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

By

Daniel Houser
George Mason University
and University of Arizona

Michael Keane
Yale University

Kevin McCabe
George Mason University

October 6, 2002

* This research was supported by grants from the Russell Sage Foundation, the CV-STARR center at NYU, and the Office of the Vice President for Research at the University of Arizona. Houser's work was also supported by a fellowship from the International Foundation for Research in Experimental Economics.

I. 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.[1] Recognizing the importance of this issue, there have been a large number of experimental studies that analyze the behavior of people confronted with complex decision problems in laboratory settings. Most of this literature adopts an "either/or" approach, asking whether subjects make optimal decisions or not, or asking what fraction of subjects behave optimally.[2]

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 a very 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. Since our approach is (largely) agnostic regarding the number and form of the rules that may exist in the population, it takes the exploratory nature of the analysis a step further than do earlier approaches to decision rule classification.

---

[1] 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)).

[2] For example, there is an extensive experimental literature that examines stopping behavior in search models (e.g., Braunstein and Schotter (1982), Cox and Oaxaca (1989, 1992), Harrison and Morgan (1990) and Hey (1987)). Although the design specifics differ, the basic idea is to compare observed and optimal search durations in order to

In contrast to our procedure, the El-Gamal and Grether (1995) approach requires the investigator to specify *a priori* both the maximum number of decision rules that may be present in the population, and the exact form of each rule. They present a statistical procedure that chooses a "best" subset of rules from the superset of all candidate rules.[3] In the particular 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, and limits the exploratory nature of the analysis. Our approach allows the data to determine both the number of rules that are present and their form.[4]

Aside from our work and that of El-Gamal and Grether (1995), 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. Again, our contribution is to develop an approach that is much less restrictive regarding the possible nature of subject heterogeneity.[5]

An attractive feature of our type classification algorithm is that it generates the posterior probability that each subject is following each decision rule, conditional on his/her observed behavior. Thus, one can classify individual subjects into types *a posteriori*. This allows one to study the characteristics of people assigned to each type. We provide Monte-Carlo evidence that our procedure performs well at assessing the number of types present in a population, and that the individual type assignments are reasonably accurate.

---

determine whether subjects make "optimal" stopping decisions. A typical finding is that only a proper subset of subjects behaves optimally, and little effort is made to describe the rules followed by subjects who play otherwise.

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

[4] Of course, the added flexibility our procedure comes at a cost. This is 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*.

[5] 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 more flexible in principle than the approaches detailed in the papers above. But this approach 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.

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 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 new type classification procedure to study the actual behavior of a sample of subjects who we confront with a particular complex dynamic optimization problem. We consider a sequential decision problem, in which subjects chose between two discrete alternatives in each of 15 time periods. In each period, a computer generates stochastic payoffs 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 (i.e., the payoffs are a controlled stochastic process), 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 very difficult to solve optimally. Analytic solution is not possible, and we must solve for the optimal decision rule numerically. In our experimental design, subjects are allowed to practice playing the game as many times as they like before they play for money.

Our prior was that few if any participants could determine the optimal decision rule in this problem, even given ample time for practice. Our interest was in characterizing the rules-of-thumb that subjects would adopt when confronted with a problem that was too difficult to solve optimally. We expected to see a great deal of heterogeneity in the rules-of-thumb that people

[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.

adopted, and we were concerned that the classification procedure might perform poorly when confronted with a large number of potential types.

We ran the experiment on 139 subjects, and we find the results to be surprising. Our classification procedure produces a very clear assignment of the population into only three distinct types. Statistical tests overwhelmingly reject the hypothesis that there are more, or that there are fewer. Visual inspection of the choice sequences of the subjects, as well as examination of exit interviews in which subjects were asked about their strategies, provide external validation that there clearly are three distinct types. And the subjects' posterior type probabilities usually assign a high probability to just one type. That is, conditional on his/her behavior, there is rarely much ambiguity about which of the three types a subject is.

The nature of the three types is also interesting. 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%. Another 40% of the subjects adopted a rule-of-thumb that is fairly simple to characterize. In fact, we will call these subjects "fatalists" because they play as if they don't appreciate the extent to which payoffs are a *controlled* stochastic process. The payoff losses for these subjects were about 12% on average. Finally, about a quarter of the subjects followed a more clearly sub-optimal decision rule, indicating they had a more severe misunderstanding of the nature of the game. Payoff losses of these "confused" subjects averaged about 19%. To give a baseline, completely myopic play, in which a subject ignores the dynamic aspect of the game and simply chooses the options with the highest current payoff in each period, would lead to payoff losses of 25% on average.

A criticism of such type assignments that we have often heard is that, while some subjects may have played sub-optimally within the context of the game, their behavior may have been optimal in some broader context. For example, perhaps the subjects we label as "confused" were just busier than other participants, and their optimal decision rule was to finish the game as soon as possible. But our data on practice rounds contradicts this view. Recall that our algorithm allows us to study the characteristics of people assigned to each type, and so, for instance, we can compare the different types in terms of the intensity with which they practiced the game. Surprisingly, we find that the confused types practiced more on average than any other type, and the near-rational types practiced the least. Thus, those subjects who played the game best also tended to learn it the quickest. On average our subjects played 66 practice rounds before playing the game for money, so it seems clear that the large majority of subjects put substantial effort into the experimental task.

In summary, we find that our type classification procedure performed very well at identifying types in the population of experimental subjects. It grouped subjects into types with a high degree of confidence, and the nature of each type was simple to characterize. We were surprised to find that, in a very difficult dynamic decision problem, more than a third of people were able to closely approximate the optimal decision rule.

The success of many participants in our experiment was particularly surprising given that the monetary rewards for optimal performance were small. A person who used the optimal decision rule could only expect to earn a few dollars more than one who behaved completely myopically.[7] In future work, we plan to examine how various experimental design features, such as (1) size of payoffs, (2) complexity of the decision problem, (3) amount of time for practice, and (3) amount of information given to the participants, effect the types of decision rules people use. Our (very) long-term goal is to be able to provide some characterization of the types of situations in which people do and do not behave close to optimally, and to characterize the types of decision rules people use in the latter cases (e.g., do certain sub-optimal behavioral patterns like "fatalism" recur in many different contexts?).

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

II. 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. Each decision rule is modeled flexibly, with the form of each rule inferred from the experimental data. The number of decision rules in the population is

[7] Experimental work that finds rejections of rational or optimal behavior is often criticized on the grounds that subjects had little incentive to behave optimally. We don't find this a compelling criticism of our findings here, since the interesting outcome was that so many people indeed behaved close to optimally, and because the practice round data shows that the large majority of subjects put substantial effort into the task.

determined using Bayesian decision theory, 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 the approach we describe 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. Appling Bellman's (1957) principle, the value to subject n of choosing alternative j from the discrete set $\{1, \ldots, J\}$ in round t of a T period game can be written:

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

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 $H(\cdot)$ is the (possibly stochastic) Markovian law of motion for the state variables.

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

Chose alternative j *iff* $V_{njt}(I_{nt}) > V_{nkt}(I_{nt})$ *for all* $k \neq j$.

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

[8] For instance, Houser (2002) applies this approach to a model with mixed discrete/continuous choice variables.
[9] If payoffs are received over time, the future component would be pre-multiplied by a discount factor, but 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.

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

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 errors. Then, we can write the future component for type k as:

$$(2) \qquad EV(I_{n,t+1} \mid I_{nt}, j) = F(I_{nt}, j \mid \pi_k) + \varsigma_{njt} \qquad\qquad k=1, \ldots, K.$$

Here $F(\cdot)$ represents the future component polynomial. It is characterized by a finite vector of parameters π_k. Note that the type k specific future component parameters π_k capture differences in the way subjects assign values to alternatives.

The random variable ς_{njt} accounts for idiosyncratic errors made when attempting to implement decision rule k. We allow the distribution of the idiosyncratic errors to vary with the decision rule, so that stochastic optimization error may be more important for some types than others.[11] Let σ_k denote the standard deviation of the optimization error for type k.

The assumption that optimization 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 (as opposed to more traditional work that simply checked whether or not subjects behavior was optimal according to particular normative models). Goeree and Holt (1999) provide an overview of this line of research.

---

[10] Sub-optimality in this sense could occur for a number of reasons. Examples include: 1) subjects do not have rational expectations, so that E is not the mathematical expectation operator, 2) subjects are unable to do the calculations necessary to solve the DP problem and compute the optimal decision rule so they use some simple approximation, 3) subjects do not understand the structure of $H(\cdot)$, 4) subjects could in principle solve the problem, but instead they solve a simpler problem to save on thinking costs, 5) subjects fundamentally misunderstand the problem, and 6) subjects are not expected wealth maximizers (e.g., they exhibit altruism), or, more generally, the problem they are solving is somehow different from the one we have posed to them.

[11] One must allow for optimization errors 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. An appealing feature of our alternative approach (i.e., introducing optimization error directly into the value function equation) is that "wrong" choices will be much more likely if the "true" values of alternatives are "close." This seems intuitive.

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

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

Denoting the deterministic part of the valuation function by $\bar{V}_{njt}(I_{nt} \mid k) \equiv w_{njt} + F(I_{nt}, j \mid \pi_k)$, we have that the probability that option j is chosen is increasing in $\bar{V}_{njt}$. The assumption that optimization error takes the form of noise appended to the $\bar{V}_{njt}$ function is attractive, because it implies that the probability of an error in implementing the decision rule is greater in situations where the several best alternatives have similar $\bar{V}_{njt}$ values, whereas the probability of an error is small when one alternative is clearly dominant in terms of its $\bar{V}_{njt}$.

It is important to note that specification (3) is extremely flexible and can nest or approximately nest (i.e., via the Weierstrass approximation theorem) many special cases of interest. For example, if the function F is sufficiently flexible, and the π_k are chosen so that $F(I_{nt}, j \mid \pi_k) \approx EV(I_{n,t+1} \mid I_{nt}, j)$, and the variance of the optimization error ς_{njt} is set close to zero, we can obtain a good approximation to the "optimal" value function given by (1). In fact, many authors have found that value functions in problems of interest to economists can be approximated extremely accurately using low order polynomials (see, e.g., Krusell and Smith (1995), Geweke and Keane (1999a, 2001), Keane and Wolpin (1994)). This observation is a key motivation for our approach.

Some other leading cases are also worth noting: Equation (3) nests myopic behavior if the π_k are set equal to zero and the variance of the optimization error is set equal to zero. And equation (3) generates purely random behavior, in which each option is chosen with equal probability, if the π_k are set equal to zero and σ_k is sent to infinity.

It is important to note what elements of the specification (3) are and are not restrictive. First, note that this specification 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}. Second, our specification may appear to impose additive separability, but this can be relaxed, since current and past payoffs and choices, and interactions between them and other state variables, may be included as elements of I_{nt} or I^{+}_{nt}. Finally, observe that non-stationarity arises as long as the index for round (or time) t is an element of I_{nt}. We comment on the restrictive aspects of our approach below.

Now we consider the problem of statistical inference in this general framework. Let $d_{nt} \in \{1, \ldots, J\}$ denote the choice made by subject n in round t. The data observed by the investigator will consist of the set of choices and payoffs for each of the N subjects:

$$\{\{\{ d_{nt} , \{w_{njt}\}_{j=1, \ldots, J}\}_{t=1, \ldots, T}\}_{n=1, \ldots, N}\}$$

The goal of the empirical analysis is 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 π_k, σ_k for each type $k=1,\ldots, K$, and (4) the population proportions of each type, which we denote by θ_k. Furthermore, we also want to construct 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 $\{\pi_k, \sigma_k, \theta_k\}$ for $k=1,\ldots, K$, *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 ς_{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 estimate a range of models with different K and P, and use the marginal likelihood to choose among competing models. Intuitively, a marginal likelihood is a likelihood-based measure of model fit that penalizes models for proliferation of parameters.[12] In any discrete choice problem, calculation of the marginal likelihood represents a very high dimensional integration problem, and this problem is

[12] The marginal likelihood is the integral of the likelihood with respect to the prior. If we increase K and/or P then the prior is specified over more parameters. Thus, the prior mass in the vicinity of any particular parameter vector, such as the posterior mode, will fall. This induces an implicit penalty on added parameters.

compounded by the presence of latent types. Again, recently developed simulation methods make this problem tractable. We describe our algorithm for calculating marginal likelihoods for our models in Appendix A.[13, 14]

At this point, some limitations of our approach can be made clear. Note that 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.

This relates to our earlier discussion of the El-Gamal and Grether (1995) approach, which requires the investigator to specify *a priori* the complete set (or super set) of decision rules that will be entertained as possibilities. Our approach requires us to specify the complete set of state variables that will be entertained as possible arguments in subjects' decision rules. In our view, our approach imposes lesser demands on the analyst's intuition, but clearly it still imposes some demands. Ultimately, it is impossible to learn anything without some identifying assumptions.

Another limitation is the requirement that the investigator make a distributional assumption on the optimization errors ς_{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 empirical application described below we have assumed that agents only use the relevant state variables (i.e., the elements of I_{nt}) when forming F, and we have assumed normality for the optimization errors. Given these assumptions, our model provides a good fit to subjects' choice behavior, and an easily interpretable classification of subjects into types. Thus, we did not see any compelling reason to generalize our model along these dimensions.

---

[13] Alternative numerical procedures for calculating the marginal likelihood are discussed in Gelfand and Dey (1994), Geweke (1997), Geweke and Keane (2001), Chib (2001), Allenby, Kim and Rossi (2002), and Lewis and Raferty (1997) among others.

[14] 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.

In the next section we describe the design of the specific experiment that we analyze, and in section IV we detail how the general framework described here can be applied to model subject's decision rules in this specific problem.

III. Experimental Design

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

1) We wanted a problem that was very difficult to solve. The entire point of our experiment is to examine how people behave when confronted with a decision problem that is too difficult for anyone to solve exactly.[15]
2) We nevertheless needed a problem whose structure was easy to explain to participants in the experiment.
3) We felt it was desirable to have a problem whose structure was in fundamental ways similar to dynamic decision problems that people actually confront in real world situations, particularly situations that economists are actually interested in.
4) We wanted to design a problem where there was some noticeable advantage to playing near-optimally, so that sub-optimal behavior would be easy to detect statistically.
5) We needed a game that could be played quickly, so that subjects would be willing to participate, so that we could collect a reasonably large amount of data, and so that subjects would be able to practice. This also prevents subjects from getting bored.

We settled on a stochastic sequential discrete choice problem, that has features similar to a human capital investment or occupational choice problem. Each subject in our experiment makes 15 decisions sequentially. The problem has a nontrivial intertemporal component in the sense that early decisions influence the distributions of payoffs for later decisions. In addition, there are stochastic components to payoffs that make solving the problem optimally very difficult (because the optimal decision rule is a function of the realization of the stochastic variables). We set up the problem so that alternative "1" can be thought of similar to "school" or "white collar" work, in that this option tends to have low initial payoffs that increase later if the subject builds up sufficient experience in "1." Option "2" has a higher mean payoff, but does not have any such "investment" component. It is importance to stress however, that no such

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

interpretation was provided to the subjects, who were simply told the mathematical form of the payoff function in each alternative.

An exact transcript of the instructions that were 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 random variable that is uniformly distributed over the interval [-5000, 5000], subject to the restriction that the payoff be non-negative. For example, if the realization of the random variable was 5000, then the immediate payoff to "2" would be 9000 points, while if the realization was negative 4000 (or less) then the payoff would be zero.[16] Note that the realization of the random variables in round *t* occurs before the decision at *t* is made, but the realization of period *t*+1's random variables occurs after the decision at *t*.

The payoff to alternative "1" was 3000 points plus the realization of an independently distributed random variable distributed uniformly over the interval [-5000, 5000], plus a "bonus" and a "cost" whose values depended on the history of the subject's choices. The bonus was 7500 points, and was available whenever alternative "1" had been chosen six, seven, eight or nine previous times (not necessarily in succession). A "switching" cost of 5000 points was incurred whenever the subject chose "1" after choosing "2" in the previous round. The subject's total payoff for the decision problem was the sum of the rewards they earned over the 15 rounds.

We chose this design to study dynamic decision-making behavior because it has, at least to some degree, all five desirable features that we listed above. First, this is a sophisticated dynamic investment problem that is nevertheless straightforward to describe. It also turns out that the problem is quite difficult in the sense that it is hard to solve for the expected wealth maximizing strategy. In fact, the dynamic optimization problem cannot be solved analytically – it must be solved numerically on a computer. 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, we found that if the game lasted only 15 periods then subjects could play the game quite quickly (i.e., in 45 seconds to a minute) allowing ample opportunity for practice.[17]

---

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

[17] We originally attempted to pattern the game quite closely on a white-collar vs. blue-collar occupational choice problem, using estimates of returns to experience and transition costs calibrated from Keane and Wolpin (1997). This meant that returns to experience were about 5% per period higher in Alternative "1" than in alternative "2."

This decision problem was coded in Visual Basic and subjects made decisions independently at a private computer terminal.[18] Figure B1 provides a screen shot of the interface seen by the subjects as they made decisions. The screen provides information on the current payoffs to both alternatives, the current round, the history of the subjects choices and payoffs, their current aggregate earnings and a summary of the decision problem's payoff structure.

We report here on results obtained from 139 subjects who completed our experiment. All experiments were conducted at the Economic Science Laboratory (ESL) at the University of Arizona. Subjects were recruited from the general student population using the ESL's standard recruiting procedures. In an effort to ensure subjects were familiar with the task when they played it for money, subjects were recruited for two laboratory sessions. On arrival for the first session they were provided with the written instructions in Appendix B and seated privately at a computer terminal. They were allowed to practice the decision problem as long as they liked (up to about 90 minutes) but did not play the decision problem for money. A typical subject practiced for about 30 minutes. The median number of practice rounds played was 38.

The second laboratory session was held two days after the first (for example, if the first was on Wednesday the second would be on Friday). Upon arrival, subjects were again provided with the written instructions and told that they could practice for as long as they liked (up to about 90 minutes). Practice was generally shorter during the second session (i.e., the median number of practice rounds was 14). When they were ready, subjects solved the decision problem one time for money. During money play, there was a 15 second forced delay between moves, which we imposed to discourage thoughtless play. In hindsight, given the large amount of time the typical subject voluntarily devoted to practice, 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 laboratory (in total).

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

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

IV. 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.

IV.A. The Functional Forms for the Decision Rules

In the experiment in section III, the current payoffs w_{njt} are simply the known immediate rewards for alternatives $j \in \{1, 2\}$ in round t. For instance, the payoff in option "1," w_{n1t}, is the (possibly censored) sum of the base reward of 3000 and round t's realization of the uniformly distributed random variable, along with any transition cost that is incurred or bonus that is available.

In our experiment, the relevant state variables for forecasting values of future states are the number of times a person has chosen each alternative, which we denote by X_{n1t} and X_{n2t} for alternatives "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 rational subjects should try to get through this bonus phase before the end of the game.

Note that current payoffs, while they are elements of the information set I_{nt}, are not useful for forecasting future payoffs, because 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}. Therefore, a polynomial approximation to RE future component would not depend on the current and/or lagged w_{njt}. As we noted at the end of section II, so far we have only estimated specifications in which the future component is a polynomial in state variables that are actually relevant for forecasting future payoffs, because such specifications fit the data quite well. In future work we plan to allow the future component to include superfluous state variables (like the payoffs) to investigate whether some subjects might use this information erroneously.

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 $H(X_1, X_2, j) = (X_1 + 1(j=1), X_2 + 1(j=2), j)$, the future component F for subjects of type k takes the following form (where we have suppressed the type subscript k on the π_k parameters):

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

Note that 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. We achieve identification by differencing the future component. Let:

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

Not all of the parameters that enter the differenced future component are identified. For example, π_1, π_2, π_{10} and π_{12}, are not separately identified, since they all form a composite constant term. Thus, our analysis is based on the following identified relative future component:

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

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

Note that the decision rule for subject n of type k in round t is:

$$
\begin{aligned}
\textit{Choose "1" iff } \; Z_{nt}(I_{nt}|k) &\equiv V_{n1t}(I_{nt}|k) - V_{n2t}(I_{nt}|k) \\
&= w_{n1t} - w_{n2t} + f(X_{n1t}, X_{n2t} \mid \pi_k^*) + \eta_{nt} > 0
\end{aligned}
$$

It is therefore intuitive to think of $f(X_{n1t}, X_{n2t} \mid \pi_k^*)$ 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 $\eta_{\text{nt}} \equiv \varsigma_{n1t} - \varsigma_{n2t}$).

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}. This is what makes it very difficult to play the game optimally. Our algorithm will allow us to infer the actual reservation payoff differential function $f(X_{n1t}, X_{n2t} \mid \pi_k)$ used by each type k. We can then compare these to the optimal f in order to characterize the manner in which play of each type of subject deviates from optimality.

IV.B. The Likelihood Function, Priors, and Joint Posterior Distribution of Parameters

We denote the round t choice of subject n by d_{nt}. Let $\tau_n \in \{1, \ldots, K\}$ be an indicator of subject n's type. If the subject uses decision rule k (i.e., $\tau_n = k$) 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 data observed by the investigator will be a sequence of current payoff realizations w_{njt} for $j=1, \ldots, 2$, $t=1, \ldots, T$ and choices d_{nt} for $t=1, \ldots, T$ for each subject n. Note that choices only depend on the value function differences Z_{nt}. The inferential problem is complicated by the fact that these, as well as the subject types τ_n, are unobserved.

We assume $\eta_{nt} \sim IIDN(0,\sigma_k^2)$ 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 π_k and σ_k parameters, so both are identified.

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

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

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

$$\mathscr{L}[\{\{d_{nt}\}_{t=1,T}\}_{n=1,N}\,|\,(\pi_k^*,\sigma_k^{-2})_{k\in K}]=\prod_n\prod_k\prod_t \mathscr{P}(d_{nt}\,|\,k)^{1(\tau_n=k)}.$$

and maximization of this likelihood with respect to $\{\pi_k, \sigma_k^{-2}\}_{k=1,K}$ would be straightforward. However, since we do not know subjects' types, we must form a likelihood function based on unconditional choice probabilities $\mathscr{P}(d_{nt}) = \sum_k \theta_k \mathscr{P}(d_{nt}|k)$, where θ_k is the probability that a person chosen at random from the population follows decision rule k. This gives:

$$\mathscr{L}[\{\{d_{nt}\}_{t=1,T}\}_{n=1,N}\,|\,(\pi_k^*,\sigma_k^{-2},\theta_k)_{k\in K}]=\prod_n\left\{\sum_k\left(\theta_k\prod_t\mathscr{P}(d_{nt}\,|\,k)\right)\right\}$$

We do not pursue maximum likelihood (ML) estimation of this model for two reasons. First, ML estimation of mixture models for discrete data is notoriously difficult due to problems with local maxima of the likelihood function. Second, the ML approach would not provide us with a rigorous and practical method of testing for the number of types.

Thus, we use a Bayesian Markov Chain–Monte Carlo (MCMC) algorithm to generate inferences about the number of types K, the order P of the F polynomials, and the model parameters $\{\theta_k, \pi_k^*, \sigma_k^{-2}\}_{k=1,K}$. The particular MCMC algorithm we employer is the Gibbs sampler. This provides draws from the joint posterior distribution of the model parameters conditional on the data. A key feature of the algorithm is that there is no separate maximization step. That is, the draws from the posterior are obtained without the need to maximize the likelihood function. Thus, the Gibbs sampler is much less sensitive to problems created by ill behaved likelihood surfaces (i.e., local maxima) than is ML.[19]

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 τ_n were observed. In our model, given a particular K and P, this is:

$$\mathscr{L}(\{Z_{nt}\}_{\substack{n=1,\ldots N,\\ t=1,\ldots,T}}, \{\tau_n\}_{n=1,N} \mid \{\theta_k, \pi_k^*, \sigma_k^{-2}\}_{k=1,K}) \propto$$

$$(4) \qquad \prod_{k=1,K} \prod_{n:\tau_n=k} \left[\theta_k \prod_{t=1,T} \frac{1}{\sigma_k} \exp\left\{ -\frac{(Z_{nt} - (w_{n1t} - w_{n2t} + Y_{nt}^{'}\pi_k^*))^2}{2\sigma_k^2} \right\} I(Z_{nt}, d_{nt}) \right].$$

where $Y_{nt}^{'}$ denotes the vector of state variables conformable with π_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:

$$(5) \qquad \begin{aligned} &\pi_k \sim N(0, \underline{\Sigma}), \text{ where } \underline{\Sigma} \text{ is a 6 x 6 diagonal matrix with entries } \underline{\Sigma}(1,1) = 20{,}000^2, \\ &\underline{\Sigma}(2,2) = \underline{\Sigma}(3,3) = 1{,}000^2, \underline{\Sigma}(4,4) = \underline{\Sigma}(5,5) = 100^2 \text{ and } \underline{\Sigma}(6,6) = 100{,}000. \\ &\sigma_k^{-2} \sim \chi^2(1),\ \sigma_1^2 > \sigma_2^2 > \cdots > \sigma_K^2 \\ &\{\theta_k\}_{k\in K} \sim Di(\{2.0\}_{\|K\|}). \end{aligned}$$

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 τ_n is simply proportional to the complete data likelihood times the prior densities $p(\pi_k^*)$, $p(\sigma_k^{-2})$, and $p(\theta_k)$.[20] 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 π_k^* vector at zero means that our prior is centered on myopia.

---

[19] For instance, in a Monte-Carlo experiment, Geweke, Houser and Keane (2001) reported no success whatsoever in attempting to estimate by ML a model rather similar to that described here. In contrast, their Gibbs sampling algorithm produced reliable inferences about the model parameters in the simulated data.

[20] Note that in the Bayesian approach there is no distinction between the parameters of interest and the latent variables, at least from the point of view of computation of the posterior.

When all element of the π_k^* vector are zero, the future component F is zero, and subjects only consider current payoffs. The issues involved in choosing $\underline{\Sigma}$ illustrate why it is not possible to have "uninformative" priors. If we made the priors on the π_k^* very flat (i.e., set the elements of $\underline{\Sigma}$ to be very large), this would imply that there is little prior mass in the vicinity of $\pi_k^* = 0$. 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 π_{k0}^* is the $f(X_{n1t}, X_{n2t} \mid \pi_k)$ function intercept, and can therefore be interpreted as the reservation payoff differential in round 1, prior to the accumulation of any experience in either option. According to the optimal decision rule, this is 3733 points (i.e., 37 cents). We decided to specify a prior mean of 0 and a prior standard deviation of 20,000 for this parameter. Thus, our prior is rather diffuse, while still leaving 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 π_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 π_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 π_k^* parameters by 50%, and find little effect.

Third, consider the prior specification for the optimization error variance, σ_k^2. Note that 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).

Also note that the prior mean and standard deviation for σ_k^2 are not defined under the $1/\chi^2(1)$ prior. We can, however, consider quantiles of the σ_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, which is quite small relative to the magnitudes of payoffs. We did this in order to favor models where the state variables largely explain behavior, as opposed to models where behavior is largely random. However, the prior density also has a very fat right tail (which accounts for the nonexistence of the mean), so that models with very large optimization errors are still given non-negligible prior mass. The prior is also quite weak. As we'll see in the next section, 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 σ_k^2 prior to $20{,}000/\chi^2(1)$, 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 situations with very unequal proportions will have substantial prior mass. For example, in the model with three types, the prior standard deviation on each of the type proportions is 18 percent.

IV.C. The Gibbs Sampling Algorithm

We now describe the Gibbs sampling algorithm that we use to approximate the marginal posteriors of the model's parameters. [21] 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 π_k^* for all $k = 1, K$.
3) Draw variance of optimization error $\sigma_k^2, k = 1, K$.
4) Draw population type probabilities $\theta_k, k = 1, K$.
5) Draw individual types $\tau_n, n = 1, N$.

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

These draws are implemented as follows:

1) Z_{nt} Step: Conditional on everything else being known, it is clear from equation (3) that a given Z_{nt} follows a truncated normal distribution, with mean $(w_{n1t} - w_{n2t} + Y_{nt}^{'}\pi_k^*)$ and variance σ_k^2. The truncation is from below at zero if $d_{nt} = 1$, and is from above at zero otherwise. It is trivial to draw from this distribution using an inverse CDF procedure.

2) π_k^* Step: Conditional on everything else being known, each vector π_k^* can be drawn using simple rejection methods. From equation (4), the source distribution can be chosen to be normal with a mean of $(Y_k^{'}Y_k)^{-1}Y_k^{'}W_k$, and variance $\sigma_k^2(Y_k^{'}Y)^{-1}$. Here, Y_k denotes the stacked array of Y_{nt} vectors for those subjects who are type k. The vector W_k is created by stacking the numbers $Z_{nt} - w_{n1t} + w_{n2t}$ conformably, again for subjects who are type k. The kernel of the normal prior for π_k^* is evaluated at each candidate draw from the source distribution, and that evaluation compared against a random number drawn from a uniform [0,1] distribution. If the uniform random variable is less than the value of the kernel evaluated at the candidate draw, the draw is accepted. Otherwise it is rejected.

3) σ_k^2 Step: Conditional on everything else being known, equations (4) and (5) imply that:

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

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 $\sigma_1^2 > \ldots > \sigma_K^2$.

4) θ_k Step: Because we assume the prior $Di(\{2\}_K)$, it is trivial to verify that the posterior conditional on everything else being known is $Di(\{2 + N_k\}_{k=1,K})$. We draw from this distribution using the procedure suggested by Anderson (1984, p. 284).

5) τ_n Step: Let $\mathscr{L}_k(n)$ denote the value of the likelihood (4) when evaluated only for individual n and under the assumption that he/she uses decision rule k and that everything else is known. Then, the conditional posterior distribution of τ_n is

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

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.

V. Empirical Results

V.A. A Basic Description of the Data

We begin with a simple comparison between the average behavior of the 139 experimental subjects and "optimal" behavior in the experiment. That is, 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 in the experiment.[22] But our hypothetical agents play the exactly optimal decision rule - which we call the "rational expectations" (RE) decision rule.[23] 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 strongly indicates that most subjects understand the investment component of option "1," and realize that they should choose "1" unless "2" offers a substantial payoff premium. This *reservation payoff differential* happens to be 3733 points (i.e., 37 cents) in round 1 according to the optimal rule. Also interesting is that the fraction of RE subjects who choose "1" is 79%, which is very close to the fraction of experimental subjects who chose "1." Of course, the fact that the aggregate percentage who chose "1" is about right does not necessarily imply near-optimal behavior at the individual level, since some types of subjects may be choosing "1" too often, while others choose it too rarely.

---

[22] While running the experiment, we recorded both the choices made by subjects and the idiosyncratic random variable realizations that they faced. Payoffs depend on both the random draws and the past history of choices. So, of course, the alternative specific payoffs in a particular round may differ between actual subjects and their RE counterparts who face the same random draws (but who may have different choice histories).

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

Figure 1 also reveals an interesting non-stationarity in choice behavior in this game that is implied by the optimal rule. 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 end.[24] 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 looking at the aggregate statistics reveals little about the play of individual subjects, because there is substantial variation in play around these 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 that different subjects are using different decision rules.

V.B. 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 π_k^* are reduced or

[24] 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 that great. But as the

increased by 50%, and (2) the scale of the optimization errors is increased by a factor of 141 = $(20000)^{1/2}$. As we would expect, since the marginal likelihood is the integral of the likelihood with respect to the prior, it changes somewhat when the prior is changed. The important question is whether the model selection is affected. In fact, the three-type model with a quadratic differenced future component is strongly preferred regardless of the prior specification.

For the preferred model, our inferences were 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 diagnostic suggested by Gelman (1996), convinced us that the Markov chain had converged by the 3,000th cycle. The draws from the last 5,000 cycles were used as simulation estimates of 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. Note that the posterior standard deviations are in all instances quite small relative to the prior standard deviations, indicating that the data is very informative about the parameters. Also, in all instances the posterior mean is within a fraction of a *prior* standard deviation of the prior mean. This suggests that the prior did not strongly influence the posterior for any parameter.[25]

Of course, as is usually the case, the polynomial coefficients for the three decision rules are difficult to interpret, and do not in themselves tell us much about the nature of the rules. 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 an exception is the intercept. As we noted earlier, the intercept 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. A striking feature of the results in Table 2 is that π^*_{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.

game 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.

[25] In contrast, in a case where the posterior mean departs from the prior mean by several *prior* standard deviations, it would typically be the case that the prior is exerting a substantial influence on the posterior.

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 (conditional on his/her observed choice history).

Interestingly, it turned out the vast majority of subjects can be assigned to one type very clearly, because the highest posterior type probability was, for 86.3% of subjects, at least 90%. This means that a subject's choice behavior in our experiment is usually highly informative about which type that subject is. In hindsight, this is not surprising. For example, if a subject consistently received good payoff draws for option "1," then 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 during the course of the game, so there is usually plenty of opportunity to reveal one's type.[26]

Figure 2 assesses the fit of the three estimated decision rules to the actual play of the subjects whom 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.

[26] Only 19 out of 139 subjects had an assignment probability that was less than 90% for their most likely type. Interestingly, 15 of these 19 had essentially all the posterior mass distributed between type 1 and type 2. Thus, when there is ambiguity, it usually arises in attempting to distinguish these two types.

Figure 3 describes the fit of the three-type model to the aggregate choice frequencies. It reports the fraction of actual and hypothetical subjects who choose alternative "1" in each round of the game. The fit seems reasonably close in all rounds, with the broad features of actual decisions, such as the peak that occurs in the ninth round, well matched by decisions under the estimated rules. Note that the model captures the departure of actual play from RE play that occurs beginning in round 11 quite well.

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

Figure 4 describes the results of 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 and that we noted earlier when commenting on Figure 1. That is, both rules generate the drop in choice frequency for option "1" to a trough in round 3, and the gradual rise 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 tracks 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 <u>all</u> the other rules. As we saw in Table 2, all three types set a reservation payoff differential close to the optimal level of 3733 in round 1, and hence they all choose option "1" at close to the optimal frequency of 79% in the first round. Under myopic play, only 40% of subjects would choose option "1" in the first round, since it has a lower median payoff. Thus, all three types recognize that there is an investment component to the choice of option "1," and hence they all set positive values for the reservation payoff differential between "1" and "2."

The myopic rule does generate a drop in choice frequency for option "1" over the first three rounds, just like the RE rule.[27] But, unlike the RE rule, the myopic rule does not generate the subsequent gradual rise in the fraction choosing "1" toward a peak in round 11. It generates a fraction choosing "1" that stabilizes at about 20% from period 3 onward. It is interesting that the type three rule also fails to generate the gradual increase in the frequency of "1" after round three, just like the myopic rule. 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.[28]

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 5A 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, since a subject could have chosen option "1" either 0, 1, 2, 3 or 4 times by the start of round five). 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 the number of times alternative one was previously chosen, 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 5A. We see that the reservation payoff differential for type ones is about right in round one. In round five, note that the reservation payoff differential is declining in the stock of "experience" in option "1." This pattern holds both for type one subjects and under the RE rule. Also, note that, holding the number of prior choices of "1" fixed at zero, the 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

---

[27] This occurs because the 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 (because some people chose "2" in round one) and so on.

[28] The astute reader may notice an apparent contradiction between the behavior of the type three subjects in Figure 4, compared to the behavior of type threes described in Figure 2 and Table 3. Figure 4 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. What is going on here 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. But Figure 4 reveals that, if faced with the same draws as other types, type threes would choose "1" much less frequently. This resolves the apparent contradiction. We will have more to say about these lucky draws for type threes in section V.D, where we ask whether the draws subjects received may have influenced the type of decision rule they adopted. To anticipate that discussion, we find no evidence of this.

reach the bonus phase. So the urgency to choose option "1" is growing. The type one subjects appear to understand this very well. They get the reservation payoff differential almost exactly right at state points $X1=1, \ldots, 4$ in round 5. They appear to be a bit off at $X1=0$, but this estimate is noisy due to limited data at that point (i.e., type one subjects are very unlikely to get to round five without having chosen option "1" at least once).

Figure 5B 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 5B, the type one subjects do not understand this. They continue to set a small but positive reservation differential during the bonus phase.[29]

One might suspect that the failure of the fitted decision rule to align well with the RE rule in period 9 at states $X1=6, 7, 8$ merely stems from a failure of our polynomial approximation to the future component in this range. But an 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 is accurately reflecting this feature. [In fact, this finding illustrates well the power of our approach. We certainly would not have thought to look for this pattern in the data unless Figure 5B 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

[29] In figure 5B, in round 13, the option value argument also applies to the $X1=8,9$ observations, and type one subjects again fail to recognize this.

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 5A and 5B. Figure 5A 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 does happen 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" if he/she has already chosen it a few times in the first few rounds. 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 the bonus phase and get through it completely before the end of the game.[30]

---

[30] In fact, 94% of the hypothetical RE subjects complete the bonus round (as do 94% of the "Near-Rational" actual subjects), compared to only 69% of the "Fatalists."

Finally, we turn to the type three subjects, whose behavior is described in the bottom panels of Figures 5A and 5B. 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 $X1<6$ (i.e., before the bonus round has started). It appears that type three's very 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 the optimization error standard deviation for type threes is much larger than for the other types (see Table 2).

V.D. Addressing Some Common Concerns

We have heard two common concerns expressed regarding latent type classification algorithms in general, and our type assignments in particular. The first is a general concern that arises out of prior work by Heckman and Singer (1984). In a Monte-Carlo study of duration models with latent discrete types, they found that the estimated type specific parameters were rather poorly estimated, even though the model fit behavior well at the aggregate level. This has generated a folklore to the effect that one shouldn't take the types one estimates in latent class models "too seriously." A related claim is that latent type models typically underestimate the true number of types that exist in a population. To help dispel these concerns, we present two Monte-Carlo experiments, which show 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 exactly zero. And we added a "future oriented" type. They have double the intercept of the "Near-Rational" future component (and the remaining π_k^* values set to zero). Given these parameter values, this type strongly prefers to choose “1.” The Gibbs sampler was run on each sample, using the same baseline prior as we described in section IV.B, and using *P=3*. The results of this exercise are reported in Table 4 and Figure 6.

In Table 4, note that 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. The results in Figure 6 indicate that the models assign subjects to the correct type with a reasonably high degree of accuracy. For example, in the three-type model, the Near-Rational subjects are assigned to the Near-Rational decision rule with a posterior probability of 70% on average. We also found that the estimated decision rules for each type closely resembled 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. 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).

Another concern that we have often encountered is more particular to our specific experiment. The 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 such an 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 classification 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 iid, 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. This point will be emphasized by some example individual choice histories we present in the next section.

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.[31] 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. If anything, the fact that type three subjects were somewhat lucky and got better than average draws for “1” makes it even more mysterious they chose “2” far too often, adding to the impression that they were “Confused.”

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.

V.E. Examination of the Individual Level Data

Having classified subjects into types, we can now examine the individual behavior and individual characteristics of the subjects classified as each type. This serves two purposes. First, it helps give face validity to the type assignments, and second, it enables us to learn more about each type.

Appendix C provides some example subject choice histories that we felt were particularly illustrative of the nature of the game and of our conclusions. Consider the data for subject number 8, who our algorithm classifies as type one (i.e., “Near-Rational”). This person happened to get very good draws for option “2” in rounds 2 through 6, making it optimal to choose “2” in each of those rounds. Note that the choice in round 6 was very difficult. The optimal reservation payoff differential for choosing “2” was 5921 (i.e., about 60 cents). The draw generated a payoff for option “1” of 1807 points. Thus, “2” would have been the optimal choice only if the draw for “2” exceeded 7728 points. The actual payoff draw for “2” was 7828, which barely exceeds the reservation level. Subject 8 made the optimal choice. This meant that

[31] The fact that type three subjects got relatively good draws for “1” explains the fact that in Figure 4, when the estimated type three decision rule is confronted with typical draws, it implies choice of “1” at a lower frequency

he/she got to period 7 with only 1 unit of accumulated experience in "1." Hence, the urgency for choosing "1" became substantial, and the optimal reservation payoff differential jumped to 10070 (i.e., about a dollar). Subject 8 then chose "1" in rounds 7 through 11, enabling him/her to reach round twelve with the requisite 6 units of accumulated experience in "1." Thus he/she was just able to get in the four bonus periods in rounds 12 through 15. This subject played exactly the RE choice sequence, which is quite impressive given the difficulty of the round 6 choice.

We also show the data for subject number 104, who our algorithm also classifies as type one (i.e., "Near-Rational"). This subject reaches the bonus phase in period 11. Since the bonus phase lasts until "1" is chosen 4 more times, and since 5 rounds remain in the game, there is an option value to choosing "2" at that point. Thus, note that the reservation payoff differential actually turns negative. Subject 104 failed to make the optimal choice in round 14. He/she chose "1" even though its payoff was only 5 cents higher than "2." Optimal play was to choose "2," thus extending the bonus phase and hoping for a better draw for "1" in round 15. Thus, subject 104 illustrates the failure of subjects to appreciate option values in this context. Nevertheless, subject 104's behavior is reasonably described as "Near-Rational." His/her error in round 14 led to an expected loss of less than 5 cents, and this subject made some fairly difficult correct decisions earlier in the game (see especially rounds 6 and 7).[32]

Of our 139 subjects, 11 played exactly the "optimal" choice sequence implied by the RE decision rule. The fact that only 8% of subjects played exactly optimally illustrates the difficulty of the game. All 11 of these subjects were classified as type one by our algorithm. Another 11 subjects who were classified as type one departed from the RE rule but actually earned more than a hypothetical RE subject.[33] Subject 104, who we discussed above, is one of these. Thus, of the 51 total subjects classified as the "Near-Rational" type, 22 either followed the RE rule exactly or departed from the rule but actually beat the RE payoff.

We also provided an exit interview to subjects after the experiment. Participation was voluntary, and only about a third of our subjects responded. The questionnaire is reproduced in

---

than the type threes actually exhibited (see Table 3 and Figure 2). We commented on this in footnote 28.

[32] *Ex post*, subject 104 got lucky. The draw for option "1" was even worse in round 15. Thus, by playing sub-optimally, subject 15 beat the RE payoff.

[33] There were five other subjects in the experiment who departed from the "optimal" choice sequence and beat the RE payoff. Three of these were classified as type two and two were classified as type three. Of course, it is to be expected that, occasionally, some subjects will play clearly sub-optimally but nevertheless achieve better than optimal payoffs, simply due to luck. It is worth noting that our algorithm recognizes the sub-optimality of such subject's choices, even though they result in high payoffs.

Appendix D. It is interesting to compare the answers that subjects of different types provided. In the survey, subjects were asked whether they followed a particular strategy and, if so, to describe it. Not all subjects indicated they were following any particular strategy. However, conditional on reporting a strategy, the classification often seems reasonable. The type one subjects who reported a strategy typically gave sensible answers that indicated some awareness of the investment value of choosing alternative "1" early in the decision problem. For example, one subject responded, "Choose 1 unless 2 is very high." Another said, "My strategy was to give myself the opportunity to have the bonus in the payoff period."

In contrast, many subjects who were classified as type three gave responses that indicated confusion about the nature of the decision problem. One explained that he/she attempted "to keep choosing 2, or amount of 5-7 dollars in average. Keep choosing 1 for more profit, but maybe less profit too." Another indicated that he/she made an effort to "Push 1 several times in the first 10 periods, and then choose the bigger payoff after that." Yet another said that his/her goal was to, "choose 1 at least 9 times in a row." Answers indicative of confusion were much more apparent among type three subjects than either type one or type two subjects. Thus, it does seem reasonable to label the type threes as "Confused."

Finally, it is interesting to examine the data on practice rounds, which we present in Table 6. A dogmatic defender of complete rationality might well argue that all subjects were, in principle, capable of solving such a 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 (i.e., just a few dollars) to elicit substantial effort. In Table 6, we see that 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.

Figure 7 presents data on the entire cumulative distribution of practice rounds by type. Note that only a small fraction of "Confused" subjects devoted little time to practice. For instance, only about 8% practiced less than 25 times, and only about 12% practiced less than 40 times. Among the "Near-Rational" subjects, a small number of practice rounds was much more common. About 15% practiced less than 20 times, and about 30% practiced less than 40 times. This shows clearly that those subjects who put in the least effort, in terms of number of practice rounds, also tended to be the ones who did best in the experiment.

VI. Conclusion

We have described a new Bayesian procedure for classification of subjects into decision rule types in choice experiments. We report two Monte-Carlo experiments in which the procedure does an excellent job of identifying both the number of nature of the decision rules that were operative in a hypothetical population of subjects. And it produced quite accurate assignments of the hypothetical subjects into their respective types.

We also applied the procedure to data from a particular experiment on choice behavior in a dynamic optimization setting. The procedure produced a clear classification of the subjects into three behaviorally distinct types. The three-type specification appeared to have considerable face validity, given inspection of the individual level data, and the characteristics of subjects classified into each type.

More than a third of the subjects in our experiment followed a rule very close to the rational expectations (expected wealth maximizing) rule. We label these subjects "Near-Rational," since their play resulted, on average, in only about a 2% payoff loss relative to optimal play. Another 40% of our subjects played according to a sub-optimal rule that resulted in a 12% payoff loss 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 the optimal strategy on average. We label this the "Confused" type. Confusion in other experimental environments is well documented (see, e.g., Andreoni (1995) or Houser and Kurzban (2002)), and our results suggest that it exists in dynamic decision problems as well.

We were surprised that more than a third of our subjects were able to "learn" to play nearly optimally in a very difficult dynamic problem after about a half hour of practice (on average). Even more surprising was the data on practice rounds. This indicated that almost all subjects devoted considerable effort to learning to play the game, and that those subjects who practiced least tended to perform best.

Our work raises two interesting questions that we will investigate in future work on dynamic choice behavior. 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.

The Bayesian procedure for decision rule classification that we have described in this paper can be applied in many settings other than that illustrated by our particular dynamic choice experiment. 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. Our method could be used to study learning by analyzing data in which subjects play for money after varying amounts of practice. And, the method could be applied to field as well as experimental data. In each of these cases, the investigator would learn in a flexible way about the number and nature of decision rules at work in a population.[34]

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,

[34] Of course, in field data, differences in observed decision rules could stem from differences in preferences as well as differences in decision-making heuristics. Moreover, there might also be heterogeneity in the nature of the problem people solve in the field (e.g., in terms of constraints or the information available). In the experimental

and those who followed the representativeness heuristic.[35] 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. We believe that developing empirical strategies to investigate and determine the implications of such heterogeneity for policy and welfare analysis provides an important research agenda.[36]

laboratory the researcher is able to exercise control over both period return functions and the nature of the problem that subjects are supposed to solve, thus more clearly isolating the sources of any behavioral heterogeneity.

[35] Typing has also been attempted in other experimental environments, such as the public goods game (see, e.g., Gunthorsdottir, Houser, McCabe and Ameden (2000)) and extensive form "trust and reciprocity" games (see, e.g., McCabe, Houser, Ryan, Trouard and Smith (2001)).

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

References

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

Allenby, Greg, Jaehwan Kim and Peter Rossi. 2002. Modeling consumer demand for variety. *Marketing Science*, forthcoming.

Anderson, T.W. 1984. An introduction to multivariate statistical analysis. 2nd Edition. John Wiley and Sons. New York.

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

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

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

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. Forthcoming. Inferring strategies from observed actions: A nonparametric, binary tree classification approach. Journal of Economic Dynamics and Control.

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. Pp. 131-43. 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. (Also available as Federal Reserve Bank of Minneapolis working paper #564, January, 1996.)

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. (Also available as Federal Reserve Bank of Minneapolis staff report #237).

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.

Gunthorsdottir, A., D. Houser, K. McCabe and H. Ameden. 2000. Excluding free riders improves reciprocity and promotes the private provision of public goods. Manuscript. University of Arizona.

Haltiwanger, J. and M. Waldman. 1985. Rational expectations and the limits of rationality: An analysis of heterogeneity. *American Economic Review*. *75*, 3, 326-340.

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

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.

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

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

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

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 estimate 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.

The marginal likelihood value for a model, which we denote by $g(\cdot)$, 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:

$$g(\{\{d_{nt}\}_{t=1,T}\}_{n=1,N} \mid M_{KP}) = \int \mathscr{L}[\{\{d_{nt}\}_{t=1,T}\}_{n=1,N} \mid (\pi_k, \sigma_k^{-2}, \theta_k)_{k\in K}]\, p((\pi_k, \sigma_k^{-2}, \theta_k)_{k\in K} \mid M_{KP})\, d((\pi_k, \sigma_k^{-2}, \theta_k)_{k\in K})$$

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

$$g(D) = \int \mathscr{L}[D \mid \xi] p(\xi) d\xi.$$

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

$$g(D) \approx (2\pi)^{\lambda/2} \mid H^* \mid^{1/2} \; p(\xi^*) \mathscr{L}(D \mid \xi^*)$$

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

(A1) $$\log\{g(D)\} \approx \frac{\lambda}{2}\log\{2\pi\} + \frac{1}{2}\log\{\mid H^* \mid\} + \log\{p(\xi^*)\} + \log\{\mathscr{L}(D \mid \xi^*)\}.$$

Lewis and Raftery (1997) call this the Laplace-Metropolis estimator. This estimator is attractive because the quantities ξ^* and H^* can both be easily derived from Gibbs sampler output. To determine ξ^* we evaluated $h(\xi)$ at each draw from the posterior simulation, and chose that parameter vector for which $h(\xi)$ 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, ξ^*) 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	
	0	if you have chosen "1" 0, 1, 2, 3, 4, or 5 previous times
Bonus:	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.

Appendix C. Example choice histories

Period	Unconditional RE Choice	Conditional RE Choice	Actual Choice	Payoff Alternative 1	Payoff Alternative 2	Experience in Alternative 1	Reservation Payoff Differential	Reservation Payoff for Alternative 2
1	1	1	1	2749	0	0	3733.05	6482.05
2	2	2	2	3173	7707	1	3601.55	6774.55
3	2	2	2	0	5577	1	3873.12	3873.12
4	2	2	2	1115	7804	1	4538.13	5653.13
5	2	2	2	0	6414	1	4927.4	4927.4
6	2	2	2	1807	7828	1	5921.1	7728.1
7	1	1	1	0	779	1	10070.09	10070.09
8	1	1	1	682	1786	2	10045.29	10727.29
9	1	1	1	1932	6167	3	10068.25	12000.25
10	1	1	1	1061	4134	4	10088.79	11149.79
11	1	1	1	0	4662	5	10106.35	10106.35
12	1	1	1	9730	1594	6	4875.82	14605.82
13	1	1	1	15062	2181	7	4819.64	19881.64
14	1	1	1	10824	8791	8	4026.49	14850.49
15	1	1	1	8339	115	9	0	8339

Total payoff (points): 85709

Subject 8, Type 1, "Near-Rational" **RE payoff (points): 85709**

Period	Unconditional RE Choice	Conditional RE Choice	Actual Choice	Payoff Alternative 1	Payoff Alternative 2	Experience in Alternative 1	Reservation Payoff Differential	Reservation Payoff for Alternative 2
1	1	1	1	5025	7733	8758.05	0	3733.05
2	2	2	2	0	5328	3601.55	1	3601.55
3	1	1	1	0	2774	3873.12	1	3873.12
4	1	1	1	5705	394	9468.09	2	3763.09
5	2	2	2	707	8467	4266.79	3	3559.79
6	2	2	2	0	5076	4171.17	3	4171.17
7	2	2	2	0	5470	4608.14	3	4608.14
8	1	1	1	1794	291	7315.73	3	5521.73
9	1	1	1	5203	6386	10533.15	4	5330.15
10	1	1	1	0	0	5059.3	5	5059.3
11	1	1	1	12212	4594	12124.71	6	-87.29
12	1	1	1	6153	2780	6032.72	7	-120.28
13	1	1	1	9426	193	9109.75	8	-316.25
14	2	2	1	9873	8960	8499.18	9	-1373.82
15	2	2	2	3649	7323	3649	10	0

Total payoff (points): 87055

Subject 104, Type 1, "Near-Rational" **RE payoff (points): 86142**

Appendix D. Questions asked in the post-experiment questionnaire

Answers to these questions were voluntary, and subjects were not paid to complete this questionnaire. The questions were as follows:

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

.

Table 1: Model Selection Based on Marginal Likelihood Values

Marginal Likelihoods under Baseline Prior

Number of Types	Order of Future Component 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
π_0^* : Intercept	0.0	$2\text{x}10^4$	4259.21	108.60	3478.47	209.06	3711.68	476.31
π_1^* : X1	0.0	10^3	-4.42	22.51	411.33	38.00	-429.17	82.87
π_2^* : X2	0.0	10^3	92.57	42.90	276.70	49.67	-447.22	120.54
π_3^* : X1^2	0.0	10^2	-29.29	1.49	-32.47	2.85	-16.97	5.82
π_4^* : X2^2	0.0	10^2	-73.35	5.39	-3.57	2.38	-2.89	6.98
π_5^* : X1*X2	0.0	$10^{5/2}$	-86.00	2.47	-1.17	4.92	-103.36	11.61
σ_η : Optimization error	Not Defined	Not Defined	208.91	58.43	863.81	29.50	2270.96	78.73
θ_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 (percent) of subjects	51 (37%)	55 (40%)	33 (24%)
Mean earnings (points)	87983	80811	75966
Mean earnings under RE (points)	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.

Figure 1: Fraction of Alternative One Choices by Round for Actual Subjects and their Rational Expectations Counterparts

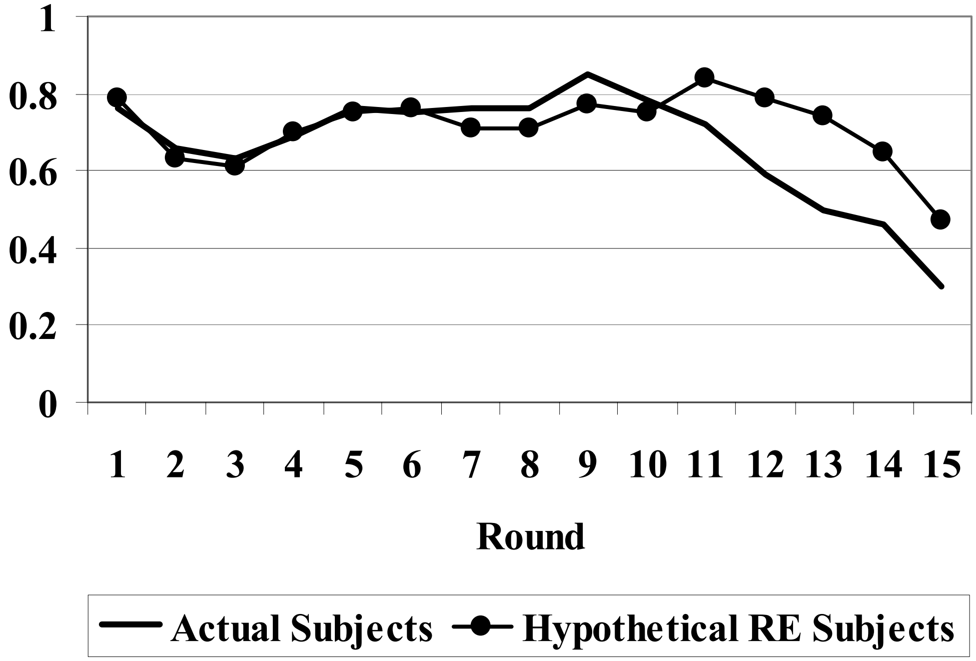
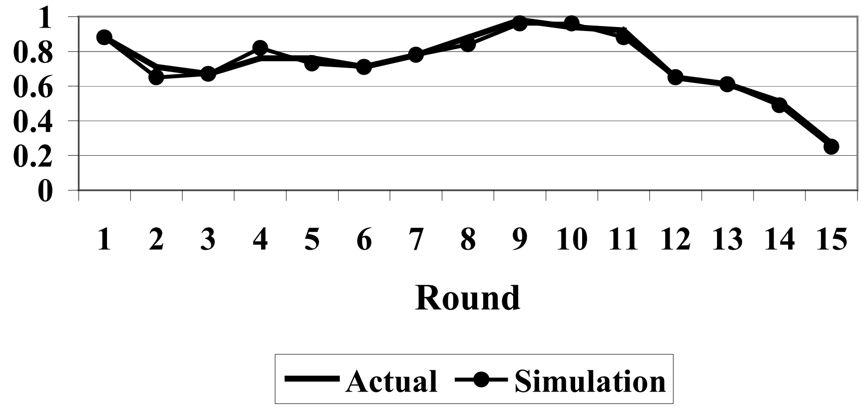
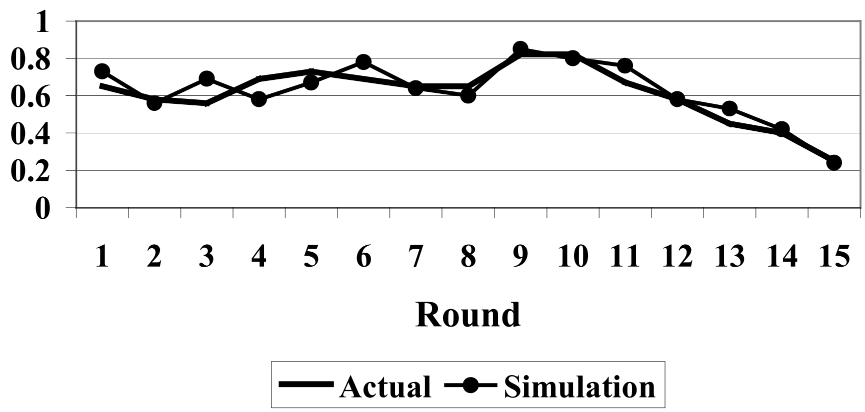


Figure 2: Fraction of Actual and Simulated Alternative One Choices by Round for Each Type

Near-Rational



Fatalist



Confused

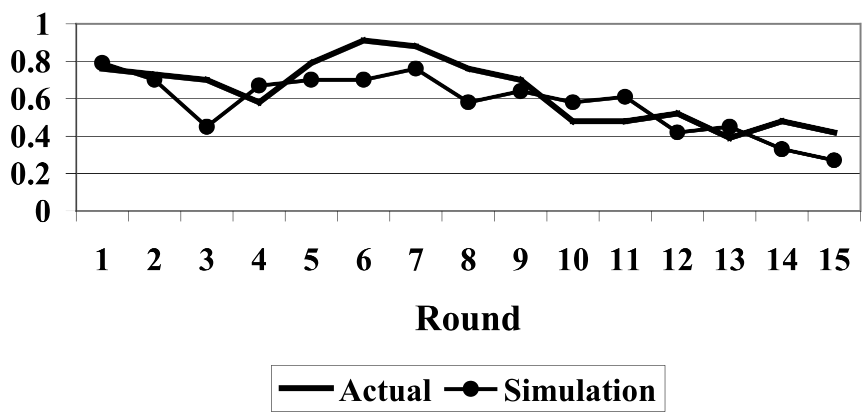


Figure 3: Aggregate Fraction of Alternative One Choices by Round for Actual Subjects and their Simulated Counterparts

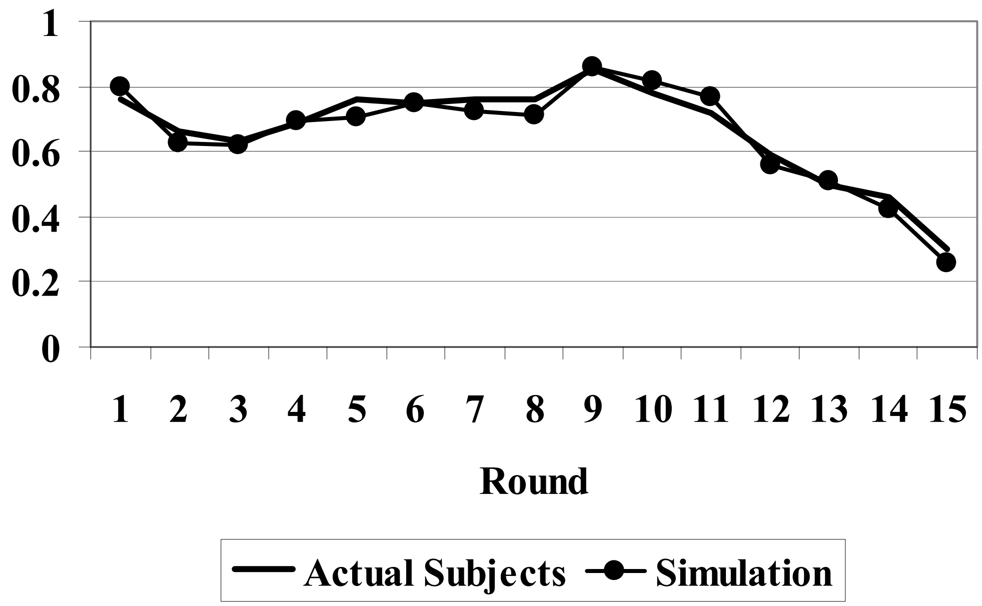


Figure 4: Fraction of Alternative One Choices Under Simulations of Various Decision Rules

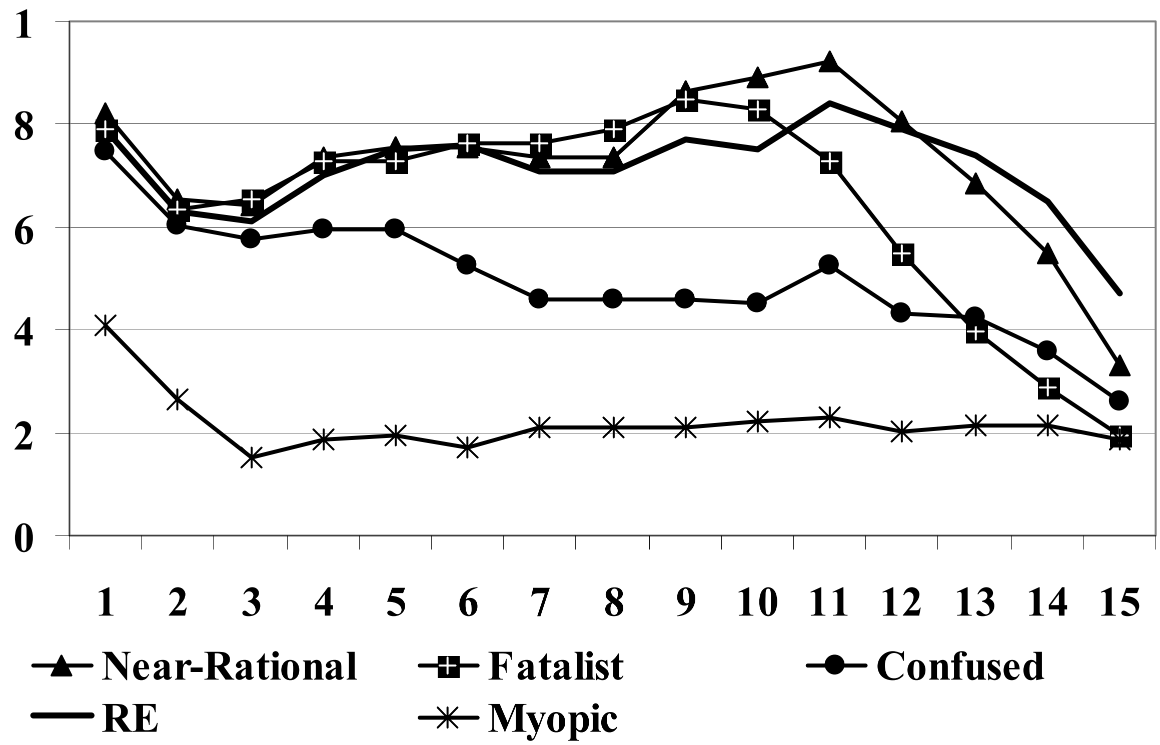


Figure 5A: Comparison of Fitted and Rational Expectations Future Component at Early State Vectors for Each Type

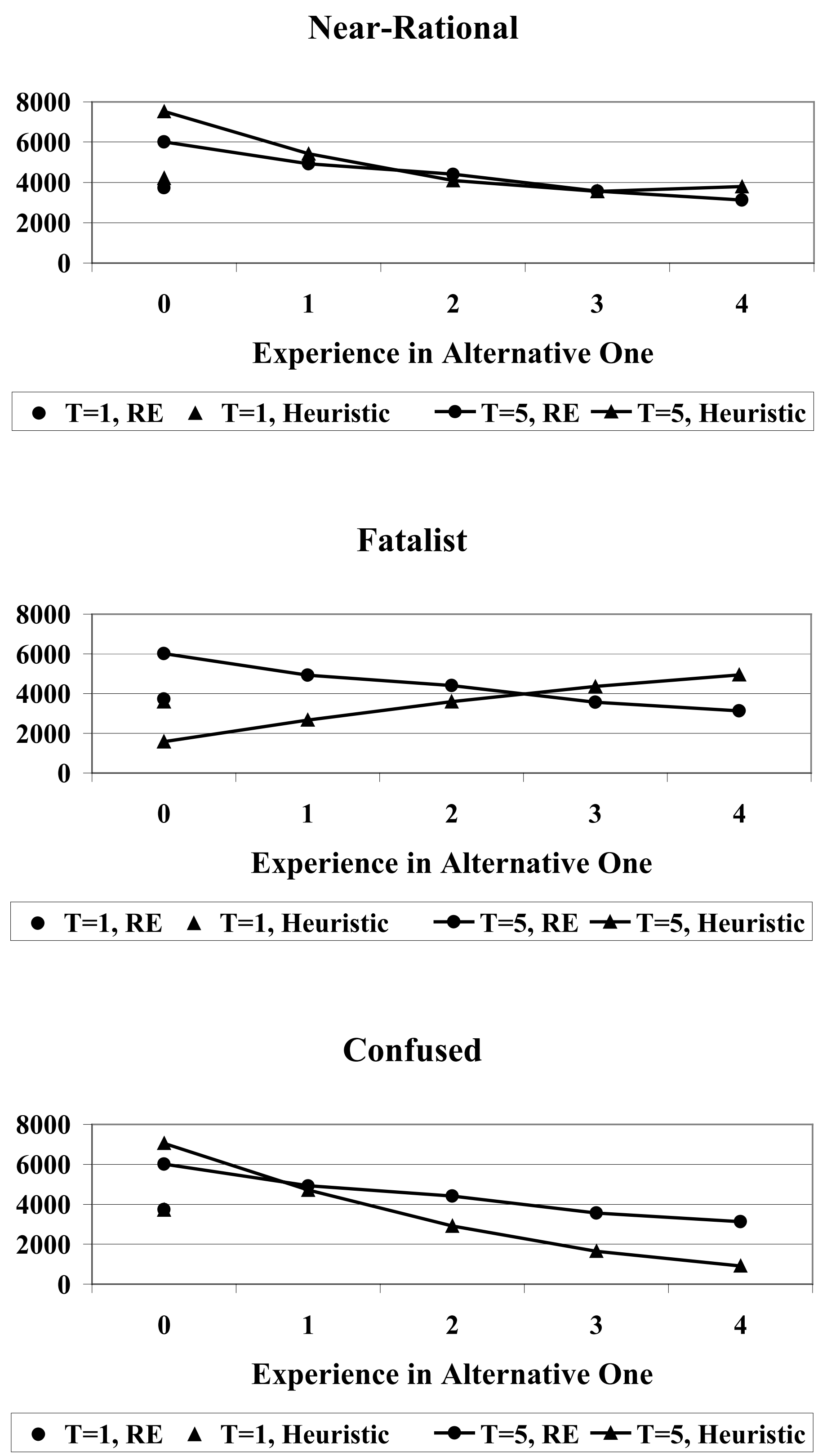
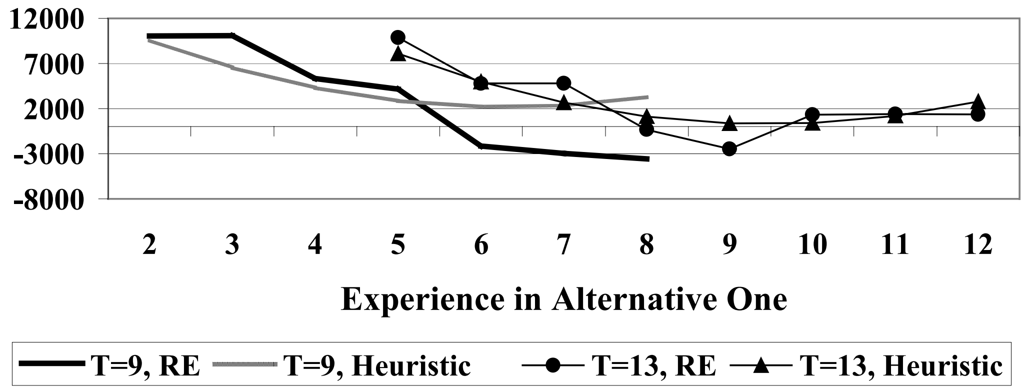
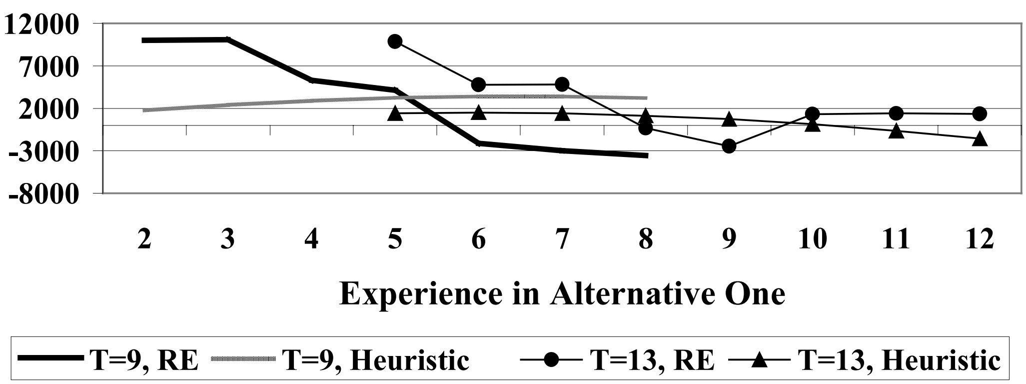


Figure 5B: Comparison of Fitted and Rational Expectations Future Component at Late State Vectors for Each Type

Near-Rational



Fatalist



Confused

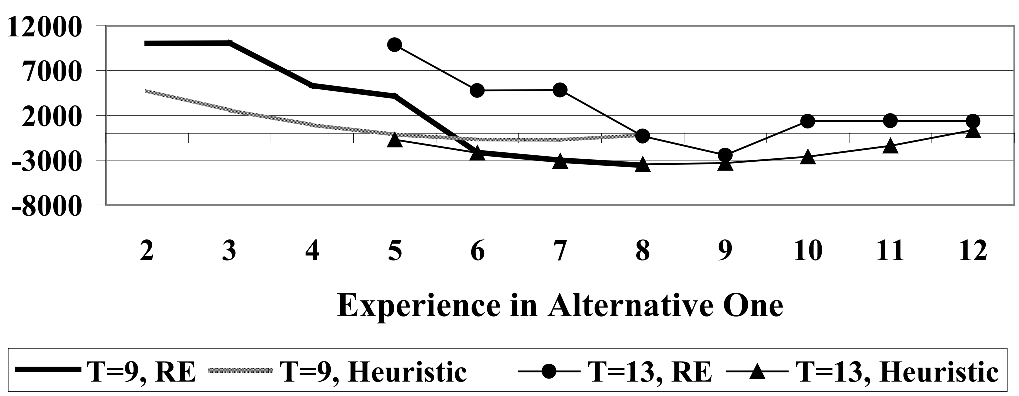


Figure 6: Average Mass Assigned to Each Type by True Type

Three Mixture Simulation

Five Mixture Simulation

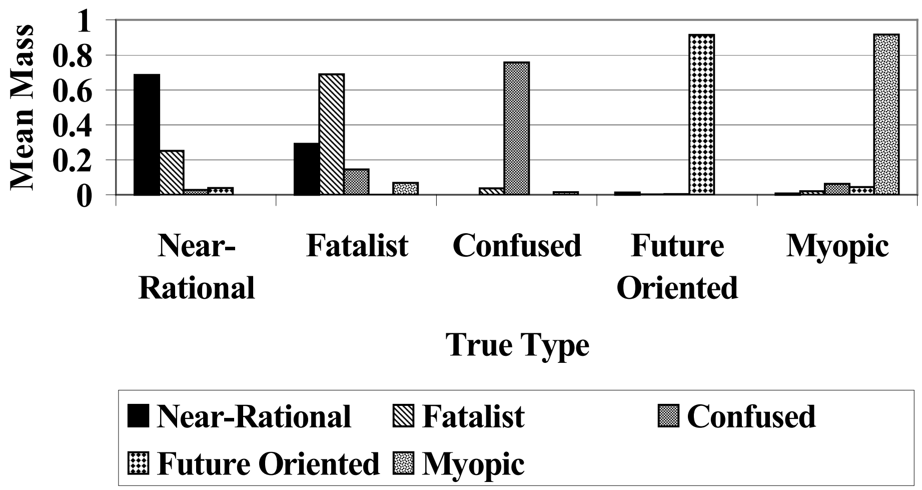


Figure 7: Empirical Cumulative Distribution Function of Total Number of Practice Games for Each Type

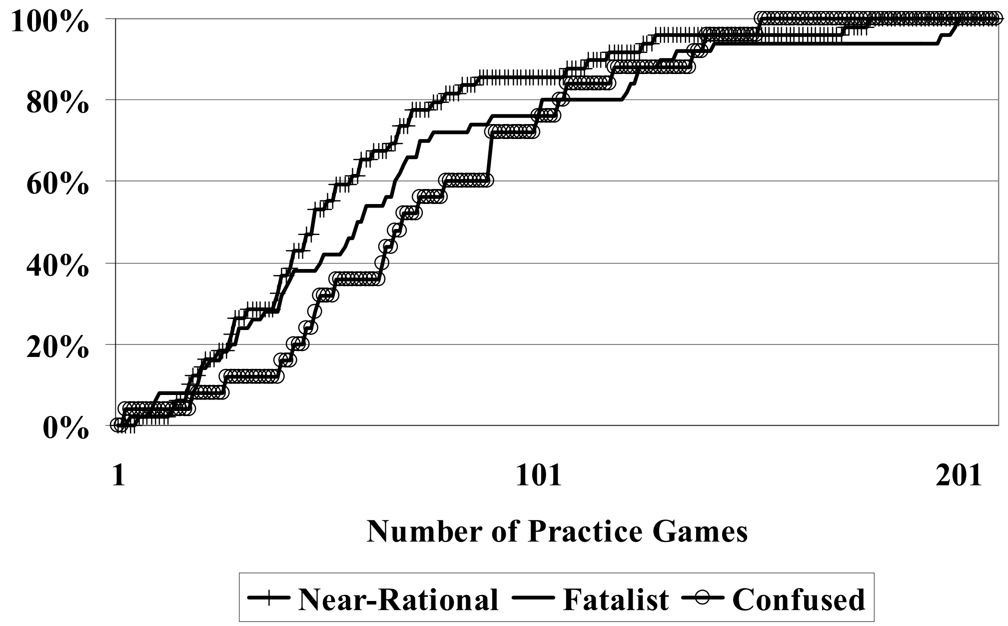


Figure B1: Example of a Subject's Game Screen