Behavior in a Dynamic Decision Problem: An Analysis of Experimental Evidence Using a Bayesian Type Classification Algorithm

By Daniel Houser, Michael Keane and Kevin McCabe[1]

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. 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. We apply our procedure to analyze the actual behavior of subjects who are confronted with a difficult dynamic stochastic decision problem in a laboratory setting. Subjects were allowed to practice as much as they liked before playing the game for money, and our analysis of decision rules is based on the money round data. Using these data, our procedure classifies the subjects into three types, who we label as "Near Rational," "Fatalist," and "Confused" based on analysis of their decision rules. Lack of effort does not seem to explain the poor performance of the "Confused" subjects, since they devoted the most time to practice before playing the game for money. There is clear evidence of continuity in subject's behavior between the practice and money rounds, as those types who performed best in practice also tended to perform best when playing for money. However, the agreement between practice and money play is far from perfect. The divergences appear to be well explained by a combination of type switching (due to learning and/or increased effort in money play) and errors in our probabilistic type assignments.

KEYWORDS: Dynamic programming, Gibbs sampling, Bayesian decision theory, Experimental economics, Behavioral economics, Heuristics

---

1. Introduction

How do people actually behave in contexts where optimal decision-making requires the solution of complex optimization problems? This question 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.[2] 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. Early work in this literature tended to adopt an "either/or" approach, asking whether subjects make optimal decisions or not, or asking what fraction of subjects behave optimally.[3]

More recently, a new literature has emerged in which investigators adopt a more exploratory approach. The goal is to discover what types of decision rules people actually use, rather than simply asking whether or not they adopt the optimal rule. Thus, behavioral heterogeneity becomes a key issue. The experimental literature has 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 to determine the number and nature of strategies operative in a population have only recently emerged. The important paper by El-Gamal and Grether (1995) was an early contribution to this literature.

Our work represents both a substantive and a methodological contribution to this emerging literature on "typing" experimental subjects. Methodologically, we provide a new Bayesian procedure for drawing inferences about both the nature and number of decision rules that are present in a population of subjects, where each subject is confronted with a dynamic decision problem. Our main identifying assumption is that the analyst must specify *a priori* the set of state variables that may enter agents' decision rules. But, conditional on the set of admissible state variables, our procedure is quite flexible in terms of letting the data determine both the number and form of the rules (i.e., they are modeled as flexible polynomial functions).

---

[2] There is an extensive 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)).

[3] For example, there is an extensive experimental literature on stopping behavior in search models (e.g., Braunstein and Schotter (1982), Cox and Oaxaca (1992), Harrison and Morgan (1990), Hey (1987)). A typical finding is that only a proper subset of subjects behaves optimally, but little effort is made to characterize suboptimal behavior.

In contrast to our procedure, the El-Gamal and Grether (1995) approach requires the investigator to specify *a priori* both the maximal number of decision rules that may be present in the population, and the exact form of each rule (including parameter values). They present a statistical procedure that chooses a "best" subset of rules from the superset of all candidate rules.[4] In the experiment they analyze, the set of plausible rules is fairly obvious. But in many contexts the requirement that the investigator be able to intuit the exact form of all potential rules *a priori* is obviously quite strong. Our assumption that the analyst can pre-specify the set of candidate state variables is somewhat weaker, because the exact form of the rules is left flexible.[5]

A number of other authors have also proposed methods for analysis of behavioral heterogeneity. These include McKelvey and Palfrey (1992), Stahl and Wilson (1995), El-Gamal and Palfrey (1995) and Camerer and Ho (1999). Like El-Gamal and Grether (1995), these approaches have in common that the set of possible subject behaviors must be pre-specified by researcher. Recently, Duffy and Engle-Warnick (2001) and Engle-Warnick (forthcoming), proposed a procedure which models decision rules as sequences of nested if-then conditions. This procedure is, in principle, more flexible. But it suffers from a curse of dimensionality that, as a practical matter, places severe restrictions on the number of decision rules that can be investigated. Engle-Warnick and Ruffle (2002) show that, after suitably constraining the space of possible decision rules, a statistical method patterned after El-Gamal and Grether (1995) can be adopted to draw inferences in this environment.

Again, our approach is less restrictive regarding the possible nature of subject heterogeneity, since we only need to specify *a priori* the possible state variables subjects might consider, and not the exact forms of possible decision rules. However, our approach would become impractical in problems where the set of state variables is sufficiently large that the curse of dimensionality in polynomial approximation becomes a problem.

In addition to the literature on subject "typing," another recent literature also attempts to advance beyond the "either/or" quality of earlier experimental work by specifying and estimating

---

[4] See Schachat and Walker (1997) and Houser and Winter (2002) for recent applications of this approach.

[5] Of course, the added flexibility our procedure comes at a cost: the loss of efficiency that comes from having to estimate the rule parameters rather than fixing them a priori. Thus, in many contexts it may be particularly efficacious to use our procedure and the El-Gamal and Grether procedure in concert. For instance, a preliminary application of our approach could determine the number and form of the candidate rules, and subsequent application of the El-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 certain 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*.

econometric models that may provide a better positive description of subjects' behavior. Much of this work is based on the McKelvey and Palfrey (1995) "quantal response equilibrium" model, in which subjects' best response functions or decision rules are subject to noise that can be interpreted as optimization error. Once subjects' choice behavior is assumed to be subject to noise, a natural next step is to estimate subjects' decision rules econometrically, including in the specification parameters that can capture various types of departure from "optimal" behavior (as defined by a particular normative model of play).[6] This work is similar in spirit to ours. But our approach is different in its emphasis on subject heterogeneity (i.e., to our knowledge the work based on McKelvey and Palfrey (1995) has assumed homogenous subjects) and in that we place less a priori structure on the potential departures from "optimality."

Our substantive contribution is to apply our type classification algorithm to data on a sample of subjects who we confront with a particular dynamic stochastic optimization problem. Subjects chose between two discrete alternatives in each of 15 time periods. In each period, stochastic payoffs are generated for each alternative according to rules that are explained to the subjects prior to the experiment. The problem is inherently dynamic, because current choices affect the distributions of future payoffs, and the optimal choice between the two options changes over time in a complex way as new information is revealed. By design, the problem is difficult in the sense that the optimal decision rule can only be obtained numerically via dynamic programming. But subjects were allowed to practice the game before playing for money.

We ran the experiment on 139 subjects, and our classification procedure produces a clear assignment of the population into only three distinct types. Statistical tests overwhelmingly reject the hypothesis that there are more, or that there are fewer. Furthermore, the subjects' posterior type probabilities usually assign a high probability to just one type, so our algorithm's prediction about a subject's type is typically unambiguous.[7]

Despite the difficulty of the game, more than one-third of the subjects adopted a decision rule that is very close to optimal. Payoff losses for subjects following this "near rational" rule (relative to what they could have earned by following the exactly optimal rule) averaged only about 2%. The other two types followed more clearly sub-optimal rules. By studying their fitted

---

[6] For instance, Goeree, Holt and Palfrey (2000a) adopt this framework to analyze "overbidding" in private value auctions relative to Nash predictions, while Goeree, Holt and Palfrey (2000b) analyze departures from Nash equilibrium behavior in matching pennies games.

decision rules, we are able to provide fairly simple characterizations of the play of each type.

The outline of this paper is as follows: In section 2 we describe the Bayesian algorithm for classifying decision rules in general terms. In section 3 we present our experimental design. In section 4 we show how decision rules in this particular experiment can be modeled using the general approach outlined in section 2. Section 5 presents results and section 6 concludes.

2. The Bayesian Classification Procedure

Our Bayesian approach to type classification enables us to draw inferences about the number and nature of decision rules present in a population of subjects, as well as the probability with which each subject uses each rule. Conditional on an assumed set of relevant state variables, each decision rule is modeled as a flexible parametric function. The number of decision rules operative in the population is determined using Bayesian decision theory (see Geweke (1997) for a recent exposition), which requires calculation of the marginal likelihood. For ease of exposition we will restrict attention to the class of discrete choice Markov decision processes (in discrete time), although in principle our approach has more general application.[8] Rust (1994) provides an excellent survey on structural approaches to inference for Markov decision processes.

We start by considering the *optimal* decision rule in a dynamic stochastic discrete choice problem. Applying Bellman's (1957) principle, the value to subject n of choosing alternative j from the discrete set $\{1, \frac{1}{4}, J\}$ in round t of a T period game can be written:

$$(1) \qquad V_{njt}(I_{nt}) = w_{njt} + EV(I_{n,t+1} \mid I_{nt}, j) \qquad\qquad t=1, \frac{1}{4}, T$$

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, given choice j. I_{nt} is the state of the subject in round t (i.e., the subject's information set). This might include, for example, the subject's 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 round given his/her current state and choice,[9] and $H(x)$ is the (possibly stochastic) Markovian law of motion for the state variables.

---

[7] This should not be confused with a statement that our type assignments are usually correct. In section 5.4 we present an extensive evaluation of the accuracy of our type assignments.

[8] For instance, Houser (2002) applies this approach to a model with mixed discrete/continuous choice variables.

[9] Notice we do not include a discount factor. In most experimental settings all payoffs are received at the same time (at the end of the game), so there is no discounting, or the time between rounds is trivial, so discounting is irrelevant.

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. The optimal decision rule is:

Chose alternative j iff $V_{njt}(I_{nt}) > V_{nkt}(I_{nt})$ for all k ^{1} j.

We assume payoffs are drawn from an absolutely continuous distribution to rule out ties.

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

Thus, 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. This follows the suggestion of Geweke and Keane (1999a, 2001). We also allow for the parameters of this function to differ across subjects of different type, denoted by k. And finally, we allow for the possibility of optimization error. Then, we can write the future component for type k as:

(2) $EV(I_{n,t+1} \mid I_{nt}, j) = F(I_{nt}, j \mid \boldsymbol{p}_k) + \boldsymbol{V}_{njt}$ $k=1, \frac{1}{4}, K.$

Here $F(\cdot)$ denotes the future component polynomial, and $\boldsymbol{p}_k$ denotes a finite vector of type specific parameters. The random variable $\boldsymbol{V}_{njt}$ accounts for idiosyncratic errors made when attempting to implement decision rule k.

We allow the distribution of the idiosyncratic errors to vary by type, so that optimization error may be more important for some types than others. Let $\boldsymbol{s}_k$ denote the standard deviation of the optimization error for type k. The assumption that optimization error may be present when implementing decision rules underlies much of the recent work in experimental economics that attempts to develop positive models of subject behavior (see, e.g., Goeree and Holt (1999)).

From (1) and (2), the value that subject n, who is type k, assigns to choice j in round t, is:

(3) $V_{njt}(I_{nt} \mid k) = w_{njt} + F(I_{nt}, j \mid \boldsymbol{p}_k) + \boldsymbol{V}_{njt}$

Denoting the deterministic part of the valuation function by $\overline{V}_{njt}(I_{nt} \mid k) \equiv w_{njt} + F(I_{nt}, j \mid \boldsymbol{p}_k)$, we have that the probability that option j is chosen is increasing in $\overline{V}_{njt}$. Letting optimization error

take the form of noise appended to the $\overline{V}_{njt}$ function is attractive, because it implies that the probability of a "mistake" in implementing the decision rule is small in situations where one alternative is clearly dominant in terms of its $\overline{V}_{njt}$. This seems intuitive.[10]

Specification (3) is quite flexible and can nest or approximately nest (i.e., via the Weierstrass approximation theorem) many special cases of interest. For example, if F is sufficiently flexible, and the $\boldsymbol{p}_k$ are chosen so that $F(I_{nt}, j \mid \boldsymbol{p}_k) \gg EV(I_{n,t+1} \mid I_{nt}, j)$, and we set $\boldsymbol{s}_k \gg$ 0, we can obtain a good approximation to the "optimal" value function. In fact, many authors have found that value functions in problems of interest to economists can be approximated very accurately using low order polynomials (see, e.g., Keane and Wolpin (1994), Krusell and Smith (1995), Geweke and Keane (1999a, 2001)). This is a key motivation for our approach.

Some other leading cases are also worth noting: Equation (3) nests myopic behavior if the $\boldsymbol{p}_k$ are set equal to zero and $\boldsymbol{s}_k = 0$. And it generates purely random behavior, in which each option is chosen with equal probability, if the $\boldsymbol{p}_k$ are set equal to zero and $\boldsymbol{s}_k$ is sent to infinity.

Specification (3) does not require that subjects understand what state variables are relevant in forecasting the value of future states. We have not made this explicit to conserve on notation, but we could easily allow for the possibility that the subjects consider superfluous information when making decisions, and denote the expanded information set by I^+_{nt}. Also, (3) may appear to impose additive separability, but this can be relaxed by including current and past payoffs and choices, and interactions between them and other state variables, as elements of I_{nt} or I^+_{nt}. And non-stationarity can be accommodated if the index for round t is an element of I_{nt}.

Now consider the problem of statistical inference in this framework. Let $d_{nt} \in \{1, ..., J\}$ denote the choice made by subject n in round t. Assume the investigator observes the set of choices and payoffs for each of the N subjects, $\{\{\{ d_{nt}, \{w_{njt}\}_{j=1, ¼, J} \}_{t=1, ¼, T} \}_{n=1, ¼, N}\}$. The goal to draw inferences about: (1) the order of the F polynomial and the set of state variables that enter the polynomial, (2) the number of decision rule types K that are present in the population of subjects, (3) the vector of parameters $\boldsymbol{p}_k$, $\boldsymbol{s}_k$ for each type $k=1, ¼, K$, (4) the population

---

[10] One must allow for optimization error in a model where K is less than the number of subjects, and payoffs are fully observed. Otherwise, the likelihood will equal zero for any subject whose behavior is not exactly explained by one of the K rules, leading to a degenerate model. El-Gamal and Grether (1995) dealt with this problem by introducing a fixed probability that a subject makes the "wrong" choice, given his/her decision rule. This has the implication that "wrong" choices are equally likely when "true" values of alternatives are close or far apart.

proportions of each type, which we denote by q_k, and (5) the posterior probability $p_n(k)$ that each subject is each type, conditional on his/her observed history of decisions and payoffs.

Consider first the simpler problem of drawing inferences about the parameters $\{p_k, s_k, q_k\}$ for $k=1, \frac{1}{4}, K$, *given* a particular choice for K, the order of F, and the set of state variables that determine F, as well as a distributional assumption on the optimization errors V_{njt}. Even this represents a fairly difficult inferential problem because we have a discrete choice model in which the type of each subject is a latent variable. However, recently developed simulation methods have made such models quite tractable (see, e.g., Geweke and Keane (2001) for a discussion). In section IV.D we describe a Gibbs sampling algorithm for Bayesian inference in this model.

Now consider the problem of drawing inferences about the number of types K and the order of the F polynomial, which we denote P. Given that we can implement a model with given (K, P), the standard approach of Bayesian decision theory is to implement a range of models with different K and P, and use the marginal likelihood to choose among them. Intuitively, a marginal likelihood is a likelihood-based measure of model fit that penalizes models for proliferation of parameters. In any discrete choice problem, calculation of the marginal likelihood represents a very high dimensional integration problem, and this problem is compounded by the presence of latent types. Again, recently developed simulation methods make this problem tractable. We describe our algorithm for calculating marginal likelihoods in Appendix A.[11]

It is straightforward to use marginal likelihood comparisons to choose the number of types K and the order P of the F polynomial. Simply increase K and/or P until the marginal likelihood begins to deteriorate, and stop there. But the use of the marginal likelihood to determine the set of state variables that agents consider when making decisions is not so mechanical. We can, of course, construct marginal likelihoods for models with and without certain extraneous state variables that agents might consider. But, ultimately, we can never know if we have failed to try some extraneous state variable that agents do in fact use. Of course, it is impossible to learn anything without some identifying assumptions. We would argue that any attempt to learn about decision rules will require some a priori narrowing of the set of potential state variables that subjects might consider.

---

[11] In contrast to our approach of using the marginal likelihood to choose among competing models, El-Gamal and Grether put a proper prior on the number of types, and then choose the number of types to maximize the posterior density of the model. Their prior explicitly favors models with fewer types.

Our approach also requires the investigator to make a distributional assumption on the optimization errors V_{njt}. However, in principle it would be straightforward to generalize normality by considering mixture-of-normals probit models, as in Geweke and Keane (1999b, 2001). Then, marginal likelihood comparisons could be used to choose the order of the mixture. This approach can parsimoniously capture a wide range of departures from normality.

In the next section we describe the specific experiment that we analyze, and in section 4 we detail how the general framework described here can be applied to that specific problem.

3. Experimental Design

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

1) It should be difficult to solve. The main point of our analysis is to examine how people behave when facing a decision problem that is too difficult for anyone to solve exactly.
2) We nevertheless wanted a problem whose structure was easy to explain to the subjects.
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 with features similar to a human capital investment or occupational choice problem. Each subject makes 15 sequential decisions. The problem is dynamic because early decisions influence the distributions of payoffs for later decisions. Period payoffs are stochastic, and the optimal decision rule, which can only be constructed numerically, is a function of the payoff draws. We set up the problem so that alternative "1" can be thought of as 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. It is important to stress however, that no such interpretation was provided to the subjects, who were simply told the mathematical form of the payoff function in each alternative.

A transcript of the written instructions given to subjects is provided in Appendix B. A precise description of the game is as follows: At the start of each of the 15 rounds, the subjects receive a draw for the payoffs in alternatives "1" and "2." The stochastic payoff to "2" is 4000 points plus the realization of a uniform random variable on the interval [-5000, 5000], subject to the restriction that the payoff be non-negative. The payoff draws for round t are seen before the decision at t is made, but the payoff draws for period $t+1$ are realized only after the decision at t.

The payoff to alternative "1" was 3000 points plus the realization of a uniform random variable on the interval [-5000, 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 added if the subject had chosen alternative "1" at least six, and no more than nine, previous times (not necessarily in succession). A "switching" cost of 5000 points was subtracted from the alternative "1" payoff if the subject had chosen "2" on the previous round. The subject's total payoff for the decision problem was the sum of the rewards they earned over the 15 rounds.

This game has, at least to some degree, all five desiderata listed above. First, it is a sophisticated dynamic investment problem that is nevertheless straightforward to describe. It is quite difficult to solve for the expected wealth maximizing strategy, which requires solving a dynamic programming problem numerically. The design also generates a non-negligible incentive for forward-looking behavior. The optimal solution earns about 25% more, on average, than the myopic strategy that simply chooses the highest payoff each period. Finally, the game only takes about 45 seconds to a minute to play, allowing ample opportunity for practice.

This game was coded in Visual Basic and subjects made decisions independently at a private computer terminal. Each subject's screen provides information on the current payoffs to both alternatives, the current round, the history of the subject's prior choices and payoffs, the subject's current aggregate earnings, and a summary of the decision problem's payoff structure.

We report on results obtained from 139 subjects who participated in this experiment, which was conducted at the Economic Science Laboratory (ESL) at the University of Arizona. Subjects were recruited from the general student population using ESL's standard procedures. In an effort to ensure subjects were familiar with the task when they played for money, subjects were recruited for two laboratory sessions. On arrival for the first session they were provided with the written instructions in Appendix B and seated privately at a computer terminal. They were allowed to practice as many times as they liked, but did not play for money.

The second lab session was held two days after the first. Upon arrival, subjects were again provided with the written instructions and told that they could again practice for as long as they liked. When they were ready, subjects solved the decision problem one time for money. During money play, we imposed a 15 second forced delay between moves, in order to discourage thoughtless play. In hindsight, since the typical subject voluntarily devoted a large amount of time to practice (i.e., the median number of practice rounds was 66), this concern seems to have been largely unwarranted. Subjects earned $8.23 on average. Subjects also received two five-dollar show-up fees, and spent about 75 minutes on average in the lab (in total).

4. Empirical Specification

In this section we first describe how we apply the general procedure outlined in section II to model the decision rules used by the subjects in our experiment. We describe the likelihood function, priors, and posterior distribution of the model parameters on which our inferences are based, and describe the Gibbs sampling algorithm that we use to sample from the posterior.

4.1. *The Functional Forms for the Decision Rules*

In our experiment, the current payoffs w_{njt} are simply the known immediate rewards that subject n draws for alternatives $j \hat{I} \{1, 2\}$ in round t. The rewards are drawn from a distribution that depends on the subject's state. 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 "1" and "2," 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 optimal play implies trying to get through this bonus phase before the end of the game.

Current payoffs are also elements of the information set I_{nt}, but they are not useful for forecasting future payoffs, since the stochastic components of payoffs 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}. A polynomial

approximation to RE future component would not depend on current and/or lagged w_{njt}. Nevertheless, it is interesting to examine whether "superfluous" state variables like these help to explain subject behavior. We tried including current/lagged payoffs and interactions of current with lagged choices in the future component polynomial to investigate whether some subjects might use this information erroneously. But this did not lead to significant improvements in model fit. Thus, we will only report on specifications in which the future component is a polynomial in state variables that are relevant for forecasting future payoffs.

To anticipate our results, our Bayesian model selection procedure selects a third order polynomial in the state variables as the preferred specification for the future component. Then, since the law of motion for the state variables is $H(X_1, X_2, j) = (X_1+I(j=1), X_2+I(j=2), j)$, the future component F for subjects of type k takes the form (suppressing the subscript k on the $\boldsymbol{p}_k$):

$$
\begin{aligned}
F(H(X_{n1t}, X_{n2t}, j) \mid \boldsymbol{p}) = {} & \boldsymbol{p}_0 + \boldsymbol{p}_1(X_{n1t} + I(j=1)) + \boldsymbol{p}_2(X_{n2t} + I(j=2)) \\
& + \boldsymbol{p}_3(X_{n1t} + I(j=1))^2 + \boldsymbol{p}_4(X_{n2t} + I(j=2))^2 + \boldsymbol{p}_5(X_{n1t} + I(j=1))(X_{n2t} + I(j=2)) \\
& + \boldsymbol{p}_6(X_{n1t} + I(j=1))^3 + \boldsymbol{p}_7(X_{n2t} + I(j=2))^3 \\
& + \boldsymbol{p}_8(X_{n1t} + I(j=1))^2(X_{n2t} + I(j=2)) + \boldsymbol{p}_9(X_{n1t} + I(j=1))(X_{n2t} + I(j=2))^2 \\
& + \boldsymbol{p}_{10}I(j=1) + \boldsymbol{p}_{11}I(j=1)(X_{n1t} + I(j=1)) + \boldsymbol{p}_{12}I(j=1)(X_{n2t} + I(j=2)) \\
& \quad + \boldsymbol{p}_{13}I(j=1)(X_{n1t} + I(j=1))^2 + \boldsymbol{p}_{14}I(j=1)(X_{n2t} + I(j=2))^2 \\
& \quad + \boldsymbol{p}_{15}I(j=1)(X_{n1t} + I(j=1))(X_{n2t} + I(j=2)) \\
& + \boldsymbol{p}_{16}I(j=2) + \boldsymbol{p}_{17}I(j=2)(X_{n1t} + I(j=1)) + \boldsymbol{p}_{18}I(j=2)(X_{n2t} + I(j=2)) \\
& \quad + \boldsymbol{p}_{19}I(j=2)(X_{n1t} + I(j=1))^2 + \boldsymbol{p}_{20}I(j=2)(X_{n2t} + I(j=2))^2 \\
& \quad + \boldsymbol{p}_{21}I(j=2)(X_{n1t} + I(j=1))(X_{n2t} + I(j=2))
\end{aligned}
$$

Note that the round t is linearly dependent on X_{n1t} and X_{n2t}, so we omit it from the polynomial. Since choices depend only on the relative values of "1" and "2," the future component is not identified in levels. Thus, our analysis is based on the differenced future component:

$$
\begin{aligned}
f(X_{n1t}, X_{n2t} \mid \boldsymbol{p}^*) \equiv {} & F(H(X_{n1t}, X_{n2t}, 1)) - F(H(X_{n1t}, X_{n2t}, 2)) \\
= {} & \boldsymbol{p}_0^* + \boldsymbol{p}_1^*(2X_{n1t} + 1) + \boldsymbol{p}_2^*(-2X_{n2t} - 1) + \boldsymbol{p}_3^*(3X_{n1t}^2 + 3X_{n1t} + 1) \\
& + \boldsymbol{p}_4^*(-3X_{n2t}^2 - 3X_{n2t} - 1) + \boldsymbol{p}_5^*(-X_{n1t}^2 + 2X_{n1t}X_{n2t} + X_{n2t})
\end{aligned}
$$

where the $\boldsymbol{p}_i^*$ are linear functions of the $\boldsymbol{p}_i$. The polynomials used here are a subset of those that appear in the differenced future component, but this set spans the same space.

The decision rule for subject n of type k in round t can be written:

Choose "1" iff $Z_{nt}(I_{nt}/k)$ º $V_{n1t}(I_{nt}/k) - V_{n2t}(I_{nt}/k)$

$$= w_{n1t} - w_{n2t} + f(X_{n1t}, X_{n2t} \mid \boldsymbol{p}_k^*) + \boldsymbol{h}_{nt} > 0$$

It is therefore intuitive to think of $f(X_{n1t}, X_{n2t} \mid \boldsymbol{p}_k^*)$ as a *reservation payoff differential*. This is the amount by which w_{n2t} must exceed w_{n1t} in order for the subject to chose "2" over "1" (subject to the added noise induced by the mean zero optimization error $\boldsymbol{h}_{nt} \equiv V_{n1t} - V_{n2t}$).

In our game, the optimal value of the reservation payoff differential varies in a complex way with the state variables X_{n1t} and X_{n2t}, making it difficult to play the game optimally. Our algorithm will allow us to infer the reservation payoff differential function 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.

4.2. *The Likelihood Function, Priors, and Joint Posterior Distribution of Parameters*

Recall that d_{nt} denotes the round t choice of subject n. Let $\boldsymbol{t}_n \in \{1, \frac{1}{4}, K\}$ indicate subject n's type. If subject n uses decision rule k (i.e., $\boldsymbol{t}_n = k$) and has information I_{nt}, then d_{nt} satisfies:

$$d_{nt} = \begin{cases} \text{"1" if } Z_{nt}(I_{nt} \mid k) > 0 \\ \text{"2" otherwise} \end{cases}$$

The investigator observes a sequence of current payoff realizations w_{njt} for $j=1,2$, $t=1$, $\frac{1}{4}$, T and choices d_{nt} for $t=1$, $\frac{1}{4}$, T for each subject n. Choices depend on the value function differences Z_{nt}, and the inferential problem is complicated by the fact that these, as well as the subject types $\boldsymbol{t}_n$, are unobserved. We assume $\boldsymbol{h}_{nt} \sim IIDN(0, \boldsymbol{s}_k^2)$ for type k. Thus, our model is formally a mixture of probit models, in which an additive part of the latent index, $w_{1nt} - w_{2nt}$, is observed. This sets the scale for the $\boldsymbol{p}_k$ and $\boldsymbol{s}_k$ parameters, so both are identified.

The probability that subject n of type k chooses alternative one in round t is:

$$P(d_{nt} = 1 \mid k) = P(V_{n1t}(I_{nt} \mid k) > V_{n2t}(I_{nt} \mid k)) = P(w_{n1t} - w_{n2t} + f(X_{n1t}, X_{n2t} \mid \boldsymbol{p}_k^*) + \boldsymbol{h}_{nt} > 0)$$

We define $P(d_{nt}/k) = I(d_{nt}=1/k) \times P(d_{nt}=1/k) + I(d_{nt}=2/k) \times [1 - P(d_{nt}=1/k)]$. Then, if all subjects' types were known, the likelihood function for the observed data would be simply:

$$L\ [\{\{d_{nt}\}_{t=1,T}\}_{n=1,N}\ |(\pmb{p}_k^*,\pmb{s}_k^{-2})_{k\in K}]=\prod_n\prod_k\prod_t P\ (d_{nt}\ |k)^{1(t_n=k)}.$$

However, since we do not know subjects' types, we must form a likelihood function based on unconditional choice probabilities $P\ (d_{nt})=\dot{\pmb{a}}_k\,\pmb{q}_k\,P\ (d_{nt}/k)$, where $\pmb{q}_k$ is the probability that a person chosen at random from the population follows decision rule k. This gives:

$$L\ [\{\{d_{nt}\}_{t=1,T}\}_{n=1,N}\ |(\pmb{p}_k^*,\pmb{s}_k^{-2},\pmb{q}_k)_{k\in K}]=\prod_n\left\{\sum_k\left(\pmb{q}_k\prod_t P\ (d_{nt}\ |k)\right)\right\}$$

Maximum likelihood (ML) estimation is problematic for two reasons. First, ML estimation of mixture models is notoriously difficult due to problems with local maxima of the likelihood function. Second, testing for the number of types is quite difficult in a classical framework.

Thus, we use a Bayesian Markov Chain–Monte Carlo (MCMC) algorithm to generate inferences about the model parameters $\{\pmb{q}_k,\ \pmb{p}_k^*,\ \pmb{s}_k^{-2}\}_{k=1,K}$. The particular MCMC algorithm we employ is the Gibbs sampler. This provides draws from the joint posterior distribution of the model parameters conditional on the data. Since these draws are obtained without the need to maximize the likelihood function, the Gibbs sampler is much less sensitive to problems created by ill behaved likelihood surfaces (e.g., local maxima) than is ML (see Geweke, Houser and Keane (2001) for a Monte-Carlo experiment that illustrates this point).

Inference via the Gibbs sampler starts with the specification of the *complete data likelihood function*, which is the hypothetical likelihood one could form if the latent indices Z_{nt} and the latent types $\pmb{t}_n$ were observed. In our model, given a particular K and P, this is:

$$L\ (\{Z_{nt}\}_{\substack{n=1,\dots,N,\\t=1,\dots,T}},\{\pmb{t}_n\}_{n=1,N}\ |\{\pmb{q}_k,\pmb{p}_k^*,\pmb{s}_k^{-2}\}_{k=1,K})\propto$$

(4)
$$\prod_{k=1,K}\prod_{n:t_n=k}\left[\pmb{q}_k\prod_{t=1,T}\frac{1}{\pmb{s}_k}\exp\left\{-\frac{(Z_{nt}-(w_{n1t}-w_{n2t}+Y_{nt}'\pmb{p}_k^*))^2}{2\pmb{s}_k^2}\right\}I(Z_{nt},d_{nt})\right].$$

where Y_{nt}' denotes the vector of state variables conformable with $\pmb{p}_k^*$, given P. 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 model is closed by specification of prior distributions for the model's parameters. We assume proper priors of a standard conjugate form for all parameters. These are as follows:

$$\boldsymbol{p}_k \sim N(0, \underline{\Sigma}), \text{ where } \underline{\Sigma} \text{ is a 6 x 6 diagonal matrix with entries } \underline{\Sigma}(1,1) = 20{,}000^2,$$

(5)
$$\underline{\Sigma}(2,2) = \underline{\Sigma}(3,3) = 1{,}000^2, \underline{\Sigma}(4,4) = \underline{\Sigma}(5,5) = 100^2 \text{ and } \underline{\Sigma}(6,6) = 100{,}000.$$

$$\boldsymbol{s}_k^{-2} \sim \boldsymbol{c}^2(1), \boldsymbol{s}_1^2 > \boldsymbol{s}_2^2 > \cdots > \boldsymbol{s}_K^2$$

$$\{\boldsymbol{q}_k\}_{k \in K} \sim Di(\{2.0\}_{\|K\|}).$$

where *Di* is the Dirichlet distribution, or multivariate Beta.

By Bayes theorem, the joint posterior of the model parameters, the latent indices Z_{nt} and the latent indicators $\boldsymbol{t}_n$ is simply proportional to the complete data likelihood times the prior densities $p(\boldsymbol{p}_k^{\,*})$, $p(\boldsymbol{s}_k^{-2})$, and $p(\boldsymbol{q}_k)$. Since we have proper priors and a bounded likelihood function, the joint posterior exists. This is a necessary condition for convergence in distribution of the Gibbs sampler draw sequence to the appropriate joint posterior.

A number of aspects of the prior specification are worth commenting upon. First, note that setting the prior mean for the $\boldsymbol{p}_k^{\,*}$ vector at zero means that our prior is centered on myopia. When $\boldsymbol{p}_k^{\,*} = 0$ subjects only consider current payoffs. The issues involved in choosing $\underline{\Sigma}$ illustrate why it is not possible to have "uninformative" priors. If the priors on the $\boldsymbol{p}_k^{\,*}$ were very flat (i.e., if the elements of $\underline{\Sigma}$ were very large), it would imply little prior mass in the vicinity of $\boldsymbol{p}_k^{\,*} = 0$. Thus, although the prior would be centered on myopia, it would say there is little prior mass on the myopic decision rule. Thus, the choice of $\underline{\Sigma}$ must be considered rather carefully.

Regarding the choice of $\underline{\Sigma}(1,1)$, note that $\boldsymbol{p}_{k0}^{\,*}$ is the $f(X_{n1t}, X_{n2t} / \boldsymbol{p}_k)$ function intercept, and is therefore the reservation payoff differential in round 1. Under the optimal decision rule, this is 3733 points (i.e., 37 cents). We specify a prior mean of 0 and a prior standard deviation of 20,000 for $\boldsymbol{p}_{k0}^{\,*}$. Thus, our prior is rather diffuse, but still leaves non-negligible mass in the vicinity of the interesting special cases of myopic and optimal play.

Second, we have a strong prior that higher order polynomial coefficients should be relatively smaller in magnitude, simply due to scale (i.e., they multiply variables that tend to be larger in magnitude). A prior with equal diagonal elements for $\underline{\Sigma}$ would place most prior mass on models where the higher order terms in the state variables dominate decisions, which is not plausible. For this reason, we put successively tighter priors on the higher order terms (e.g., prior standard deviations of 1000 on the linear X terms, 100 on the X^2 terms, etc.).

In our empirical results below, we find that the posterior mean for each of the $\boldsymbol{p}_k^{\,*}$ parameters is within one prior standard deviation of our prior mean. We also find that in every

instance the posterior standard deviation is about 8 to 200 times smaller than our prior standard deviation. Together, these results suggest that the data is very informative about all the $\boldsymbol{p}_k^*$ parameters, and that in no instance was the prior "too tight" to "let the data speak." We also report below on the sensitivity of our inferences to both reducing and increasing the prior standard deviations on all the $\boldsymbol{p}_k^*$ parameters by 50%, and find little effect.

Third, consider the prior on the optimization error variance, $\boldsymbol{s}_k^2$. The ordering restriction on the type specific variances is simply 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) for a discussion).

The prior mean and standard deviation for $\boldsymbol{s}_k^2$ are undefined under the $1/\chi^2(1)$ prior. But we can consider quantiles of the $\boldsymbol{s}_k^2$ distribution. Our prior puts 95% of the mass on models where the optimization error standard deviation is less than roughly 16 experimental points. This is quite small relative to the magnitudes of payoffs. We did this to favor models where the state variables largely explain behavior, as opposed to letting behavior be largely random. But the prior density has a very fat right tail, so models with large optimization errors are still given non-negligible prior mass. The prior is also quite weak. As we'll see in section IV.C., it has an impact on inference equivalent to adding a single observation with a squared error term equal to one.

We also report below on the sensitivity of our inferences to scaling up the $\boldsymbol{s}_k^2$ prior to $20,000/\boldsymbol{c}^2(1)$, which increases the 95th percentile point to roughly 2250 experimental points. This corresponds to roughly the highest level of optimization error we found in any of the type specific decision rules. We find little effect of this change in prior on our inferences.

Fourth, the Dirichlet prior is centered on equal type proportions (i.e., *1/K* each) but it is sufficiently diffuse that models with very unequal proportions will have substantial prior mass. For example, with three types, the prior standard deviation on the type proportions is 18 percent.

4.3. *The Gibbs Sampling Algorithm*

We now describe the Gibbs sampling algorithm that we use to approximate the marginal posteriors of the model's parameters.[12] The product of the complete data likelihood (4) and the set of prior densities implied by the prior structure in (5) define the joint posterior used to construct the Gibbs sampler. The sampler includes the following five steps:

1) Draw latent utility values Z_{nt}.

2) Draw decision rule coefficients $\boldsymbol{p}_k^*$ for all $k = 1, K$.

3) Draw variance of optimization error $\boldsymbol{s}_k^2, k = 1, K$.

4) Draw population type probabilities $\boldsymbol{q}_k, k = 1, K$.

5) Draw individual types $\boldsymbol{t}_n, n = 1, N$.

These draws are implemented as follows:

1) $\underline{Z_{nt}}$ Step: Conditional on everything else being known, (4) implies that Z_{nt} is a truncated

   $N[(w_{n1t} - w_{n2t} + Y_{nt}'\boldsymbol{p}_k^*), \boldsymbol{s}_k^2]$, with truncation from below at zero if $d_{nt} = 1$, and from

   above at zero otherwise. We draw from this distribution using an inverse CDF procedure.

2) $\underline{\pi_k^*}$ Step: Conditional on everything else being known, the $\boldsymbol{p}_k^*$ vectors can be drawn using

   rejection methods. Based on (4), the source distribution is $N[(Y_k'Y_k)^{-1}Y_k'W_k, \boldsymbol{s}_k^2(Y_k'Y_k)^{-1}]$.

   Here, Y_k denotes the stacked array of Y_{nt} vectors for the type k subjects, and W_k is created

   by stacking the quantities $Z_{nt} - w_{n1t} + w_{n2t}$ conformably. Take a candidate draw for $\boldsymbol{p}_k^*$

   form this source distribution. Evaluate the kernel of the normal prior density for $\boldsymbol{p}_k^*$ at

   this draw, and call this u. Note that $u \in (0,1)$. Obtain a draw U from a uniform $[0,1]$

   distribution. If $U < u$ the draw for $\boldsymbol{p}_k^*$ is accepted. Otherwise it is rejected.

3) $\underline{\sigma_k^2}$ Step: Conditional on everything else being known, equations (4) and (5) imply that:

$$\frac{1 + \sum_{\substack{t=1,T \\ n=1,N}} (Z_{nt} - (w_{n1t} - w_{n2t} + Y_{nt}'\boldsymbol{p}_k^*))^2 I(\boldsymbol{t}_n = k)}{\boldsymbol{s}_k^2} \sim \boldsymbol{c}^2(N_k T + 1),$$

   where N_k is the number subjects who are type k. We draw from this distribution using

   standard software, and reject any draw that does not satisfy the ordering $\boldsymbol{s}_1^2 > ... > \boldsymbol{s}_K^2$.

4) $\underline{\boldsymbol{q}_k}$ Step: The prior for $\boldsymbol{q}_k$ is $Di(\{2\}_K)$, so the conditional posterior is $Di(\{2 + N_k\}_{k=1,K})$. We

   draw from this Dirichlet distribution using standard procedures.

5) $\underline{\boldsymbol{t}_n}$ Step: Let $L_k(n)$ denote the likelihood contribution for subject n given that he/she uses

   decision rule k and with everything else is known. Then, the distribution of $\boldsymbol{t}_n$ is

---

[12] Our FORTRAN 77 code, which makes extensive use of IMSL subroutines, is available on request.

$$\Pr(t_n = k') = \frac{\mathsf{L}_{k'}(n)}{\sum_{k=1,K} \mathsf{L}_k(n)}.$$

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

Finally, we consider the choice of the number of types K and the polynomial order P. We consider models with several different values of K and P, and then use Bayesian decision theory to choose among the models. This requires construction of the marginal likelihood for each model. The procedure for constructing marginal likelihoods is described in Appendix A.

5. Empirical Results

5.1. *A Basic Description of the Data*

We begin by simply comparing the average behavior of our 139 subjects to "optimal" behavior in the experiment. To do this, we construct choice histories for 139 hypothetical agents who face the exact same realizations for the random components of payoffs as did the human subjects.[13] But our hypothetical agents play the exactly optimal decision rule - which we call the "rational expectations" (RE) decision rule.[14] We will refer to these as "RE subjects."

Figure 1 compares the fraction of actual and RE subjects who choose alternative "1" in each round of the game. In round 1, the median payoff for "2" is 1000 points higher than for "1." Yet, quite interestingly, over 75% of actual subjects chose "1." This implies that most subjects understand the investment component of option "1," and realize that they should choose "1" unless "2" offers a substantial payoff premium. The fraction of RE subjects who choose "1" in round 1 is 79%, which is very close to the fraction of experimental subjects who chose "1."

Figure 1 also reveals an interesting non-stationarity in choice behavior in this game. The fraction choosing "1" if subjects play optimally should drop after round one, fall to a trough in round 3, rise (non-monotonically) to a peak in round 11, and then drop off rather sharply at the

---

[13] Period payoffs depend on both the random draws and the past history of choices. So, of course, even though actual subjects and their RE counterparts face the same random draws, their alternative specific payoffs in a particular round may differ because they have different choice histories.

[14] To be precise, we construct the decision rule that would be used by rational, *expected wealth maximizing* agents. We ignore the possibility of risk aversion when constructing the optimal decision rule, because the payoffs at stake in the game are rather small. Note, however, that our inferential procedure does not impose risk neutrality on the estimated decision rules of the experimental subjects. Departures from risk neutrality would be captured by the π's.

end.[15] The aggregate choice frequencies of subjects in the experiment match this complex pattern implied by the optimal rule quite closely over the first six rounds. But from rounds seven to nine, actual subjects chose alternative "1" slightly more frequently than their RE counterparts. And from round 11 onward, the choice frequencies diverge substantially. The actual subjects chose option "1" much less frequently than the RE subjects during the later rounds of the game.

On average, RE subjects choose "1" 10.7 times during the game, compared to only 10.0 times for actual subjects. While 94% of RE subjects complete the "bonus phase," only 64% of actual subjects do so. The actual subjects earn 10.0% less than the RE subjects on average.

But averages reveal little about the play of individual subjects, because there is substantial variation in play around the averages. For instance, while on average our subjects chose "1" less often than the RE subjects, 38 subjects chose "1" more often, and 44 chose "1" exactly as often. This is consistent with the notion of decision rule heterogeneity.

5.2. *Model Selection and Evaluation of Fit*

Next, we use the Bayesian type classification procedure described in Section IV to learn about the nature and number of decision rules operative in the population. We implemented the our Gibbs sampling algorithm on models where the number of decision rule types K ranged from one through four, and in which the order P of the future component polynomial F ranged from 3 through 5. Thus, the order of the differenced future component ranged from quadratic through quartic. Table 1 contains marginal likelihood comparisons among the 12 candidate models. The model with 3 types and a quadratic for the differenced future component is clearly preferred.

In the bottom panel of Table 1 we check the sensitivity of our model selection to the choice of prior. We report marginal likelihood values under alternative priors in which: (1) the posterior standard deviations of the future component polynomial parameters $\boldsymbol{p}_k^*$ are reduced or increased by 50%, and (2) the scale of the optimization errors is increased by a factor of 141 = $(20000)^{1/2}$. The three-type model with a quadratic differenced future component is strongly preferred regardless of the prior specification.

For the preferred model, our inferences are based on the final 5,000 cycles of an 8,000 cycle Gibbs sampler run. Inspection of the draw sequence, as well as the split-sequence

---

[15] The intuition for this pattern is as follows: Early in the game, when there is plenty of time left to accumulate the six choices of "1" needed to reach the bonus phase, the "urgency" for choosing "1" is not great. But as the game

diagnostic suggested by Gelman (1996), convinced us that the Markov chain had converged by the 3,000[th] cycle. The draws from the last 5,000 cycles are used to estimate the posterior means and standard deviations of the model parameters. These are reported in Table 2.

Table 2 reports on the parameters of each of the three decision rules, as well as the population type proportions. The table also reports our prior mean and standard deviation for each parameter. The posterior standard deviations are quite small relative to the prior standard deviations, indicating that the data is very informative about the parameters. Also the posterior means are all within a fraction of a *prior* standard deviation of the prior means. This suggests that the prior did not strongly influence the posterior for any parameter.

Of course, as is usually the case, the polynomial coefficients in the decision rules are difficult to interpret. Thus, we will leave the coefficients largely uncommented, and instead turn to simulations of behavior under the each of the three rules in order to understand their behavioral implications. But the intercept is an exception, since it has a clear interpretation as the reservation payoff differential in round 1 (when all state variables equal zero), and it is 3733 under the "optimal" rule. In Table 2, the posterior mean for $\boldsymbol{p}^{*}_{k0}$ is rather close to 3733 for all three types. Thus, it appears that all three types play nearly optimally in round 1. They all understand that there is an "investment" value to choosing "1" in the first round.

In order to characterize the behavior of each type, and to better assess the fit of the three-type model to the data, we assigned each subject to a type based on his/her highest posterior type probability. Recall that step 5 in the Gibbs sampling algorithm in Section IV.C is to draw a subject's latent type. The fraction of draws in which a subject is assigned to a particular type is a simulation consistent estimator of the posterior probability that the subject is that type.

The vast majority of subjects can be assigned to one type very clearly, because the highest posterior type probability is at least 90% for 86.3% of the subjects. This means that a subject's choices in our experiment are usually highly informative about his/her type. In hindsight, this is not surprising. For example, if a subject consistently received good payoff draws for option "1," it would be easy for him/her to make optimal choices, and his/her history would not be very revealing. But, since draws are iid and each subject must make 15 decisions, such a scenario is highly unlikely. Most subjects have to make at least a few "tough" choices

---

progresses, the urgency to choose option "1" tends to increase, and the reservation payoff differential between "1" and "2" should grow, holding the number of times one has chosen "1" constant.

during the course of the game, so there is usually plenty of opportunity to reveal one's type.

Figure 2 assesses the fit of the three estimated decision rules to the actual play of the subjects who we classify as following each rule. To do this, we simulated the hypothetical decisions that each subject would have made under his/her assigned decision rule, given the realizations of the payoff draws that he/she actually experienced in each round. Figure 2 reports the fraction of actual and hypothetical subjects of each type who choose "1" in each round. The main features of each type's play are well matched by the simulated choices. For example, type one subjects are extremely likely to choose option "1" in rounds 9 through 11 (over a 95% chance), while type 2 subjects only choose "1" about 80% of the time in those rounds. Our fitted decision rules capture this difference in behavior rather well. The type three subjects choose option "1" much less often during rounds 9 to 11 (i.e., about 70% in round 9 and only about 50% in rounds 10 to 11). The model captures this basic pattern, although it somewhat understates the degree of the difference (i.e., it predicts that types 3s would choose "1" about 60% of the time in rounds 10-11). In general, the fit for type three is not quite as good as that for types one and two.

The gray line in Figure 1 shows the fit of the three-type model to the aggregate choice frequencies. It shows the fraction of hypothetical subjects who choose alternative "1" in each round. The fit is reasonably good in all rounds, with the broad features of actual decisions, such as the peak that occurs in the ninth round, well matched by the model. The model captures the departure of actual play from RE play that occurs beginning in round 11 quite well.

5.3. *Characterization of the Decision Rules*

In this section, we attempt to characterize the nature of decision rule used by each of the three types. Table 3 compares the play of the subjects we assign to each type along a number of dimensions. There is a clear ranking of the types in terms of how well they play the game. The subjects who we classify as type one do best. On average, they earn 87983 experimental points, or about $8.80. We simulate that hypothetical RE subjects, facing the exact same random draws, would earn about nine dollars on average. Thus, on average, type one subjects only lose about 21 cents, or 2.3%, of what they could have earned by playing exactly optimally. In contrast, type two subjects lose 11.7% and type three subjects lose 18.6%.

Next, to get a better sense of the behavior implied by each decision rule, we simulated the play of hypothetical subjects under each rule. 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 draw sequences, 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, and a fifth in which the subjects are myopic (i.e., they chose the highest payoff alternative in each period).

Figure 3 describes the results of this simulation exercise. It reports the fraction of hypothetical subjects of each type who choose "1" in each round. Note that rule one tracks the optimal RE rule quite closely. These two rules imply nearly identical behavior through the first eight rounds. In rounds 9 through 11 rule one generates a slightly higher probability of choosing "1" than does the RE rule, and in rounds 13 through 15 it generates a slightly lower probability, but these differences are fairly minor.

Rule two also tracks the RE rule quite closely through the first several rounds. But beginning at round nine it starts to generate a lower frequency of option "1." This divergence becomes greater as the game progresses, and becomes quite dramatic in rounds 12 to 15. Over the last four rounds rule two generates a 20 to 35% lower frequency of option "1" than does the RE rule. It is interesting that both rules one and two capture the complex non-stationary pattern in choice behavior that is implied by the optimal rule, whereby the choice frequency for option "1" drops to a trough in round 3, and gradual rises to a peak later in the game. Rule one gets the timing of the peak exactly right (round 11), while rule two misses the peak slightly.

Rule three diverges in much more obvious ways from the RE rule. Interestingly, it does track the RE rule quite closely for the first three rounds. It generates a similar high frequency for option "1" in round one, and the sharp drop off to a trough in round 3. However, under rule three, the fraction choosing "1" does not recover later in the game. By the sixth round rule three finds only half of the subjects in alternative "1," while the other rules lead about 75% of the subjects to make this choice.

A critical point is that myopic play differs substantially from all three rules. Under the myopic rule, only 40% of subjects would choose option "1" in the first round, since it has a lower median payoff. But, as we've noted, all three types choose option "1" at close to the optimal frequency of 79% in the first round. Thus, all three types recognize that there is an

investment component to the choice of option "1."

The myopic rule generates a drop in choice frequency for option "1" over the first three rounds, just like the RE rule.[16] But, unlike the RE rule, it does not generate the subsequent rise in the fraction choosing "1." It is interesting that the type three rule also fails to generate this increase in the frequency of "1" after round three. Thus, it seems that type three subjects play close to optimally at the start of the game, but behave more myopically as the game progresses.[17]

A good way to gain greater insight into the behavioral implications of the different decision rules is to compare the features of the estimated future components graphically. An important aspect of our procedure is that, by providing estimates of the future component for each type, it allows the investigator to perform such a graphical comparison.

Figure 4A graphs the reservation payoff differentials for each type at selected state points. Values are plotted for round one (when there is only one possible state point) and for round five (when there are four possible state points, $X_{n1t} = 0$, 1, 2, 3 or 4). In each case the vertical axis denotes the value of the reservation payoff differential. Larger values indicate that option "1" has a larger continuation value relative to option "2," so a greater current payoff differential is needed to induce choice of "2." The horizontal axis indicates X_{n1t}, which is the only state variable relevant for calculating the reservation payoff differential. The figure also graphs the reservation payoff differential under the RE rule for comparison purposes.

Consider first the type one subjects, who are described in the top panel of Figure 4A. The reservation payoff differential for type ones is about right in round one. In round five, their reservation payoff differential is declining in the stock of "experience" in option "1." This pattern also holds under the RE rule. Also note that, holding the number of prior choices of "1" fixed at zero, the optimal (RE) reservation payoff differential grows from 3733 to about 6000 as we move from round one to round five. Intuitively, this occurs because, as rounds go by, time is growing short to accumulate the six choices of "1" needed to reach the bonus phase. So the urgency to choose option "1" is growing. The type one subjects appear to understand this very

---

[16] This drop is driven by the current payoff structure. Since there is a transition cost for moving into "1" from "2," but not vice versa, the median payoff for option "1" drops in round two (since some people chose "2" in round one).
[17] The astute reader may notice an apparent contradiction between the behavior of the type three subjects in Figure 3, compared to their behavior as described in Figure 2 and Table 3. Figure 3 implies that type threes choose "1" much less often than other types. But Table 3 indicates that type threes chose "1" more often than type twos. Figure 2 indicates that type threes chose "1" particularly often in rounds 5-7. The explanation is that actual type three subjects were "lucky" in that they got statistically significantly better than average draws for "1" in rounds 5-7. This induced them to chose "1" very frequently in those rounds. We discuss this further in section 5.4.

well. They get the reservation payoff differential almost exactly right at state points $X_1 = 1, \ldots, 4$ in round 5. They appear to be a bit off at $X_1 = 0$, but this estimate is noisy due to limited data at that point (i.e., type ones rarely get to round five without choosing option "1" at least once).

Figure 4B provides similar information for rounds 9 and 13. Again consider the type one subjects, who are described in the top panel. The close agreement between their fitted decision rule and the RE rule is again quite remarkable, with one important exception. If a subject is in round 9 and has chosen option "1" either 6, 7 or 8 prior times, then that subject is in the bonus phase of the game, in which the option "1" base payoff is raised 7500 points. Recall that the bonus phase lasts until the subject has chosen option "1" four more times – not necessarily consecutively. Thus, there is an option value to choosing option "2," because it prolongs the bonus phase. As a result, the reservation payoff differential goes negative at this point according to the optimal (RE) rule. That is, one should actually demand a premium to choose option "1." As we see in Figure 4B, the type one subjects do not understand this. They continue to set a small but positive reservation differential during the bonus phase.

One might worry that the failure of the fitted decision rule to align well with the RE rule in period 9 at states $X_1 = 6, 7, 8$ could stem from a failure of our polynomial approximation in this range. But inspection of the individual level data suggests that this is not the case. In fact, once the bonus phase is under way, no subject in our entire data set ever chooses option "2," even when it is optimal to do so. Thus, the failure to understand the option value of "2" during the bonus phase is clearly a feature of the data, and our polynomial approximation to the future component accurately reflects it. [This finding illustrates the power of our approach. We would not have thought to look for this pattern in the data unless Figure 4B had pointed us toward it.]

Based on the evidence presented so far, we decided to label the type one subjects as "Near-Rational." They use a decision rule that is nearly identical to that of hypothetical RE subjects, except during the bonus phase, when they fail to grasp the option value of "2." But this failure only costs them about a 2.3% loss in earnings, on average, relative to what they could have earned by playing exactly optimally. Understanding the investment value of "1," and how this varies over states, is much more important for overall success in this game, and the type one subjects grasp this feature of the game very well.

Now consider the type two subjects, whose behavior is described in the middle panels of Figures 4A and 4B. Figure 4A indicates that these subjects value alternative "1" in about the

same way as RE subjects in the first round of the game. However, by round five it is apparent that their reservation payoff differentials differ from the RE values in both level and shape. While the RE future component assigns less value to option "1" as experience in "1" increases, type two subjects do exactly the opposite.

Based on this inversion in the shape of their future component, we decided to label the type two subjects "Fatalists." The reason is as follows: If type two subjects don't choose option "1" many times in the early part of the game (say, because they happen to get a set of good payoff draws for option "2"), they reduce their reservation payoff differential for choosing "2" rather than increasing it. Thus, just when they should be increasing the urgency with which they attempt to choose "1," they instead reduce it. This means, in effect, that if they don't happen to choose "1" a few times early in the game, they start to "give up" on reaching the bonus phase.

In contrast, suppose a type two happens to choose "1" a few times early in the game (say, because he/she happens to get some good payoff draws for option "1"). The optimal rule says to reduce the reservation payoff differential, because, loosely speaking, you can now "relax" because you will almost surely get to the bonus phase before the end of the game even if you choose "2" the next few rounds. But a type two acts differently. He/she shows a greater urgency to chose "1" in this case. This means, in effect, that once the bonus phase appears to be easily "within reach," type two subjects strive harder to reach it.

This behavioral pattern seems to be well described by "fatalism," meaning that type two subjects assign too much significance to the luck of the draw in determining the outcome of the game. They fail to appreciate that by properly modifying the reservation payoff differential as the state evolves, one can almost guarantee that one will reach and complete the bonus phase before the end of the game (i.e., 94% of hypothetical RE subjects complete the bonus phase).

Finally, we turn to the type three subjects, whose behavior is described in the bottom panels of Figures 4A and 4B. As we have already noted, type three and RE subjects have about the same reservation payoff differential in first round. What is interesting is the difference that emerges in round five. The type threes let their reservation payoff differential for choosing "2" decline much too rapidly as they accumulate experience in "1." Then, in round nine, their reservation payoff differential is much too small all states in which $X_1 < 6$ (i.e., before the bonus round has started). It appears that type threes quickly become overconfident about reaching the bonus phase if they happen to choose "1" a few times early in the game. And then they seem to

largely forget about the investment value of option "1" by round nine. In light of this pattern, we decided to label the type three subjects "Confused." Also consistent with this characterization is that σ_k is much larger for type threes than for the other types (see Table 2).

5.4. *Addressing Some Common Concerns*

We have heard four common concerns expressed regarding latent type classification algorithms in general, and our type assignments in particular. The first is a concern, arising out of work by Heckman and Singer (1984), that in mixture models the type specific parameters tend to be poorly estimated, and the number of types tends to be underestimated. To help dispel this concern, we present two small Monte-Carlo experiments, which suggest that our algorithm does do a good job of accurately uncovering the latent types that exist in a population.

In the first experiment, we generated a hypothetical sample of 139 subjects, using the estimated decision rules and type proportions from Table 2. In the second experiment, we constructed an artificial sample of N=200 in which 5 decision rules are operative, and type proportions are 20% each. The 5 rules include the three rules from Table 2, along with two new rules. We added a myopic type, for whom the future component is zero, and a "future oriented" type. The latter have double the intercept of the "Near-Rational" future component, and the remaining $\boldsymbol{p}_k^*$ values set to zero. We ran the Gibbs sampler on each sample, using the same baseline prior as given in (5), and using *P*=3. The results are reported in Table 4.

In Table 4, the marginal likelihood correctly chooses the three-type model on the three type data set, and the five type model on the five type data set, so there is no tendency to underestimate the number of types. The estimated decision rules for each type closely resemble the actual decision rules used to generate the data, but we do not report the large number of type specific polynomial parameters to conserve on space. Furthermore, the models assign subjects to types with a reasonably high degree of accuracy. In the three-type model, the Near-Rational subjects are assigned to the Near-Rational, Fatalist and Confused decision rules with posterior probabilities of .703, .293 and .004 on average. The comparable numbers for Fatalist subjects are .244, .690 and .066, and those for the Confused subjects are .046, .209 and .747. What drives the difference between our results and those of Heckman and Singer is that our experiment is structured to be much more informative about a subject's type (i.e., we observe 15 choices for each subject while they consider only search durations).

A second common concern is that the draw sequence that a subject receives may somehow make him/her appear to be a particular type. For example, a person who was lucky enough to get all good draws for option "1" would never face a "tough" decision where the RE rule says "1" is optimal even though "2" has a considerably higher payoff. It would then be "easy" for this person to make optimal decisions. But this argument shows a misunderstanding of both our classification algorithm and the game. First of all, our algorithm would *not* clearly classify such a person as "Near-Rational." Give such a history, the likelihood of this person's draw sequence would be high under any type. Thus, our algorithm would conclude that the data is uninformative, and the person's type classification would be ambiguous. Second, given the length of the game, such a scenario is highly unlikely. Since the game is 15 rounds, and payoff draws are i.i.d., almost every subject faces at least a few "tough decisions" where the RE rule implies the two alternatives have close to the same value, even though one may have a much better current payoff than the other.

We can also examine directly whether the type of draws that subjects received tended to differ by type. In Table 5 we report the mean draws for option "1" and "2" among subjects who were classified as each of the three types. If, for example, the "Near-Rational" types tended to get relatively good draws for option "1" and poor draws for option "2," making it "easy" for them to choose "1" frequently and hence reach the bonus phase, this would show up when we look at the mean draws. In fact, the only significant departure from mean zero draws was among subjects classified as "Confused." These subjects actually got relatively good draws for option "1" and poor draws for "2," although the latter is not significant. We interpret this as simply a chance outcome. It is hard to develop a story in which getting good draws for option "1" would make the game harder and somehow induce poor play.

One might also wonder if draws in early rounds are particularly important. For example, perhaps a subject who gets good draws for "1" in the first few rounds will start down the path of investing, while a subject who gets good draws for "2" early will become myopic. To address this concern we ran logits for whether a subject was one of the three types on the subject's draws in the first few periods. Early draws were not significant predictors of type.

The third common concern is that the behavioral types we have uncovered do not really reflect differential ability of the subjects to perform the task, but merely differential effort. A dogmatic defender of complete rationality might well argue that all subjects were, in principle,

capable of solving our decision problem, and that those who performed poorly in the game were simply not trying. This might occur because the expected rewards to good performance were too small to elicit substantial effort. We believe that the data on practice rounds can be used to address this question. In Table 6 we present data on the number of times subjects of each type practiced. The mean number of practice rounds is 66 (the median is 58). This fact is in itself significant, because it suggests that the typical subject devotes a substantial amount of time to practice. Given that it takes roughly 45 to 60 seconds to play the game, the typical subject is devoting roughly 45 to 60 minutes to practice.

Even more interesting is the ordering of practice rounds by type. "Near-Rational" subjects practiced 58 times on average, while "Fatalists" practice 69 times and "Confused" subjects practice 75 times. The medians were 48, 59 and 69, respectively. Thus, the subjects who perform worst in the experiment actually devoted the most effort to practice. It therefore appears difficult to rationalize their poor performance by lack of effort.

The fourth common concern is that the behavioral types we uncover are not stable characteristics of subjects. Would subjects be consistently assigned to the same type if they played the game multiple times? We can't address that question directly, but the practice round data can shed some light on the issue. We don't expect that the subjects would use the same decision rules in the practice rounds as in the money round, simply because rational behavior would imply using the practice rounds to experiment. However, if the types are stable characteristics of subjects, we would expect to see the types practice differently.

In Figure 5 we report information on the how well the subjects performed in the practice rounds. Specifically, for each type, we report mean earnings in each practice round. Quite interestingly, the relative performance of the three types in the very first practice round on day one is almost exactly the same as in the money round. The same is true of their relative performance in the first practice round on day 2. Thus, there is clear continuity of subject type behavior between the practice rounds and the money round.

Note that all three types of subjects do noticeably better when they play for money. Also note that the "Near-Rational" subjects, after doing clearly better (on average) than the other types for the first several practice rounds on day one, are less clearly superior in subsequent rounds on that day. Both these observations suggest that subjects are not trying to maximize points in each practice round, conditional on their knowledge of the game up to that point. Rather, since no

money is at stake, they may be making moves designed in part to enhance their knowledge.

We can gain further insight into the relation between practice and money play by running our algorithm on some of the practice round data. We decided to run the algorithm on data from the first two practice rounds on day 2. By then, subjects should be familiar with the game. Also, it seems that performance deteriorates during a subject's last few practice rounds (perhaps due to boredom). Thus, we anticipated that behavior in the first two practice rounds on day 2 would most resemble money play. To examine stability of types, we allowed a subject's type to differ between the two practice rounds (i.e., we pretended we had data on 2·139 = 278 subjects).

We again found clear evidence for 3 distinct decision rules.[18] To characterize these rules, we assigned each of the 278 "subjects" to a rule, and constructed the same statistics as we reported in Table 3. The type one subjects complete the bonus phase 95% of the time, compared to 94% for the Near-Rational subjects in money play, and they have the smallest payoff loss relative to RE play of the three types.[19] Thus, it again makes sense to describe their decision rule as "Near-Rational." The type two subjects have exactly the same mean payoff loss relative to RE play as the Fatalist subjects in money play (11.7%), and they get through the bonus phase with roughly the same frequency (60% vs. 69%). The type three subjects again do considerably worse than the other two types (an 18.5% payoff loss vs. 18.6% for the Confused subjects in money play), so it again makes sense to describe their decision rule as "Confused."

A clear difference between practice and money play is that the percent of subjects assigned to the three rules - which, given their broad similarities to the money play rules, we will continue to call Near-Rational, Fatalist and Confused - are 26%, 30% and 44%, respectively, as opposed to 37%, 40% and 24% in money play. Thus, not surprisingly, subjects are considerably more likely to play the Near-Rational rule when they play for money. This may be partly due to additional learning subsequent to the first two practice rounds, and partly due to enhanced effort.

To what extent do subject's type assignments exhibit persistence over the two practice rounds and the money round? There are 27 possible sequences of type assignments over the three rounds. Figure 6 reports the number of subjects assigned to each type sequence. It also reports the expected number under the "Naïve" assumption that assignments are independent across

---

[18] The marginal likelihood values were -3056, -2816 and -2825 for models with 2, 3 and 4 types, respectively.

[19] The type ones in the practice data choose option "1" a bit too readily relative to the RE rule, and so have a mean 7% wealth loss relative to RE, as opposed to 2% for the Near-Rational type in money play. Still, their decision rule resembles the Near-Rational rule, subject to some additional fine-tuning of the reservation payoff differential.

rounds, given the marginal type proportions in each round. The cell counts depart substantially from what one would expect given random assignment. The $\chi^2(20)$ test for the hypothesis of independent assignments across rounds is 71.4, compared to the 1% critical value of 37.6.

A good deal of the variation in type across rounds that we do observe in Figure 6 is driven by improvement between practice and money play (due either to learning or enhanced effort). Note that only 3 subjects are assigned to the Confused rule in money play after having done better than that in both practice rounds (these are the RRC, FRC, RFC and FFC cells). But 21 subjects are assigned to the Near-Rational rule in money play after never having done that well in the two practice rounds (these are the FFR, CFR, FCR and CCR cells). Furthermore, among subjects whose performance varied between the two practice rounds, there are 34 cases where money play was as good as the subject's best practice round (these are the RFR, FRR, RCR, CRR, FCF and CFF cells), and only 17 cases where money play was as bad as the subject's worst practice round (these are the RCC, CRC, FCC, CFC, RFF and FRF cells).

Table 7 provides another way of looking at the data that helps to clarify the extent of persistence in type assignments between practice play and the money round. Clearly, practice round play is quite predictive of money play type assignments, but the agreement is far from perfect. An interesting question is the extent to which the variability in type assignments across the three rounds is attributable to type switching vs. errors in our probabilistic type assignments. To address this, we estimate a simple model of an econometrician trying to infer subject types.

A general model would have 20 free parameters: 2 type proportions in practice play, 2 type proportions in money play, a 3 by 3 matrix of type transition probabilities containing 4 free parameters (once the type proportions in practice and money play are given), and two 3 by three matrices of classification rates in practice and money play, each of which has 6 free parameters (since the classification probabilities in each row must sum to one). Thus 2+2+4+6+6 =20.

We impose three restrictions to obtain a simpler model. First, we assume that classification error is unbiased. This means the restriction $P(k^*) = P(R) P(k^*|R) + P(F)P(k^*|F) + P(C) P(k^*|C) = P(k)$ must hold for k=R, F, C, where $P(k^*)$ denotes the probability a subject is classified as type k, while $P(k)$ denotes the proportion of type k in the population. This reduces the number of free parameters in the classification rate matrices from 6 down to 4. It also means that the 4 free type proportion parameters are equal to the observed population type frequencies.

Second, we restrict the 3 diagonal elements in the classification matrix in the money round to differ from those in practice play by a common additive constant. This is motivated by the fact (apparent in Figure 5) that the degree of type separation in money plays seems to be greater, so we expect that the overall rate of correct classification in money play may be higher. By constraining the model in this way, we leave only two free parameters in the money round classification rate matrix. Third, we assume that type improvement between practice and money play is possible (due to either learning or increased effort), but that deterioration is not. This zeros out three elements in the type transition matrix, leaving it with just one free parameter.

Thus, we are left with just 4+2+1=7 parameters to be estimated. We estimate these by maximum likelihood using the 27 cell frequencies in Figure 6. The results are in Table 8. The $\chi^2(20)$ statistic for our "Simple" model is 10.25, compared to the 10% critical level of 28.41. Thus, the model reproduces the cell frequencies in Figure 6 very accurately. We were unable to significantly improve the likelihood by relaxing any of our restrictions. On the other hand, any further restrictions on the classification rate or type transition matrices are rejected.

Turning to the estimates, we obtain probabilities of correct assignment for types R, F and C in practice play of 51%, 53% and 78%, respectively, while in the money round these figures are 69%, 71% and 95%.[20] The improvement in the rate of correct assignment in money play is 18%.[21] Unconditionally, the probability of correct classification in the practice round is 64%, while that in the money round is 76%. There is also strong evidence of improvement between practice and money play. For instance, the estimates imply that 37% of the F subjects improve to R. These figures clarify the extent of type switching and classification error that is necessary to explain the variability in type assignments across rounds that we observe in Figure 6.

6. Conclusion

We have described a new Bayesian procedure for classification of subjects into decision rule types in choice experiments. We applied the procedure to experimental data from a sequential discrete choice setting in which optimal decision making would require subjects to

---

[20] A completely uninformative algorithm would assign subjects to each type with rates equal to the population type proportions (i.e., 26%, 30% and 44% for R, F and C in practice play, and 37%, 40% and 24% in money play).
[21] Our model constrains type to be invariant between the two practice rounds. Type changing between the two practice rounds would therefore be reflected in a lower rate of correct type assignment. For instance, Glenn Ellison points out to us that (using a slightly different model) he calculates rates of correct assignment of about 58% and 68% for the R and F types in practice if he assumes a 40% chance of changing type between the two rounds.

solve a difficult dynamic programming problem. The procedure produced a clear classification of the subjects into three behaviorally distinct types.

More than a third of the experimental subjects followed a rule very close to the optimal (expected wealth maximizing) rule. We labeled them "Near-Rational," since their play resulted, on average, in only about a 2% payoff loss relative to optimal play. We were surprised that so many subjects learned to play nearly optimally in a very difficult dynamic problem after about a half hour of practice (on average), particularly since the monetary rewards were fairly small.

The remaining subjects departed more clearly from optimal play. About 40% followed a sub-optimal rule that resulted in 12% payoff losses on average. We labeled these subjects "Fatalists," because their behavior implied too much reliance on the luck of the draw and a failure to appreciate the extent to which payoffs in the game were a controlled stochastic process. About a quarter of our subjects performed substantially less well, following a rule that earned about 19% less than optimal play on average. We labeled them "Confused" (see Andreoni (1995) or Houser and Kurzban (2002) on confusion in other experimental settings).

Experimental work that finds departures from 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, and because the practice round data shows that most subjects put substantial effort into the task. Furthermore, it was the types that performed worst in the game who tended to practice most.

In future work, we plan to examine how various experimental design features, such as (1) size of payoffs, (2) complexity of the problem, (3) amount of practice time allowed for learning about the game, and (4) amount of information given to the participants, effect the types of decision rules people use. Our (very) long-term goal is to try to provide some characterization of the types of situations in which people do and do not behave close to optimally, and to ascertain if certain sub-optimal behavioral patterns recur in many different contexts.

Our work raised two particular questions that we will investigate in future work. One is whether the "fatalistic" type behavior that we uncovered is common in other dynamic stochastic choice problems. A second question is whether the notion of an option value is generally much harder to understand than that of an investment value. All the subjects in our experiment understood, at least to some extent, the notion of an "investment value" of a choice. That is, all

subjects chose options that had low current payoffs, but that raised future expected payoffs, far more often than would myopic subjects. But not even the best performing subjects showed any understanding of the notion of an "option value." That is, no subject ever declined a high payoff alternative in order to defer the option of choosing it to a future round. In future work, we will examine whether the notion of an option value becomes more salient if the rewards to understanding the concept are increased.

Our Bayesian procedure for decision rule classification could potentially be applied in many settings besides the dynamic discrete choice problem we have considered. For instance, it could be used to model decision rules in strategic games, in which case the polynomial approximation to the continuation value would typically include state variables characterizing the play of other subjects. And it could be applied to field as well as experimental data. In each case, the key identifying assumption is that the investigator must specify *a priori* the set of state variables that he/she will entertain as potential arguments in subjects' decision rules.

Our finding that behavioral heterogeneity is important in experimental data is consistent with prior results. For instance, El-Gamal and Grether (1995), in their experiment on Bayesian learning, found evidence that subjects fell into three types: Bayesians, conservative Bayesians, and those who used the representativeness heuristic (see also McCabe et al (2001)). Given the accumulating evidence that decision rule heterogeneity is important in laboratory environments, we believe it is reasonable to suspect that such heterogeneity is also important in field data. Developing empirical strategies to model such heterogeneity is an important research agenda.

Interdisciplinary Center for Economic Science and Dept. of Economics, George Mason University, 4400 University Blvd., MSN 1B2, Fairfax, VA 22030, U.S.A.; dhouser@gmu.edu; www.ices-gmu.org,
and
Dept. of Economics, Yale University, 37 Hillhouse Ave., New Haven, CT 06511, U.S.A.; michael.keane@yale.edu,
and
Interdisciplinary Center for Economic Science and Dept. of Economics, George Mason University, 4400 University Blvd., MSN 1B2, Fairfax, VA 22030, U.S.A.

References

Akerlof, G. and J. Yellen. 1985. The macroeconomic consequences of near-rational rule-of-thumb behavior. *Quarterly Journal of Economics*, 823-38.

Andreoni, J. 1995. Cooperation in public goods experiments: kindness or confusion? *American Economic Review*, 891-904.

Bellman, R. 1957. Dynamic Programming. Princeton University Press. Princeton, NJ.

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

Camerer, Colin and Teck-Hua Ho. 1999. Experience-weighted attraction learning in normal form games. *Econometrica* 67, 827-74.

Chib, Siddhartha. 2001. Markov chain monte carlo methods: Computation and Inference. In J. Heckman and E. Leamer (eds), *Handbook of Econometrics*, Vol. 5. North Holland, p. 3569-3649.

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.

Duffy, John and Jim Engle-Warnick. 2001. Using symbolic regression to infer strategies from experimental data. In S.H. Chen (ed.), Evolutionary Computation in Economics and Finance. Springer-Verlag, NY.

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

El-Gamal, Mahmoud A. and Thomas R. Palfrey. 1995. Vertigo: Comparing structural models of imperfect behavior in experimental games. Games and Economic Behavior 8, 322-48.

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

Engle-Warnick, Jim. 2003. Inferring strategies from observed actions: A nonparametric, binary tree classification approach. *Journal of Economic Dynamics and Control*, forthcoming.

Engle-Warnick, Jim and Bradley J. Ruffle. 2002. The strategies behind their actions: A method to infer repeated-game strategies and an application to buyer behavior. Working paper, University of Oxford.

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.

Gelman, A. 1996. Inference and monitoring convergence. In W.R. Gilks, S. Richardson and D.J. Spiegelhalter (eds.) Markov Chain Monte Carlo in Practice. London: Chapman & Hall.

Geweke, J. 1997. Posterior simulators in econometrics. *Advances in economics and econometrics: Theory and applications: Seventh World Congress, 3,* 128-65.

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. 466-493.

Geweke, J. and M. Keane. 1999a. 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.

Geweke, J. and M. Keane. 1999b. Mixture of normals probit models. In Hsiao, Lahiri, Lee and Pesaran (eds.), Analysis of Panels and Limited Dependent Variable Models, Cambridge University Press, 49-78.

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, p. 3463-3568.

Goeree, Jacob K. and Charles A. Holt. 1999. Stochastic game theory: For playing games, not just doing theory. *Proceedings of the National Academy of Sciences of the United States of America*, 96(19), 10564-10567.

Goeree, Jacob K., Charles A. Holt and Thomas R. Palfrey. 2000a. Quantal response equilibrium and overbidding in private-value auctions. Working paper, California Institute of Technology.

Goeree, Jacob K., Charles A. Holt and Thomas R. Palfrey. 2000b. Risk averse behavior in asymmetric matching pennies games. Working paper, California Institute of Technology.

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.

Heckman, J. and B. Singer. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. Econometrica 52, 271-320.

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

Houser, D. 2002. Bayesian analysis of a dynamic stochastic model of labor supply and saving. *Journal of Econometrics*, forthcoming.

Houser, D. and R. Kurzban. 2002. Revisiting kindness and confusion in public goods experiments. *American Economic Review*, forthcoming.

Houser, D. and J. Winter. 2002. How do behavioral assumptions affect structural inference? Evidence from a laboratory experiment. *Journal of Business and Economic Statistics*, forthcoming.

Keane, M. and K. Wolpin. 1994. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. *Review of Economics and Statistics*, *76*, 648-72.

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.

Lewis, Steven M. and Raftery, Adrian E. 1997. Estimating Bayes Factors via Posterior Simulation with the Laplace-Metroplis Estimator. Journal of the American Statistical Association, 92. 648-655.

McCabe, K., D. Houser, L. Ryan, T. Trouard and V. Smith. 2001. A functional imaging study of cooperation in two person reciprocal exchange. *Proceedings of the National Academy of Sciences of the United States of America*, 11832-11835.

McKelvey, Richard D. and Thomas R. Palfrey. 1992. An experimental study of the centipede game. Econometrica 4, 803-36.

McKelvey, Richard D. and Thomas R. Palfrey. 1995. Quantal response equilibria for normal form games. Games and Economic Behavior 10, 6-38.

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

Rust, J. 1994. Structural estimation of Markov decision processes. In R.F. Engle and D.L. McFadden (eds.), Handbook of Econometrics. Vol 4. North Holland, Amsterdam.

Schachat, J. and M. Walker. 1997. Unobserved heterogeneity and equilibrium: An experimental study of Bayesian and adaptive learning in normal form games. Working paper, University of Arizona.

Stahl, Dale O. and Paul W. Wilson. 1995. On players' models of other players: Theory and experimental evidence. Games and Economic Behavior 10, 218-54.

Appendix A: Numerical Procedure for Calculating Marginal Likelihood Values

We construct the marginal likelihood for each of our candidate models using the procedure developed by Lewis and Raftery (1997). Their algorithm combines posterior simulation with the Laplace-Metropolis estimator. Alternative numerical procedures for calculating the marginal likelihood are discussed in Gelfand and Dey (1994), Geweke (1997), Geweke and Keane (2001), and Chib (2001), among others.

The marginal likelihood value for a model, which we denote by $g(x)$, is the integral of the model's likelihood function with respect to the model's prior. For the model we present in Section IV, the marginal likelihood is:

$$g(\{\{d_{nt}\}_{t=1,T}\}_{n=1,N} \,|M_{KP}) =$$
$$\int L\ [\{\{d_{nt}\}_{t=1,T}\}_{n=1,N} \,|(\boldsymbol{p}_k, \boldsymbol{s}_k^{-2}, \boldsymbol{q}_k)_{k\in K}]\, p((\boldsymbol{p}_k, \boldsymbol{s}_k^{-2}, \boldsymbol{q}_k)_{k\in K} \,|M_{KP})\, d((\boldsymbol{p}_k, \boldsymbol{s}_k^{-2}, \boldsymbol{q}_k)_{k\in K})$$

where M_{KP} indicates a model with K types of subjects, in which the order of the polynomial F is P, and in which the prior is $p(x/M_{KP})$. Dropping the notational dependence on the model, and letting $\boldsymbol{x}$ denote the parameter vector and D the vector of observed decisions, this can be written

$$g(D) = \int L\ [D\,|\,\boldsymbol{x}]\, p(\boldsymbol{x})d\boldsymbol{x}.$$

The Laplace method generates the following approximation for the marginal likelihood:

$$g(D) \approx (2\boldsymbol{p})^{l/2}\, |H^*|^{1/2}\, p(\boldsymbol{x}^*) L\ (D\,|\boldsymbol{x}^*)$$

where $\boldsymbol{x}^*$ is the value of $\boldsymbol{x}$ at which $h(\boldsymbol{x}) \equiv \log\{p(\boldsymbol{x}) L\ (D\,|\,\boldsymbol{x})\}$ attains its maximum (i.e., the posterior mode), H^* is minus the inverse Hessian of h evaluated at $\boldsymbol{x}^*$, and l is the dimension of the parameter space. Taking logarithms, this can be rewritten as

$$(A1) \quad \log\{g(D)\} \approx \frac{l}{2}\log\{2\boldsymbol{p}\} + \frac{1}{2}\log\{|H^*|\} + \log\{p(\boldsymbol{x}^*)\} + \log\{L\ (D\,|\boldsymbol{x}^*)\}.$$

Lewis and Raftery (1997) call this the Laplace-Metropolis estimator. It is attractive because the quantities $\boldsymbol{x}^*$ and H^* can both be easily derived from Gibbs sampler output. To determine $\boldsymbol{x}^*$ we evaluate $h(\boldsymbol{x})$ at each draw from the posterior, and chose that parameter vector for which $h(\boldsymbol{x})$ was the largest. The quantity H^* is asymptotically equivalent to the posterior variance matrix, so one may use the sample covariance matrix of the simulation output as an estimate of its value.

The marginal likelihood tends to favor more parsimonious models for the following reason: If we increase K and/or P then the prior is specified over more parameters. Thus, ceteris paribus, the prior mass in the vicinity of any particular parameter vector (such as the posterior mode, $\boldsymbol{x}^*$) will fall. This tends to reduce the value of (A1), inducing an implicit penalty on added parameters.

Appendix B: 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

Bonus:
    0 if you have chosen "1" 0, 1, 2, 3, 4, or 5 previous times
  7500 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.

Table 1: Model Selection Based on Marginal Likelihood Values

Marginal Likelihoods under Baseline Prior

	Order of Future Component		
Number of Types	P=3	P=4	P=5
1	-1712	-1721	-1736
2	-1451	-1364	-1442
3	**-1224**	-1277	-1369
4	-1312	-1342	-1449

Prior Sensitivity
A. Prior Std. Dev. Of π^* Reduced 50%

Types	P=3	P=4	P=5
1	-1708	-1704	-1728
2	-1445	-1373	-1420
3	**-1213**	-1271	-1347
4	-1275	-1316	-1370

B. Prior Std. Dev. Of π^* Increased 50%

Types	P=3	P=4	P=5
1	-1731	-1707	-1741
2	-1459	-1366	-1502
3	**-1231**	-1282	-1439
4	-1297	-1286	-1437

C. Prior on σ^2 Scaled Up

Types	P=3	P=4	P=5
1	-1706	-1716	-1731
2	-1442	-1363	-1432
3	**-1209**	-1263	-1355
4	-1285	-1319	-1429

Note: The preferred model under each prior is highlighted in bold.

Table 2: Prior and Posterior Means and Standard Deviations of Future Component Parameters

	Prior Distribution		Type 1: N=51 "Near-Rational"		Type 2: N=55 "Fatalist"		Type 3: N=33 "Confused"	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
p_0^*: Intercept	0.0	2×10^4	4259.21	108.60	3478.47	209.06	3711.68	476.31
p_1^*: X1	0.0	10^3	-4.42	22.51	411.33	38.00	-429.17	82.87
p_2^*: X2	0.0	10^3	92.57	42.90	276.70	49.67	-447.22	120.54
p_3^*: X1^2	0.0	10^2	-29.29	1.49	-32.47	2.85	-16.97	5.82
p_4^*: X2^2	0.0	10^2	-73.35	5.39	-3.57	2.38	-2.89	6.98
p_5^*: X1*X2	0.0	$10^{5/2}$	-86.00	2.47	-1.17	4.92	-103.36	11.61
s_h: Optimization error	Not Defined	Not Defined	208.91	58.43	863.81	29.50	2270.96	78.73
q_k: population type probability	0.33	0.18	0.36	0.06	0.40	0.06	0.25	0.05

Note: "X1" Denotes experience in alternative "1," and "X2" denotes experience in alternative "2." The round is not included as a state variable in the polynomial since it is perfectly collinear with X1 and X2. Lagged choice is not included because it drops out of the differenced future component (and is subsumed in the constant).

Table 3: Descriptive Statistics for the Play of Each Type

	Type 1: "Near-Rational"	Type 2: "Fatalist"	Type 3: "Confused"
Number of subjects (percent)	51 (37%)	55 (40%)	33 (24%)
Mean earnings (points)	87983	80811	75966
Mean earnings under RE	90047	91546	93316
Percent loss relative to RE	2.3%	11.7%	18.6%
SD of earnings	9620	13727	14189
Number who earn at least as much as RE subjects	22	3	2
Number who earn exactly as much as RE subjects	11	0	0
Mean number of times alternative "1" is chosen	11	9.2	9.6
Number who complete all bonus rounds	48 (94.1%)	38 (69.1%)	12 (36.3%)

Note: "Mean earnings under RE" reports the mean earnings for hypothetical subjects who follow the optimal (expected wealth maximizing) decision rule, given that they face the same draws for the stochastic component of payoffs as did the actual subjects.

Table 4: Marginal Likelihoods of Various Models Using Simulated Data

Number of Distributions in Mixture	3-Mixture Simulation	5-Mixture Simulation
1	-1388	-2651
2	-826	-2000
3	**-604**	-1359
4	-802	-1113
5	-NA-	**-1039**
6	-NA-	-1052

Table 5: Mean Value of Draw by Alternative and Type

Posterior Type Classification	Draw for Alternative 1	Draw for Alternative 2	Difference (1-2)
Near-Rational	108 (0.29)	33 (0.75)	75 (0.59)
Fatalist	31 (0.76)	99 (0.34)	-69 (0.64)
Confused	269 (0.05)	-58 (0.64)	327 (0.08)

Note: p-values for two-sided t-test that mean is zero are in parentheses.

Table 6: Practice Rounds by Day and Type

Type	N	Day 1 Practice Rounds				Day 2 Practice Rounds				Total Practice Rounds			
		Mean	Median	Min	Max	Mean	Median	Min	Max	Mean	Median	Min	Max
Near-Rational	49	43	29	0	166	15	14	2	44	58	48	6	180
Fatalist	50	52	42	0	195	17	12	2	95	69	59	4	201
Confused	25	54	46	1	131	21	18	2	50	75	69	3	154
All	124	49	38	0	195	17	14	2	95	66	58	3	201

Note: The file containing practice round data for 15 of our 139 subjects was inadvertently deleted. This table reports data for the remaining 124 subjects.

Table 7: Probability of Type in Money Round Conditional on Practice Play

Practice	N	Money Game Near Rational	Fatalist	Confused
RR	14	42.9%	42.9%	14.3%
FR, RF	24	70.8%	29.2%	0.0%
FF	19	47.4%	47.4%	5.3%
RC, CR	22	31.8%	45.5%	22.7%
FC, CF	21	28.6%	47.6%	23.8%
CC	39	15.4%	33.3%	51.3%

Table 8: A Simple Model of Type Assignments

Classification Probabilities - Practice Rounds

True Type	Classified Type R	F	C
R	0.51	0.39	0.10
F	0.23	0.53	0.24
C	0.14	0.09	0.78

Transition Rates - Practice to Money

Practice	Money R	F	C
R	1.00	0.00	0.00
F	0.37	0.63	0.00
C	0.00	0.45	0.55

Classification Probabilities - Money Round

True Type	Classified Type R	F	C
R	0.69	0.28	0.03
F	0.29	0.71	0.00
C	0.01	0.04	0.95

Model Fit Statistics

Model	Chi-Square	Log-likelihood
Simple	10.3	-421.7
Naïve	71.4	-448.4
Unconstrained	0.0	-416.1

Note: The "simple" model has 7 free parameters in the classification rate and transition rate matrices. The "Naive" model assumes that type assignments are independent across rounds. The "Unconstrained" model sets the probability of each cell in Figure 6 equal to its empirical frequency. The 10% critical value of the $\chi^2(20)$ statistic is 28.41.

Figure 1: Fraction of Alternative One Choices by Round for Actual Subjects, and their Rational Expectations and Simulated Counterparts
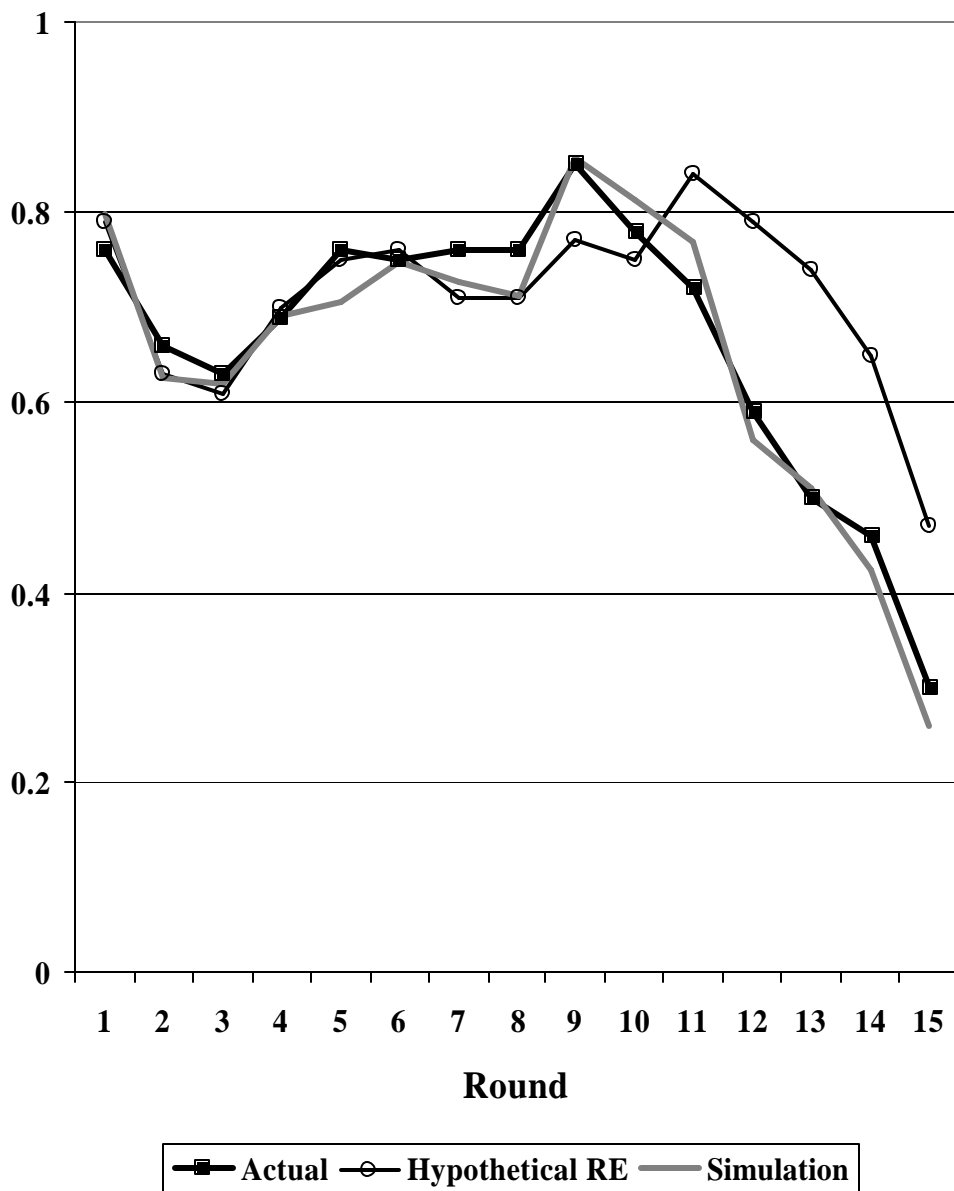
## Figure 2: Fraction of Actual and Simulated Alternative One Choices by Round for Each Type

(a) Near-Rational

Round

(b) Fatalist

Round

(c) Confused

Round

— Actual —●— Simulation
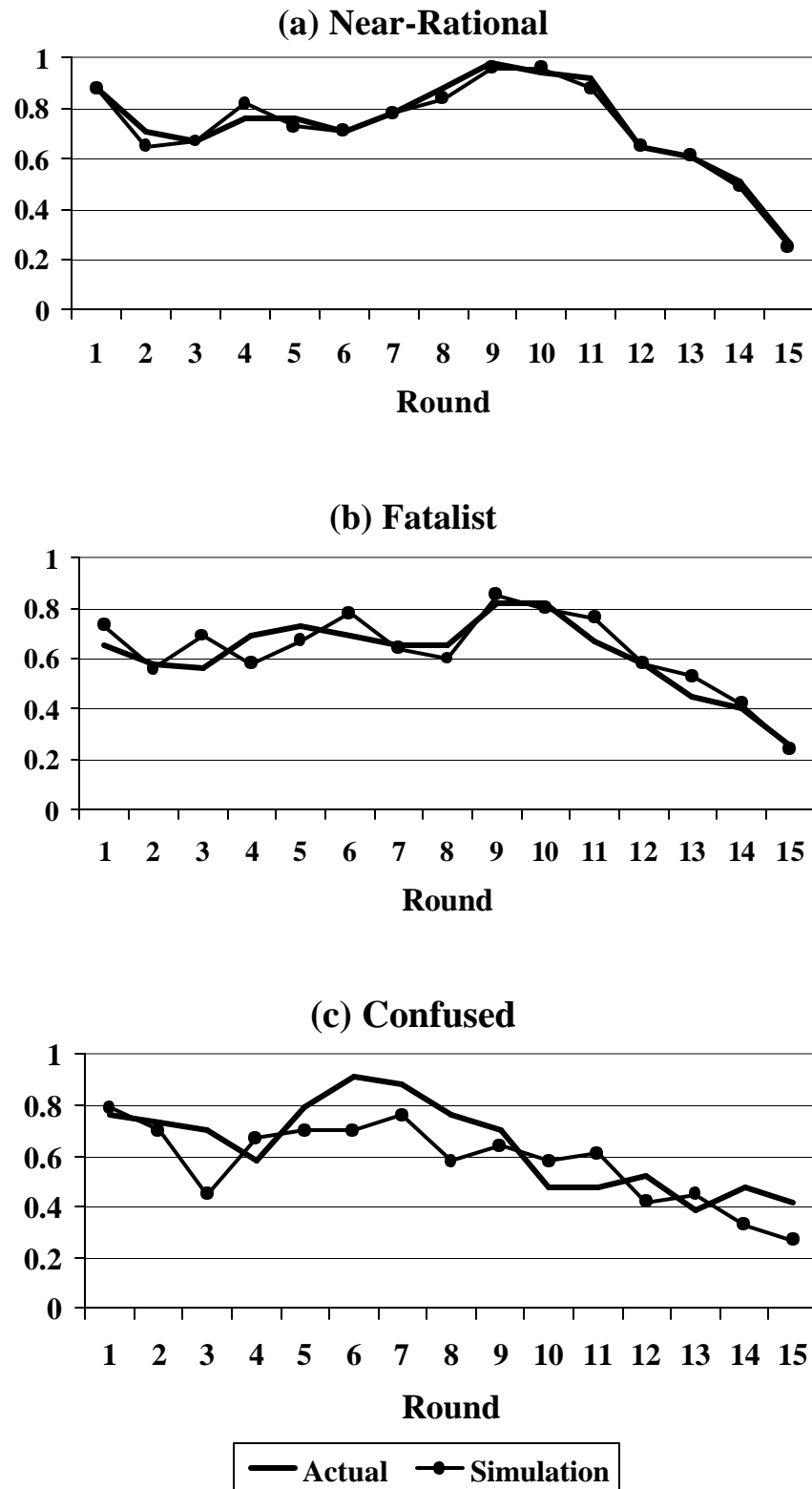
**Figure 3: Fraction of Alternative One Choices Under Simulations of Various Decision Rules**
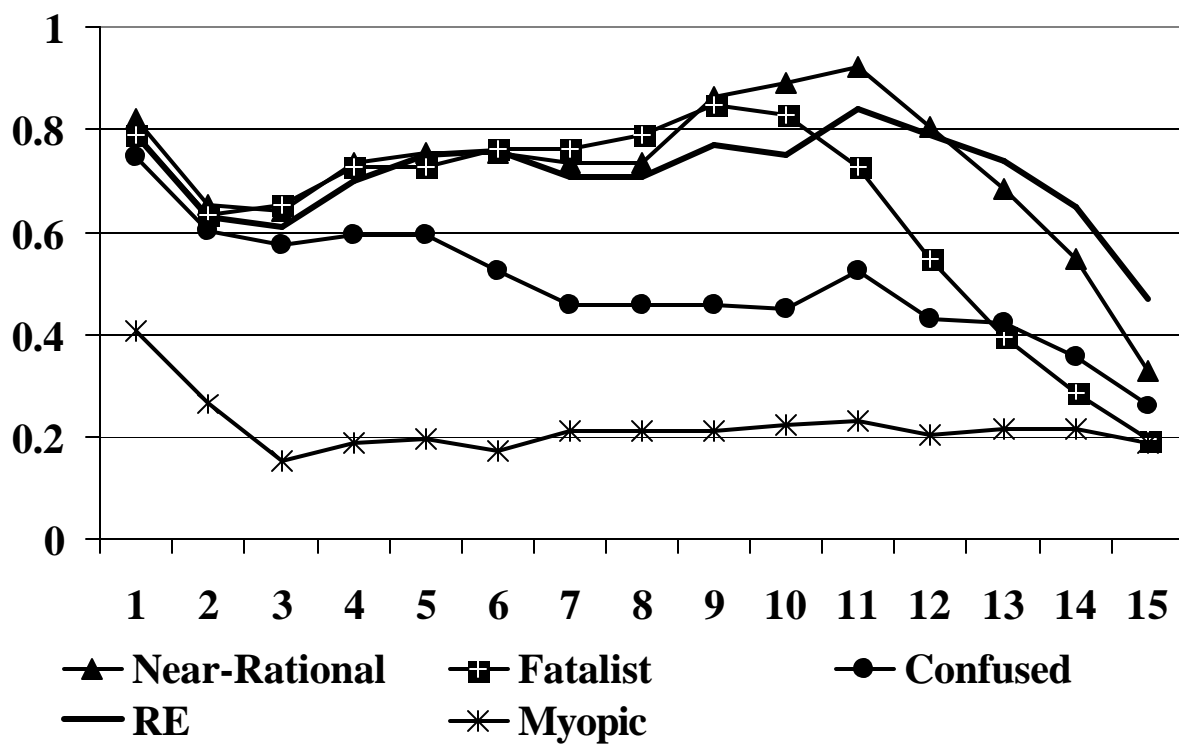
Figure 4A: Comparison of Fitted and Rational Expectations Future Component at Early State Vectors for Each Type

Near-Rational

Fatalist

Confused
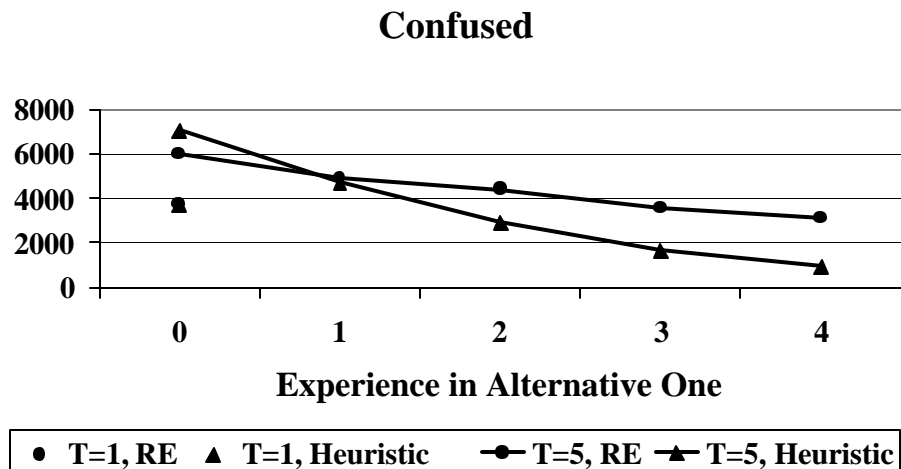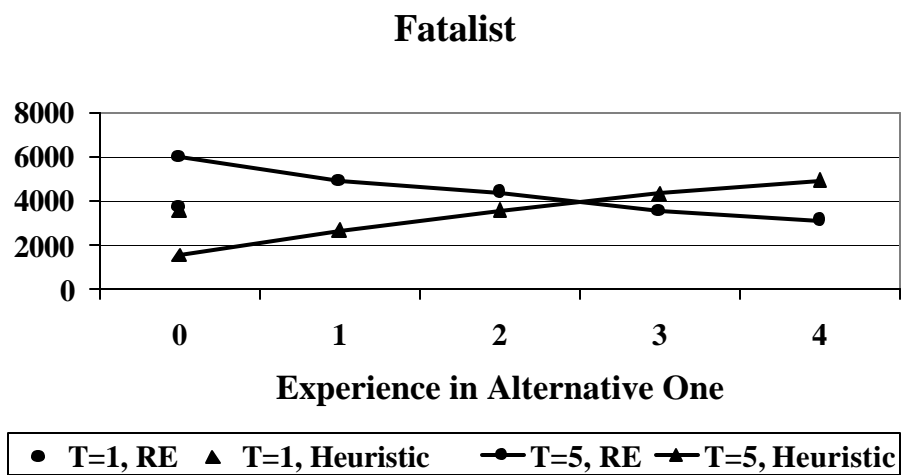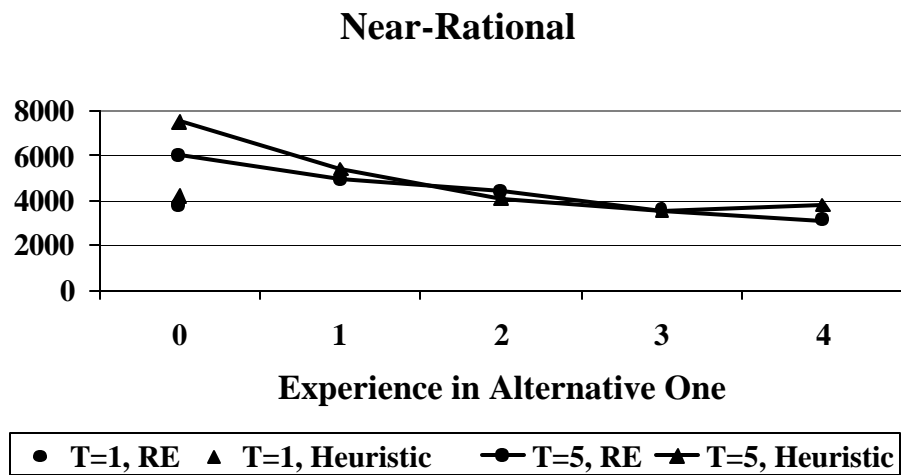
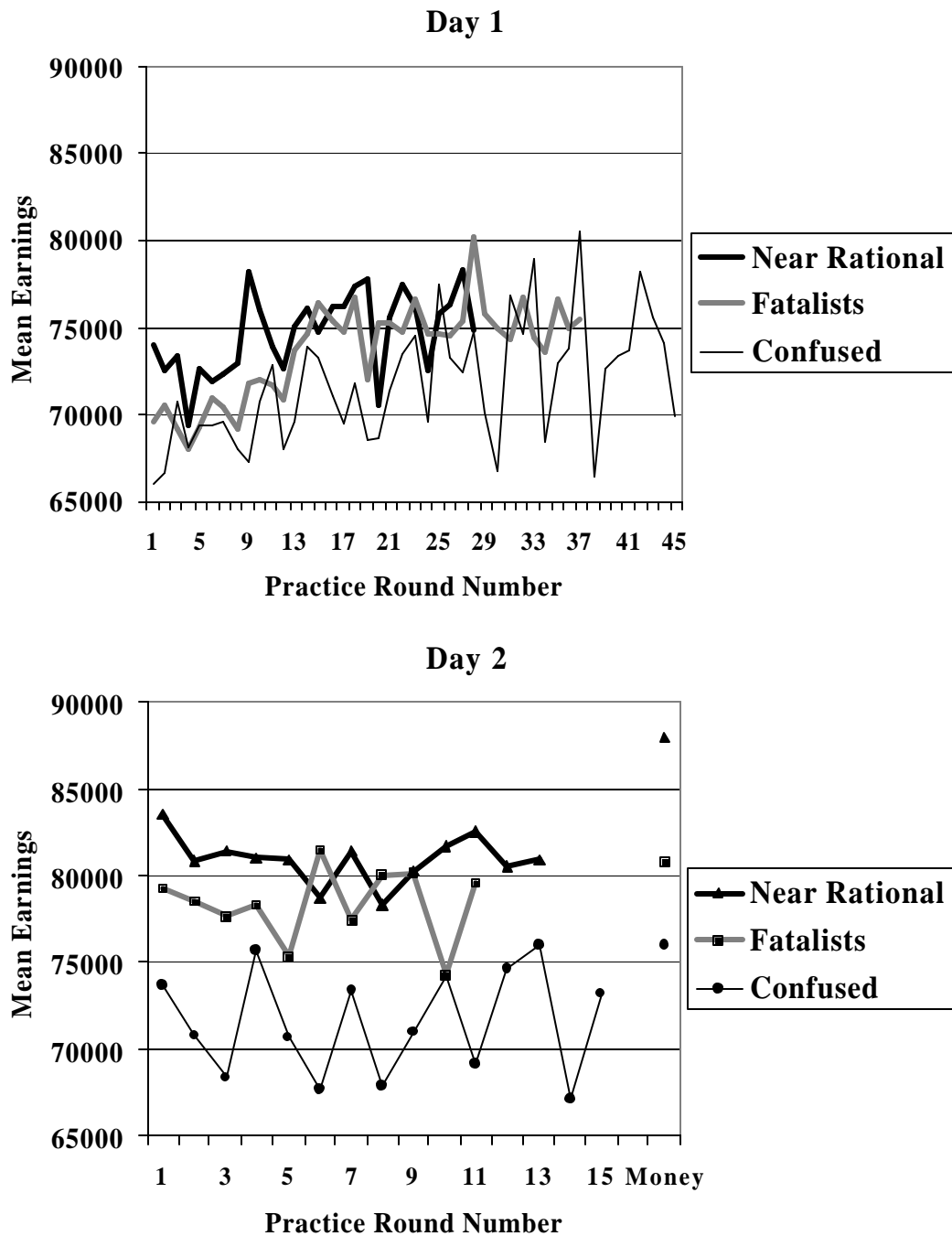# Figure 4B: Comparison of Fitted and Rational Expectations Future Component at Late State Vectors for Each Type

Near-Rational

Fatalist
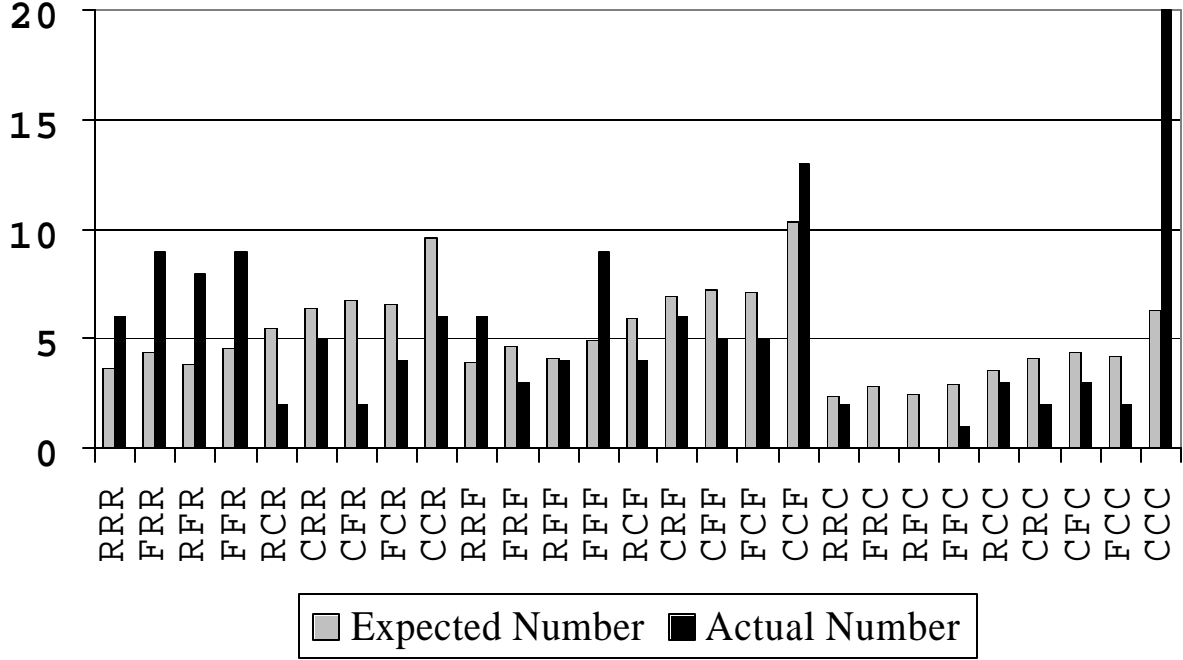

Confused
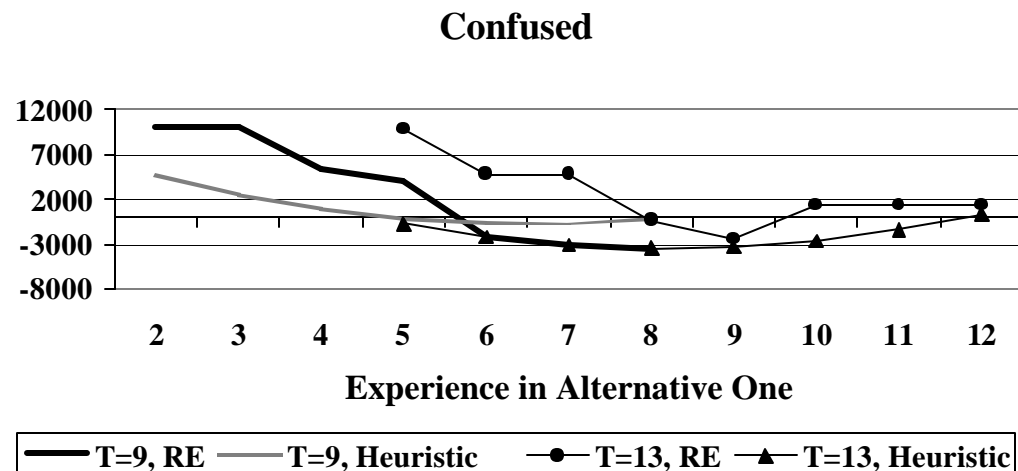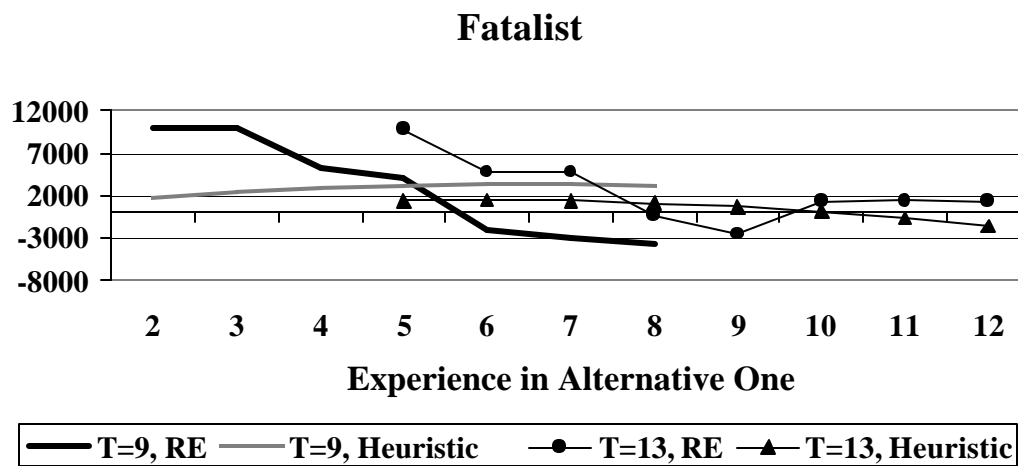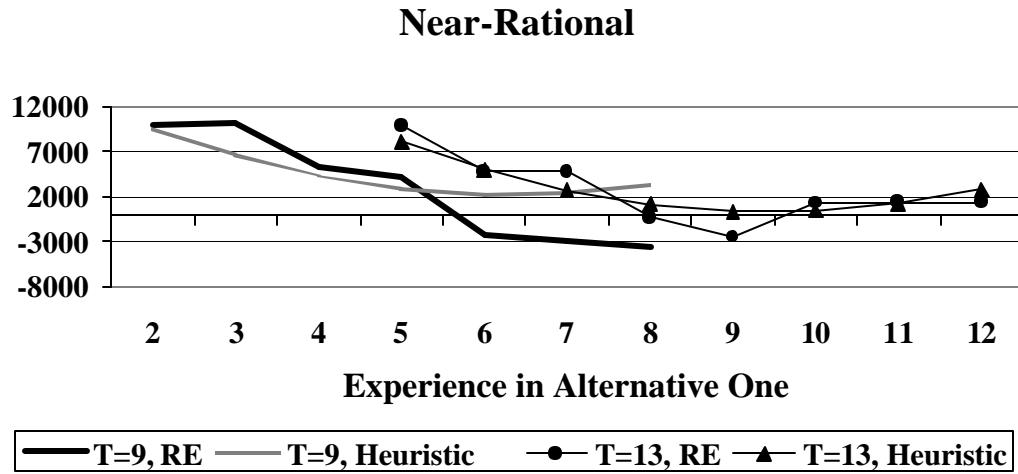
Figure 5: Mean Earnings During Practice by Round and Type

Day 1

Day 2

Note: Means that include a small number of subjects are highly volatile over rounds. Hence, for each type we include only rounds where over half of our subjects practice. In each panel, the number of subjects underlying each mean is greater than 25, 25 and 13 for near rational , fatalist and confused subjects, respectively.

Figure 6: Expected and Actual Cell Frequencies

Note: The labels on the horizontal axis indicate sequences of rules followed in the first and second practice round and the money round, in that order. "R" denotes rule 1 ("Near-Rational"), "F" denotes rule 2 ("Fatalist") and "C" denotes rule 3 ("Confused"). For example, the cell labeled "CCR" gives the number of subjects assigned to the Confused rule in the two practice rounds and the Near-Rational rule in the money round. The expected number of subjects in each cell under independence of type assignments across rounds is calculated using the marginal type probabilities for each round. Specifically, if $p(X,j)$ represents the fraction of subjects classified as type X in round j, then the expected proportion of subjects in cell XYZ is the product $p(X,1)p(Y,2)p(Z,3)$.