Figure 5B. Fitted and actual differenced EMAX values by type at various state-vectors.

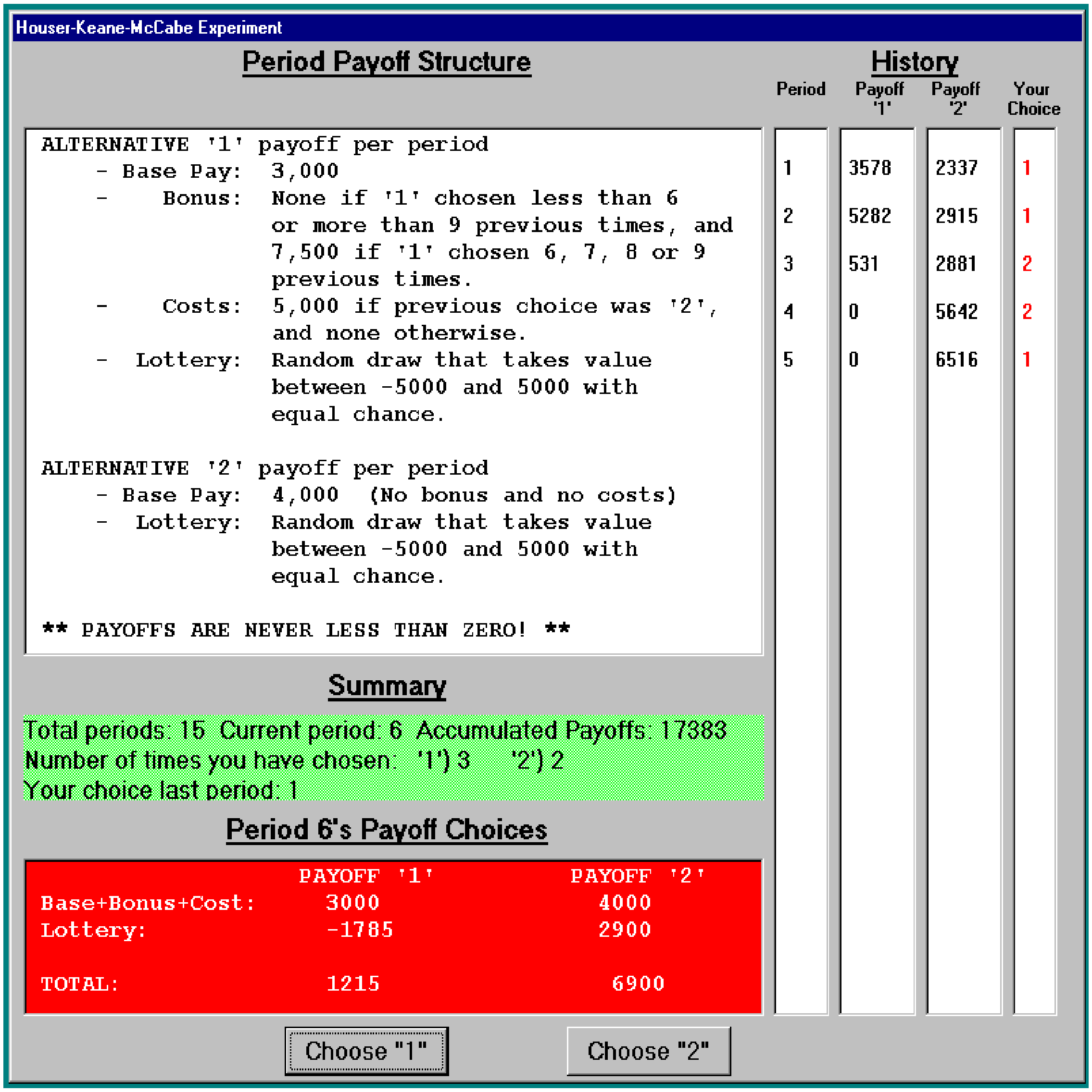

Figure A1. Subjects' screen.