Be My Guinea Pig: Information Spillovers in a One-Armed Bandit Game*

John R. Boyce†
David M. Bruner‡
Michael McKee§

February 16, 2009

Abstract: This paper tests the Nash equilibrium predictions of a two-period, two-player one-armed bandit problem with information spillovers. Public information allows individuals to learn from their own private outcomes and/or from the other player's outcomes, creating an incentive to free-ride. The results from our experiment suggest players undervalue information, yet they also behave strategically based on the value they place on information. When only one player pulls the arm in the first period, it is the player predicted to do so at least 96% of the time. Furthermore, players respond to information in a manner consistent with Bayesian updating/reinforcement learning with deviations being attributable, in part, to salience and cognitive dissonance. Estimated regression coefficients are statistically significant and consistent with the summarized results.

Keywords: risk, uncertainty, sequential sampling, Bayesian, public goods, free-riding, experiments

JEL Classifications: C91,D81

---

*Acknowledgments: Funding for this was provided by a grant from the Social Sciences and Humanities Research Council of Canada. This research was undertaken at the University of Calgary Behavioral and Experimental Economics Laboratory (CBEEL). We would like to thank Diane Bischak, Rachel Croson, Curtis Eaton, and Francisco Gonzalez, for helpful comments and suggestions.

†Professor of Economics, Department of Economics, University of Calgary, 2500 University Drive, N.W., Calgary, Alberta, T2N 1N4, Canada. email: boyce@ucalgary.ca.

‡Assistant Professor of Economics, Appalachian State University, Boone, North Carolina, 28608 U.S.A. email: brunerdm@appstate.edu.

§Professor of Economics, Appalachian State University, Boone, North Carolina, 28608 U.S.A. email: mckeemj@appstate.edu.

1 Introduction

The most popular approach to modeling problems in the literature on strategic experimentation has been the armed-bandit problem. In the classic one-armed bandit problem a gambler faces a slot machine that pays a fixed return with an unknown probability. In each period, the gambler may choose either to play the slot machine, at a cost, or not. Each time the gambler plays he obtains additional information about the true probability that the machine pays. Traditionally, the gambler's problem is a tradeoff between experimentation and exploitation. The solution to the problem yields an optimal stopping rule for the gambler to quit playing the slot machine (Gittins, 1979).[1] In the game we investigate, others also observe the gamble and its outcome; hence there is a pure informational externality. In the presence of information spillovers, exploitation may be achieved through free-riding on others' experimentation.

A Bayesian agent will form an expected probability of winning from her prior belief and information, if any, provided from her own past draws as well as those draws she may observe of other agents. Taking a draw when the outcome will be public knowledge is akin to providing a public good—information. The Nash equilibrium in such public good settings is zero provision unless the information is a spillover from an action taken for private payoff, in which case, the provision level will be positive. The question is, do people actually behave according to the Nash equilibrium? If not, theoretical models of many different decision settings may fail to generate accurate predictions.

There are many decision settings in which such information spillovers may affect the level of risk taking and potential policy actions to increase the production of beneficial information by encouraging risk taking and/or compelling the public reporting of the outcomes. An (incomplete) list of such settings would include research and development activities, especially in settings when new information would inform others of the possibility of a bad outcome (Dixit and Pindyck, 1999); initial public offerings (IPOs) in which the informational spillover would inform others of the thickness of the security market—measured as the gap between

[1] The bandit problem owes its origins to (Robbins, 1952). In their survey, Bergemann and Välimäki (2006) describe the different faces that economists have put on the slot machine. Rothschild (1974) was the first to employ it, in the context of a firm experimenting with prices to learn about market demand. Weitzman (1979) and Roberts and Weitzman (1981) applied the bandit setup to the problem of investment in R&D. Jovanovic (1979) has used the bandit framework in a model of competitive labor markets where matching is important. Bergemann and Välimäki (1996) and Felli and Harris (1996) use the bandit problem to study the division of surplus under uncertainty. Bergemann and Hege (1998) and Bergemann and Hege (2005) apply the bandit problem to corporate finance decisions of venture and innovation. Caplin and Leahy (1998) investigate informational spillovers between firms regarding market demand. Sah (1991) and Lochner (2007) consider rational-cheater models of criminal behavior where agent's form beliefs about the enforcement regime based on their own experience (Lochner, 2007) or both their own experience and that of others (Sah, 1991). In addition, Aghion et al. (1991), Bolten and Harris (1999), and Keller et al. (2005) all investigate strategic experimentation in various bandit scenarios.

actual and forecast price of the IPO shares (Benveniste et al., 2003); entering new (e.g., geographic) markets such as new retail space (Caplin and Leahy, 1998); mineral exploration in a new region (Hendricks and Porter, 1996); and the introduction of new products such as the Apple ipod®. While there are insights to be gained from traditional empirical approaches, the laboratory offers the control and observability necessary to test the theoretical model. After all, it is important to understand the behavioral reactions to such information spillovers and identify the nature of the information provision if we are to inform the policy debate.

To this end, this is the first paper, to the best of our knowledge, to report the results of an experimental test of the Nash equilibrium predictions in a one-armed bandit problem with information spillover. We implement a two-player, two-period game in which players' have commonly known heterogeneous returns to the safe option. In each period, players simultaneously choose between a risky option and a safe option. Choosing the safe option in the first period yields a guaranteed return, but no information about the expected payoff to the risky option that might inform the second period decision. Information is valuable because the probability that the risky option pays out is unknown and a draw will update this probability. The free-rider problem arises because if *either* player chooses the risky action in the first period, the outcome of that choice is observable by the other player. The public good is the expected benefit information provides in making the subsequent decision. However, because the players' safe returns differ, their incentives for free-riding differ. It turns out that there is a range of safe returns for which players have a conditional best-response. This is because for players with safe returns in this range the expected value of the message service (i.e. the perceived future expected benefits of the information) when one player chooses the risky option is greater than the expected value of the message service when two players choose the risky option. When paired with a player whose safe return is below the lower bound on this range, a player can free-ride, as the other player is should choose the risky option (i.e. they have a dominate strategy to choose the risky option). On the other hand, when paired with a player whose safe return is above the upper bound on this range, no information should be provided by the other player (i.e. they have a dominate strategy to choose the safe option). Thus, for players with a conditional best-response, their incentive to free-ride depends on the safe return of the other player. Since the expected value of information provision is endogenously determined, and the game is dynamic, the problem we consider is more subtle than the standard public good problem.[2] Nonetheless, our results

[2]There have been numerous experimental tests of Nash equilibrium predictions of both provision and of free-riding in a public goods setting. Overwhelmingly the evidence suggests the people do not free-ride as often as theory predicts. Experiments

are consistent with previously reported findings in that players free-ride less than predicted when free-riding is a best response, and more than predicted when pulling the arm of the bandit is a best response.

Because the game is played for two periods, we can identify whether players can correctly deduce the dynamically optimal first period strategies and whether players' second period choices are consistent with Bayesian updating/reinforcement learning.[3] In order for players to choose the Nash best-response strategies in the first period of the game, they must satisfy three increasingly sophisticated cognitive requirements: (*i*) players must be able to formulate and update their beliefs using Bayes' rule;[4] (*ii*) players must be forward-looking, so that they recognize the value of information; (*iii*) players must also be strategically rational in order to recognize when the incentive to free ride is a dominant strategy and when it is not. In order for a player to be strategically rational, it must be that information has value and in order for information to be of value, it must be that the player updates her beliefs. If a player fails to satisfy any of these requirements, her behavior will deviate from the theoretical prediction.[5]

Choices made by players in the second period of the game allow us to identify whether or not players utilize information in manner consistent with the model. Consistent behavior in the second period of the game lends support to the notion that players understood the value of information in the first period of the game. While there is considerable published evidence that players are not "perfect Bayesians"—people place too little weight on prior beliefs and too much weight on new information—our experiment is designed to afford players the best shot at being "good Bayesians." For example, when there are a small number of possible data generating processes, players are able to form their posterior beliefs on heuristics such as whether the sample is "representative" of one of the possibilities (e.g. El-Gamal and Grether (1995); Grether (1980, 1992); Kahneman and Tversky (1972, 1973); Tversky and Kahneman (1971, 1973)). Because we implicitly allow for a very large number of possible data generating processes, we minimize the influence of prior beliefs so that a heuristic such as representativeness has little explanatory power. Furthermore, we employ a very simple

routinely report that subjects contribute to the public good despite this being a dominated strategy. Explanations for such behavior have typically taken the form of either warm-glow (i.e. the satisfaction from giving) or altruism (i.e. interdependent utility functions).

[3] Charness and Levin (2005) tested Bayesian predictions against those of reinforcement learning. Their results suggest that both heuristics are used, and when the predictions of the two are aligned, as in our experiment, people respond as expected.

[4] In the absence of Bayesian updating the behavior in the one-armed bandit setting is degenerate. Classical statistical decision theory would imply that players always take two draws to maximize the size of the sample. Such myopic behavior is inconsistent with a decision setting in which information from prior draws has value.

[5] There is mixed evidence that people understand the value of information as modeled in the economic literature. Kraemer et al. (2006) investigate the extent to which people rationally acquire costly information (i.e. understand the value of information). There results suggest that people overvalue information (i.e. they purchase too much information). However, earlier experiments provide a body of evidence that suggests the people undervalue information (McKelvey and Page, 1990; Meyer and Shi, 1995; Banks et al., 1997; Anderson, 2001; Gans et al., 2007).

case of Bayesian updating, a binomial sample, in which the sample proportion asymptotically approaches the no information prior Bayesian estimate. Despite these simplifications to the decision setting, we find evidence that when players make a choice in the first period of play, they are more likely to repeat that choice in the second period. Thus, like Charness and Levin (2005), we find that players who make errors in the second period do so when information is in dissonance with their beliefs.[6] While our first period error rates indicate that players undervalue information, they respond to information in a manner consistent with Bayesian updating/reinforcement learning, suggesting they place some value on information. Deviations from Bayesian predictions appear to be attributable, in part, to salience and cognitive dissonance. Players for whom the cost of errors is high are less likely to commit errors. Indeed, when only one player pulls the arm, it is the player predicted to do so at least 96% of the time. Therefore, our results suggest that players are somewhat myopic, yet are forward-looking "enough" to behave strategically.

As we have noted, when each player can observe the other's actions and results, information is a classic public good and the gambler may have an incentive to free-ride on the play of another.[7] It could be argued that there are other effects at work. For example, in games with three or more periods and publicly observable outcomes, Bolten and Harris (1999) show that players face two opposing incentives: an incentive to free-ride on information provision by others and an incentive to provide information in the current period to encourage information provision by others in the future. Like Bolten and Harris (1999), we assume that both the choice of and the result of a draw is perfectly observable. However, by restricting our attention to a two-period setting, we eliminate the encouragement effect; information gathered in the second period of the one-shot game cannot affect future play. This allows us to focus our attention solely on the free-riding effect. Thus, unlike Gans et al. (2007), who focus on how players form heuristics to simplify the repeated sampling problem in infinite time bandit problems, the complexity of the problem our players face is lessened.

[6]While previous results suggest people place too little weight on priors, our observation of cognitive dissonance suggests the contrary. We believe there is plausible explanation for the apparent discrepancy. Previous experiments testing Bayes rule versus other heuristics explicitly induced prior beliefs, provided new information, and then ask the subject to make a guess about the source of the information. However, players do not have to use information in our experiment (they can simply choose the safe option in the second period of the game). Indeed, some players *never* choose the risky option in the second period of certain treatments (see footnote 15). This could explain why players seem to ignore information in our experiment.

[7]The informational externality considered here differs from the literature concerning "herding" and "information cascades." These latter models assume players have private information regarding the true state of nature and observation of other players' actions causes inference about that private information. As in Bolten and Harris (1999), our model allows players to form beliefs about the true state of nature through observation of other players' outcomes.

2 Information Spillovers in a Bandit Game

Two risk neutral player's, $i = 1, 2$, play for two periods, $t = 0, 1$. In each period the players simultaneously choose whether or not to take the safe option, paying S_i, or to take the risky option, a sample of n Bernoulli trials, where each trial pays R for a success and zero for a failure. The results of the period one draws, if any, become public information prior to period two. The payoffs to each player are the undiscounted sum of first and second period returns.

2.1 Equilibrium in the Two-Person, two-period One-Armed Bandit Game

The probability of a success in each draw from the risky option is θ, where $0 < \theta < 1$, but θ is unknown to the players. Since the probability of success is unknown, players must make their decisions in each period based upon their subjective beliefs. Let $\xi(\theta)$ denote the distribution of each player's prior beliefs about θ. Assume that $\xi(\theta)$ is a beta distribution with parameters α and β, so that $\xi(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}$ (DeGroot, 1970, p. 40). A draw of N samples from the risky option results in a sample of outcomes, $x_1, \ldots, x_N$, each of which is distributed as a Bernoulli random variable, so that the likelihood we observe X successes in N trials is given by $f_N(x_1, \ldots, x_N|\theta) \propto \theta^X(1-\theta)^{N-X}$, where $X = \sum_{i=1}^{N} x_i$ is the number of observed successes in N trials. As each player may draw n samples, $N \in \{0, n, 2n\}$ denotes the number of draws in the first period. Zero draws occur when neither player selects the risky choice; n draws occur when one of the players selects the risky choice and the other the safe choice; and $2n$ draws occur when both players select the risky choice. The posterior distribution of beliefs is given by $\xi(\theta|x_1, \ldots, x_N) \propto \theta^{\alpha+X-1}(1-\theta)^{\beta+N-X-1}$, which is a beta distribution with parameters $\alpha + X$ and $\beta + N - X$ (DeGroot, 1970, p. 160, Theorem 1). Thus, the information available to players at the beginning of period two is the number of draws, N, and the number of successes, X. Let $I = \{N, X\}$ denote the information set. The possible information sets after one period are $I_0 = \{\emptyset\}$, $I_1(X) = \{n, X\}$, $I_2(X) = \{2n, X\}$, where the subscript on the information set represents the number of players who have chosen the risky option. Let $p_t = E[\theta|I]$ denote each player's common expectation of θ in period t, given the information, I, available at that time. A beta distribution with parameters α and β has expectation $(\alpha + X)/(\alpha + \beta + N)$. Thus $p_0 = \frac{\alpha}{\alpha+\beta}$ and $p_1(I) = \frac{\alpha+X}{\alpha+\beta+N}$. Since players are completely uninformed about the value of θ, it is natural to assume that each player has uninformative prior beliefs about the probability of success. Therefore, we assume that $\alpha = \beta = 1$, which implies that $\xi(\theta)$ is a uniform distribution over the interval zero to one. Hence, the expectation of the uninformed prior is $p_0 = \frac{1}{2}$ and the expectation of the posterior distribution is $p_1 = \frac{1+X}{2+N}$.

Given players are assumed to be risk neutral, the current period subjective expected utility (EU_{it}) to each player i from choosing the risky option is simply the expected value from the risky option, which is given by

$$EU_{it} = R \sum_{X=0}^{n} f_n(X|I)X = nRp_t(I). \tag{1}$$

In the Nash equilibrium, each player solves the dynamic game using backward induction. Player i's decision rule in the second period is to choose the risky option if, and only if:

$$nRp_1(I) > S_i. \tag{2}$$

Thus, a risk-neutral player chooses the risky option if, and only if, the return to the safe option, S_i, is less than the expected return to the risky option, $nRp_1(I)$, given the expectation about θ given by the posterior distribution as informed by the information set I. Therefore, we can write the second period expected payoffs as $\max\big[S_i, Rnp_1(I)\big]$. For a Bayesian expected utility player, there is a critical number of successful draws, $X_N^{S_i}$, that depends upon the opportunity cost to the player and the number of draws observed, such that if $X > X_N^{S_i}$ the player is induced to choose the risky option. From the posterior expectations, those values satisfy $X_N^{S_i} = \frac{S_i(2+N)}{nR} - 1$ for $N \in \{n, 2n\}$.

The game in period one is complicated by both strategic and information concerns. The normal form of the first period game is depicted in Figure 1. Player j's expected payoffs are depicted in the upper-right of each cell and player i's expected payoffs are depicted in the lower left of each cell. When neither player chooses the risky option, the second period payoffs are simply $\max[S_i, Rnp_0]$, since no draws from the risky option have been observed. When one or both of the players chooses the risky option, however, the terms in summation give the decision rule for each possible information set weighted by the probability of observing that information set given the prior beliefs.

From Figure 1, we may derive the decision rule for player i to choose the risky option when player j has chosen the safe option. That rule is to choose the risky option if, and only if the following inequality holds:

$$Rnp_0 - S_i + \sum_{X=0}^{n} f_n(X|I_0)\max\Big[S_i, Rnp_1\big(I_1(X)\big)\Big] - \max\Big[S_i, Rnp_1(I_0)\Big] > 0. \tag{3}$$

		Player j	
		Safe	Risky
Player i	Safe	$S_j + \max\Big[S_j, Rnp_1(I_0)\Big]$ $S_i + \max\Big[S_i, Rnp_1(I_0)\Big]$	$Rnp_0 + \sum_{X=0}^{n} f_n(X\|I_0)\max\Big[S_j, Rnp_1\big(I_1(X)\big)\Big]$ $S_i + \sum_{X=0}^{n} f_n(X\|I_0)\max\Big[S_i, Rnp_1\big(I_1(X)\big)\Big]$
	Risky	$S_j + \sum_{X=0}^{n} f_n(X\|I_0)\max\Big[S_j, Rnp_1\big(I_1(X)\big)\Big]$ $Rnp_0 + \sum_{X=0}^{n} f_n(X\|I_0)\max\Big[S_i, Rnp_1\big(I_1(X)\big)\Big]$	$Rnp_0 + \sum_{X=0}^{2n} f_{2n}(X\|I_0)\max\Big[S_j, Rnp_1\big(I_2(X)\big)\Big]$ $Rnp_0 + \sum_{X=0}^{2n} f_{2n}(X\|I_0)\max\Big[S_i, Rnp_1\big(I_2(X)\big)\Big]$

Figure 1: Normal Form of the First Period Game

The difference in the first two terms in (3) is the expected opportunity cost of obtaining information in the first period. This equals the expected return from choosing the risky option less the forgone return to choosing the safe option. The difference in the last two terms in (3) is the expected value of the message service (Hirshleifer and Riley, 1992, p. 180). It is well known that for any given return to the safe option relative to the risky option, the expected value of the message service is non-negative. This can be shown by writing the expected value of the message service as

$$EVMS_{S_i}(n) = \sum_{X=0}^{n} f_n(X|I_0)\left\{\max\Big[S_i, Rnp_1(I_1(X))\Big] - \max\Big[S_i, Rnp_1(I_0)\Big]\right\} \geq 0. \tag{4}$$

Information has value because it may change the decision in the second period. The expected value is the difference between the weighted average of expected returns from the optimal decision, $\max\Big[S_i, Rnp_1\big(I_1(X)\big)\Big]$, weighted by the expected probability of the particular information signal, $f_n(X|I_0)$, and the expected return from the optimal decision in the absence of information, $\max\big[S_i, Rnp_1(I_0)\big]$. This difference is greater than zero whenever the player chooses an action that differs from what he would have chosen absent information. Since the opportunity cost of obtaining information is positive for all values of $S_i > Rnp_1(I_0)$, and negative for $S_i < Rnp_1(I_0)$, it is a best-response for a player with a safe return $S_i \leq Rnp_0$ to choose the risky option whenever the other player chooses the safe option.

In contrast, if player j has chosen the risky choice in period one, player i's decision is no longer a comparison between zero information and a positive amount of information. Rather, player i's choice is now between two message services, one in which a n draws are provided and one in which $2n$ draws are provided. When player j has chosen the risky option, player

i's decision rule is to choose the risky option if, and only if,

$$Rnp_0 - S_i + \sum_{X=0}^{2n} f_{2n}(X|I_0)\max\Big[S_i, Rnp_1\big(I_2(X)\big)\Big] - \sum_{X=0}^{n} f_n(X|I_0)\max\Big[S_i, Rnp_1\big(I_1(X)\big)\Big] > 0. \tag{5}$$

The difference in the first two terms in (5) is the expected opportunity cost of obtaining the message service that provides $2n$ draws in the first period rather than the message service that provides only n draws, measured in the safe return foregone. This cost is identical to the case where the other player does not make the risky choice. The expected value of the message service that provides $2n$ draws relative to the message service that provides n draws is captured by the difference between the two summations in (5). This equals the weighted average of expected returns the player expects to receive from the second period optimal decision, given that both players have chosen the risky option, less the weighted average of expected returns the player expects to receive from the second period optimal decision, when only the other player has chosen the risky option. We now show that the net expected value of the message service that provides $2n$ draws relative to the message service that provides n draws is negative for some players.

2.2 Nash Equilibrium Illustrated when $n = 1$

To gain some intuition about how players are predicted to play, consider the case where $n = 1$. Suppose that player j has chosen the safe option. If player i chooses the safe option, his expected return is

$$EU_i(\text{Safe}|\text{Safe}) = S_i + \max\Big[S_i, \tfrac{1}{2}R\Big], \tag{6}$$

and if player i chooses the risky option, his expected return is

$$EU_i(\text{Risky}|\text{Safe}) = \tfrac{1}{2}R + \tfrac{1}{2}\max\Big[S_i, \tfrac{1}{3}R\Big] + \tfrac{1}{2}\max\Big[S_i, \tfrac{2}{3}R\Big], \tag{7}$$

where the fractions outside the max functions are the probability of the information set given the uninformed prior, and the fractions inside the max functions are the posterior expectations of a success. Equating the right-hand sides of (6) and (7) and solving for the safe return value, $\bar{S}$, such that for $S_i < \bar{S}$, the expected return from choosing the risky option exceeds the expected return from the safe option, yields $\bar{S} = \frac{5}{9}R$. For all players for whom $S_i < \frac{5}{9}R$, the best-response to the other player choosing the safe option is to choose the risky option, and for all players for whom $S_i \geq \frac{5}{9}R$, the best response to player j choosing the safe option is to choose the safe option.

Suppose instead that player j has chosen the risky option. Then the expected return to player i of choosing the safe option is

$$EU_i(\text{Safe}|\text{Risky}) = S_i + \tfrac{1}{2}\max\Big[S_i, \tfrac{1}{3}R\Big] + \tfrac{1}{2}\max\Big[S_i, \tfrac{2}{3}R\Big]. \tag{8}$$

In contrast, if player i chooses the risky option, his expected return is

$$EU_i(\text{Risky}|\text{Risky}) = \tfrac{1}{2}R + \tfrac{1}{4}\max\Big[S_i, \tfrac{1}{4}R\Big] + \tfrac{1}{2}\max\Big[S_i, \tfrac{1}{2}R\Big] + \tfrac{1}{4}\max\Big[S_i, \tfrac{3}{4}R\Big]. \tag{9}$$

Solving for the value of the safe return, $\underline{S}$, such that for $S_i \leq \underline{S}$ the expected return from choosing the risky option exceeds the expected return from choosing the safe option, yields $\underline{S} = \frac{29}{60}R$. Thus, all players for whom $S < \frac{29}{60}R$, the best-response to the other player choosing the risky option is to also choose the risky option. This also means that all players for whom $S_i \geq \frac{29}{60}R$, the best response to player j choosing the risky option is to choose the safe option.

Therefore, players whose safe return is less than $\underline{S}$ have a dominant strategy of choosing the risky option and players whose safe return is greater than $\bar{S}$ have a dominant strategy of choosing the safe option. Players for whom $\underline{S} < S_i < \bar{S}$, however, have a conditional best-response; to choose risky if the other player chooses safe and to choose safe if the other player chooses risky. This is demonstrated by plotting the expected second period return as a function of the value of S_i relative to R as shown in Figure 2 (where the vertical scale has been truncated). The thick solid chord with a single kink at $S_i = \frac{1}{2}R$ is the expected return of the second period optimal decision when neither player has chosen the risky option in the first period ($N = 0$). The dashed chord with kinks at $S_i = \frac{1}{3}R$ and $S_i = \frac{2}{3}R$ is the expected return of the second period optimal decision when one player has chosen the risky option in the first period ($N = 1$). The thin solid chord with kinks at $S_i = \frac{1}{4}R$, $S_i = \frac{1}{2}R$, and $S_i = \frac{3}{4}R$ is the expected return of the second period optimal decision when both players have chosen the risky option in the first period ($N = 2$). The vertical distance between the thin solid chord and the thick solid chord is the expected value of the message service when one player has chosen the risky option in the first period. The vertical distance between the dashed chord and the thick solid chord is the expected value of the message service when both players have chosen the risky option in the first period. Between $\underline{S}$ and $\bar{S}$, the expected value of the message service associated with two draws is less than the expected value of the message service associated with one draw. This difference is maximized at $S_i = \frac{1}{2}R$. The two horizontal dotted chords show the expected value of the message services

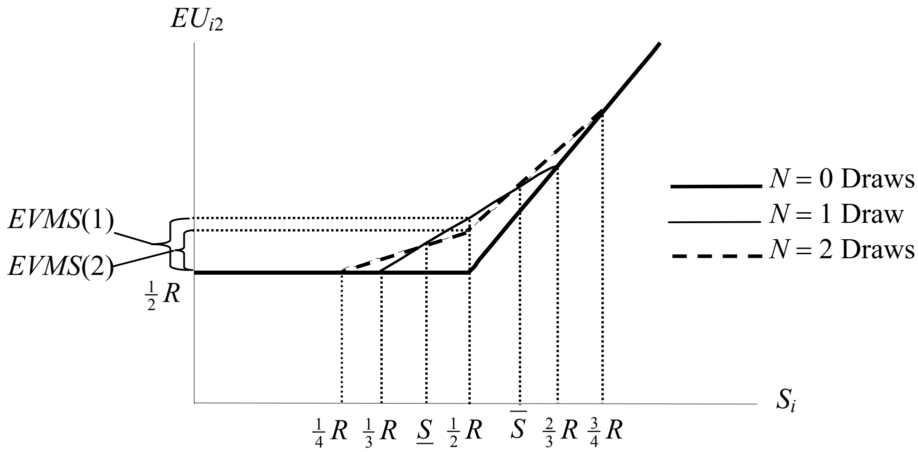

Figure 2: Expected Value of Message Services with n Draws and $2n$ Draws, $n = 1$.

for the player for whom $S_i = \frac{1}{2}R$, for the cases of one draw and two draws, respectively. For the player for whom $S_i = \frac{1}{2}R$, the expected value of the message service associated with one draw is equal to $EVMS_{\frac{R}{2}}(1) = \frac{1}{12}$; the expected value of the message service associated with two draws is equal to $EVMS_{\frac{R}{2}}(2) = \frac{1}{16}$. The difference for this player is $EVMS_{\frac{R}{2}}(2) - EVMS_{\frac{R}{2}}(1) = \frac{1}{16} - \frac{1}{12} = -\frac{1}{48}$.

With one draw, relative to his uninformative prior of $p_0 = \frac{1}{2}$, the player's posterior expectation moves to either $p_1 = \frac{1}{3}$ with probability $\frac{1}{2}$ or to $p_1 = \frac{2}{3}$ with probability $\frac{1}{2}$. With two draws, given his uninformative prior, the player obtains a posterior expectation of $p_1 = \frac{1}{4}$ with probability $\frac{1}{4}$, a posterior expectation of $p_1 = \frac{3}{4}$ with probability $\frac{1}{4}$, or a posterior expectation of $p_1 = \frac{1}{2}$ with probability $\frac{1}{2}$. Therefore, with two draws, it is equally likely that the information from the two draws will cancel each other out, in which case the player is left with no more information than from his prior, of that the two draws will reinforce each other, in which case the player is more informed than with his prior. In contrast, with one draw, the player's posterior expectation always differs from his prior expectation. A risk neutral Bayesian player only cares about the expected value of the lottery relative to the safe return. Thus, for a player for whom $\underline{S} < S_i < \bar{S}$, taking two draws is expected to be less valuable in making his subsequent second period decision than is taking only one draw.

2.3 Solution to the Game Used in the Experiment

In the experiment the payoff to a successful draw is $R = \$5$, and choosing the risky option gives the player $n = 3$ draws. There are three types of players, indexed by their safe returns: $S_L = \$4$, $S_M = \$8$, and $S_H = \$12$. Thus we can obtain exact values for the expectations, assuming a no information prior, depending upon the value of S_i each player faces. In the experiment, $S_L = \$4$ and $S_H = \$12$ players are always paired with $S_M = \$8$ player. Figure 3 displays the normal form of the first period games used in the experiment. The numbers in Figure 3 represent the expected (undiscounted) present value stream of payoffs from playing a particular strategy, given optimal second period behavior and an uninformative prior. The $S_M = \$8$ player's payoffs are the first number in each cell and best-responses are in bold-face font. There is one pure-strategy Nash equilibrium in each of the two games in Figure 3. In the Nash equilibrium for the $\{\$4, \$8\}$ game, the $S_M = \$8$ player chooses the safe option and the $S_L = \$4$ player chooses the risky option. In the Nash equilibrium for the $\{\$12, \$8\}$ game, however, the $S_M = \$8$ player chooses the risky option and the $S_H = \$12$ player chooses the safe option.

Observe also that for players of types $S_H = \$12$ and $S_L = \$4$ the cost of decision error is much higher relative to the cost of decision error for the $S_M = \$8$ player type (Smith and Walker, 1993). For any player for whom $\underline{S} < S_i < \bar{S}$, the cost of decision error is relatively low, since the expected value of the choosing the risky option is very close to the return from the safe option. Unfortunately, this is also the type of player who is the most interesting because he does not have a dominant strategy. This feature of the problem is dictated by the underlying the one-armed bandit game.[8]

3 Experimental Design and Hypotheses

The experiment implements the two period game described above. Each player participates in twenty rounds of play in a within-subjects design. In each period of each round, players choose between a guaranteed amount, predetermined to be one of $S_i = \{\$4, \$8, \$12\}$, and a lottery with an unknown probability distribution. The lottery is framed as three draws with replacement from an urn containing 100 balls, composed of an unknown proportion θ of red balls and proportion $1-\theta$ of blue balls, where $0 \leq \theta \leq 1$. Each red ball pays \$5 and each blue ball pays \$0. Players are informed that θ is held constant in both periods of a round

[8] The small difference in the payoff space implies that we are more likely to err in the direction of rejecting Nash equilibrium predictions.

$S_i = \$8$ (rows), $S_i = \$4$ (columns)

	Safe	Risky
Safe	$\frac{128}{8}$ $\frac{92}{8}$	$\mathbf{\frac{4320}{256}}$ $\mathbf{\frac{121}{8}}$
Risky	$\mathbf{\frac{131}{8}}$ $\frac{5952}{512}$	$\frac{4149}{256}$ $\mathbf{\frac{7709}{512}}$

(a) $\{\$4, \$8\}$ Game

$S_i = \$8$ (rows), $S_i = \$12$ (columns)

	Safe	Risky
Safe	$\frac{128}{8}$ $\mathbf{\frac{48}{2}}$	$\mathbf{\frac{4320}{256}}$ $\frac{39}{2}$
Risky	$\mathbf{\frac{131}{8}}$ $\mathbf{\frac{12288}{512}}$	$\frac{4149}{256}$ $\frac{9993}{512}$

(b) $\{\$12, \$8\}$ Game

Figure 3: First Period Game Expected Utilities

of the experiment; but varies across rounds. Thus, information obtained in the first period could inform the decision the player is to make in the second period; however, information has no value from one round of the experiment to the next.

Players are randomly assigned to one of the safe payoff values in each round of a session. The model predicts that the $S_i = \$8$ player type will systematically vary his behavior depending upon which type of player he is paired with, $S_i = \$4$ or $S_i = \$12$ player types. Thus, an $S_i = \$8$ player type is always paired with either an $S_i = \$4$ player type or an $S_i = \$12$ player type. Therefore, in any round of play 25% of players are type $S_i = \$4$, 50% are type $S_i = \$8$, and 25% are type $S_i = \$12$. This experimental design yields four treatments consisting of the combinations of a player's own type and the type of the other player $\{S_i, S_j\} = \{\{\$4, \$8\}, \{\$8, \$4\}, \{\$8, \$12\}, \{\$12, \$8\}\}$. Each treatment pair is assigned a probability of success from a distribution with mean 0.5 in accordance with the non-informative prior. The player types, player pairing, and underlying probability of success are drawn randomly in each round of play, and remain fixed for that round.[9] Treatments are drawn with replacement. Player pairing and the underlying probability of success in each round are drawn without replacement. Players never learn the identity of their partner in any round, but in each round they know their own and their partner's type.

In all sessions the instructions are read aloud, as well as presented on computer screens, to ensure common knowledge.[10] Participants are required to pass a basic test of statistical skills to reduce errors due to misunderstanding.[11] A total of 52 players have participated (18 in session 1, 20 in session 2, and 14 in session 3). Since each player makes choices in each of

[9]The motivation for randomizing the player types is to facilitate the backward induction process. That is, actually playing as another type should enable a player to deduce the optimal strategy for that type.

[10]Screen images are available upon request from the authors.

[11]Participants could not proceed until they answered all questions correctly. The questions are provided in the screen images.

two periods of twenty rounds of play, there are 2080 total observed risky/safe choices. The experiments were conducted at the University of XXXX Laboratory. The participant pool is composed of volunteer students at the university. Participants are recruited by email via the lab's Online Recruitment System for Experimental Economics (ORSEE) (Greiner, 2004). The experiment is programmed and conducted with the software Z-Tree (Fischbacher, 2007). Players are paid on the basis of the outcome of one randomly chosen round. Experimental sessions last approximately 90 minutes, and participant earnings average \$23, of which \$5 was a show-up fee.

3.1 Hypotheses Regarding First Period Choices

We present the equilibrium predictions for the first period choices implied by the Bayesian behavior, and derive hypotheses about the deviations from those predictions. Let ΔEU^1_{ijt} denote the latent undiscounted net gain in the sum of first and second period subjective expected utility to a player of type i playing against a player of type j from choosing the risky option in round t. Then the probability that the player chooses the risky option in the first period is given by

$$Pr(\Delta EU^1_{ijt} > 0) = \beta^{\mathbf{1}}_{\mathbf{ij}}\mathbf{D_{ij}} + \beta^{\mathbf{1}}_{\mathbf{t}}\mathbf{R_t} + \beta^{\mathbf{1}}_{\mathbf{P}}\mathbf{P_i} + \epsilon^1_{ijt}, \tag{10}$$

where $\mathbf{D_{ij}}$ is a vector of treatment dummy variables where i denotes a player's own payoff from the safe choice and j denotes the partner's payoff from the safe choice, $\mathbf{R_t}$ is a vector of round fixed-effects, $\mathbf{P_i}$ is a vector of player fixed-effects, and ϵ^1_{ijt} is the unobserved error.[12]

The $\beta^{\mathbf{1}}_{\mathbf{ij}}$ parameters correspond to the (conditional) mean probabilities that each type chooses the risky option in the first period. Under the null hypothesis that the round and player effects are zero, these correspond to the unconditional mean probabilities that each type chooses the risky option in period one. The Nash equilibrium to the game yields a set of behavioral predictions that strictly follow the theory. These are that first period behavior satisfies $\beta^1_{4,8} = \beta^1_{8,12} = 1$ and $\beta^1_{8,4} = \beta^1_{8,12} = 0$. If players err, their errors are necessarily one sided: those players who should play the risky (safe) choice with probability one can only err by choosing the risky (safe) choice with probability less than one. The following set of behavioral hypotheses takes player error into account:

Hypothesis 1. $\beta^1_{4,8} > \beta^1_{12,8}$: *players in treatment* $\{\$4, \$8\}$ *should choose the risky option more often than players in treatment* $\{\$12, \$8\}$*, in the first period.*

[12]One of the strengths of a within-subjects design is that the player fixed-effects should be uncorrelated with the treatments, allowing for a cleaner test of treatment effects.

Hypothesis 2. $\beta^1_{8,12} > \beta^1_{8,4}$: *players in treatment* $\{\$8, \$12\}$ *should choose the risky option more often than players in treatment* $\{\$8, \$4\}$*, in the first period.*

To achieve the Nash predictions, players must satisfy three requirements: (i) they must be Bayesian, (ii) they must recognize the value of information, and (iii) they must behave strategically. Let ϵ^1_{ij} denote the average error by treatment type. Player error can be caused by several factors. Thus, deviations from theoretical predictions in the first period could be due to *myopic* behavior (undervaluation of information) or a failure to be *strategically rational.* The two types of violations, however, do not produce symmetric errors, as hypotheses 3 and 4 relate:

Hypothesis 3. $\epsilon^1_{8,12} \equiv 1 - \beta^1_{8,12} > \beta^1_{8,4} - 0 \equiv \epsilon^1_{8,4}$: *in the first period, myopic players in treatment* $\{\$8, \$12\}$ *should exhibit a higher error rate by choosing the safe option more often than players in treatment* $\{\$8, \$4\}$ *choose the risky option.*

Hypothesis 4. $\epsilon^1_{8,12} \equiv 1 - \beta^1_{8,12} < \beta^1_{8,4} - 0 \equiv \epsilon^1_{8,4}$: *in the first period, strategically irrational players in treatment* $\{\$8, \$12\}$ *should exhibit a lower error rate by choosing the safe option less often than players in treatment* $\{\$8, \$4\}$ *choose the risky option.*

If players are myopic, they will choose the guaranteed amount in both periods of the $\{\$12, \$8\}$, $\{\$8, \$12\}$, and $\{\$8, \$4\}$ treatments. Thus, we expect myopia to produce a higher error rate in the first period of the $\{\$8, \$12\}$ treatment relative to the $\{\$8, \$4\}$ treatment.[13] On the other hand, if players understand the expected value of information but fail to recognize when the other player will provide it, they should choose the risky option in the first period of the $\{\$8, \$12\}$ and $\{\$8, \$4\}$ treatments. These players recognize the expected value of information but fail to be strategically rational. Thus, we expect strategic irrationality to produce a higher error rate in the $\{\$8, \$4\}$ treatment than in the $\{\$8, \$12\}$ treatment.[14] However, it is clear that Hypotheses 3 and 4 are mutually exclusive: Hypothesis 3 implies $1 < \beta^1_{8,12} + \beta^1_{8,4}$ while Hypothesis 4 implies that $1 > \beta^1_{8,12} + \beta^1_{8,4}$.

Another potential cause of player error is lack of payoff salience (Smith and Walker, 1993). This is an issue for any experimental design — a failure to implement a basic precept of experimental methodology (Smith, 1982). By virtue of our design, players in treatments $\{\$4, \$8\}$ and $\{\$12, \$8\}$ each have much more to gain than a player of type $\{\$8, \$4\}$ and

[13] Myopia causes a bias towards the safe option, which simultaneously raises the error rate in the $\{\$8, \$12\}$ treatment and lowers the error rate in the $\{\$8, \$4\}$ treatment.

[14] Strategic irrationality causes a bias toward choosing the risky option, which simultaneously raises the error rate in the $\{\$8, \$4\}$ treatment and lowers the error rate in the $\{\$8, \$12\}$ treatment.

$\{\$8, \$12\}$ from making the correct choice, hence their decision errors should be smaller.[15] The mean decision error for players in treatment $\{\$4, \$8\}$ is $\epsilon^1_{4,8} = 1 - \beta^1_{4,8}$ and the mean decision error for players in treatment $\{\$8, \$12\}$ is $\epsilon^1_{8,12} = 1 - \beta^1_{8,12}$. Thus, salience implies that $1-\beta^1_{4,8} < 1-\beta^1_{8,12}$, or that $\beta^1_{8,12} < \beta^1_{4,8}$. Similarly, while players in treatments $\{\$12, \$8\}$ and $\{\$8, \$4\}$ are each expected to choose the safe option, the salience is higher for a player in treatment $\{\$12, \$8\}$, thus we expect that $\epsilon^1_{12,8} = \beta^1_{12,8} - 0 < \beta^1_{8,4} - 0 = \epsilon^1_{8,4}$. The design allows us to investigate the contribution of salience to decision errors. Specifically:

Hypothesis 5. $\epsilon^1_{4,8} \equiv 1 - \beta^1_{4,8} < 1 - \beta^1_{8,12} \equiv \epsilon^1_{8,12}$*; salience implies that in period one players in treatment* $\{\$4, \$8\}$ *should have lower errors rates than players in treatment* $\{\$8, \$12\}$*.*

Hypothesis 6. $\epsilon^1_{12,8} \equiv \beta^1_{12,8} < \beta^1_{8,4} \equiv \epsilon^1_{8,4}$*; salience implies that in period one players in treatment* $\{\$12, \$8\}$ *should have lower errors rates than players in treatment* $\{\$8, \$4\}$*.*

3.2 Hypotheses Regarding Second Period Choices

The Bayesian predictions are that players choose the risky option in the second period when the inequality in (4) is satisfied, i.e., that a player of type S_i chooses the risky option whenever $X \geq X_N^{S_i}$. When $N = 3$, $X_3^{\$4} = 1$, $X_3^{\$8} = 2$, and $X_3^{\$12} = 3$; and when $N = 6$, $X_6^{\$4} = 2$, $X_6^{\$8} = 4$, and $X_6^{\$12} = 6$. This implies that the second period choice does not depend upon the first period choice, except in so far as that choice determines the number of draws. It also implies that what matters is whether or not X is greater or less than $X_N^{S_i}$; it does not matter by how much the two numbers differ.

While Bayesian updating implies that all that is important is the sign of $X - X_N^{S_i}$, salience implies that the larger is the difference between X and $X_N^{S_i}$, the larger will be the difference in expected utility. This is measured by including a variable equal to the difference $X - X_N^{S_i}$ in a regression equation explaining second period choices. Salience implies that the coefficient on this variable will be positive in sign.

The first period choice made by a player could matter if the player suffers cognitive dissonance (Akerlof and Dickens, 1982). Cognitive dissonance is a psychological condition induced when an individual is confronted with information that contradicts their prior belief. This results in placing too much weight on prior beliefs in the updating process. In our experiment, cognitive dissonance implies that players will be more likely to repeat their first

[15] The percentage difference in payoffs between choosing the best response and the alternate strategy in Figure 3 is at most 4% for a player in treatments $\{\$8, \$4\}$ and $\{\$8, \$12\}$, while the percentage difference between choosing the best response and the alternate strategy is at least 23% for a player in treatment $\{\$12, \$8\}$ and at least 29% for a player in treatment $\{\$4, \$8\}$.

period choices regardless of the outcome of that choice.[16] Thus, the player is more likely to choose the risky option in period two if she chose the risky option in period one and to choose the safe option in period two if she chose the safe option in period one. In a regression model in which second period choices are coded as a one if the risky option is chosen and a zero otherwise, the effect of cognitive dissonance can be analyzed by including a dummy variable that is equal to one when the first period choice was the risky option. Cognitive dissonance implies that this coefficient will be positive in sign.

Let ΔEU^2_{ijt} denote the latent net gain in expected utility of a player of type i playing against a player of type j in period two of round t. Then the preceding discussion implies that we may write a regression model of the probability that a player chooses the risky option in period two as

$$Pr(\Delta EU^2_{ijt} > 0) = \beta^{\mathbf{2}}_{\mathbf{ijA}}\mathbf{D_X}\mathbf{D_{ij}} + \beta^{\mathbf{2}}_{\mathbf{ijB}}(\mathbf{X} - \mathbf{X^{S_i}_N})\mathbf{D_{ij}} + \beta^2_{CD}D_1 + \beta^{\mathbf{2}}_{\mathbf{t}}\mathbf{R_t} + \beta^{\mathbf{2}}_{\mathbf{P}}\mathbf{P_i} + \epsilon^2_{ijt}, \quad (11)$$

where $\mathbf{D_X}$ is a dummy variable equal to one if $X > X^{S_i}_N$ and equal to zero otherwise; $\mathbf{D_{ij}}$ is a dummy variable equal to one if the treatment is type $\{\$i, \$j\}$; $\mathbf{X} - \mathbf{X^{S_i}_N}$ is a vector of the differences between the number of observed successes and the critical number of successes to induce the risky choice, which is set equal to zero when neither player chooses the risky choice in period one; D_1 is a dummy variable equal to one if player i chose the risky choice in the first period; and as in period one, $\mathbf{R_t}$ is a vector of round fixed-effects and $\mathbf{P_i}$ is a vector of player fixed-effects. The β^2_k are parameters to be estimated.

The Bayesian predictions imply that $\beta^2_{ijA} = 1$, and that $\beta^2_{ijB} = \beta^2_{CD} = \beta_t = \beta_P = 0$. Since player error can only lower the proportion of players who satisfy these strict hypotheses, we state the Bayesian hypotheses as follows:

Hypothesis 7. $\beta^2_{ijA} \geq 0$: *Players apply Bayes' rule when making second period decisions. The alternative hypothesis is that players are less likely to choose the risky option in the second period when they have observed the critical proportion of successes.*

The salience hypothesis is:

Hypothesis 8. $\beta^2_{ijB} \geq 0$: *Salience increases the probability that a player chooses the risky option. The alternative hypothesis is that* $\beta^2_{4,8B} < 0$, $\beta^2_{8,4B} < 0$, $\beta^2_{8,12B} < 0$ *and* $\beta^2_{12,8B} < 0$, *or that salience has no effect upon second period choices.*

[16] We do not explicitly induce priors in the experiment, in order to maximize the value of information in the first period of the game. However, the result is an inability to distinguish cognitive dissonance from strong prior beliefs. That is, players may indeed update their beliefs in accordance with the Bayesian process, yet begin with such strong priors that they appear to be ignoring the information provided in the first period of the game.

The cognitive dissonance hypothesis is:

Hypothesis 9. $\beta^2_{CD} \geq 0$: *Choosing the risky option in period one increases the likelihood of choosing the risky option in period two. The alternative hypothesis is that* $\beta^2_{CD} < 0$, *cognitive dissonance has no effect upon second period choices.*

Finally, players in the $\{\$8, \$4\}$ treatment should be indistinguishable from players in the $\{\$8, \$12\}$ treatment once the second period arrives, since their incentives are identical at that period. This implies the following hypothesis:

Hypothesis 10. $\beta_{8,4A} = \beta_{8,12A}$ *and* $\beta_{8,4B} = \beta_{8,12B}$: *Whether a player is in the* $\{\$8, \$4\}$ *or* $\{\$8, \$12\}$ *treatment does not matter in the second period. The alternative hypothesis is that one or more of the equalities does not hold.*

4 Analysis of Results

4.1 Analysis of First Period Choices

We begin by examining aggregate measures of first period choices by treatment. Figure 4 depicts the proportion of players choosing the lottery and Table 1 presents the error rates conditioned only by treatment.

Table 1: First Period Error Rates by Treatment

Treatment	$\{\$4, \$8\}$	$\{\$8, \$4\}$	$\{\$8, \$12\}$	$\{\$12, \$8\}$
Error	$1 - \beta_{4,8}$	$\beta_{8,4} - 0$	$1 - \beta_{8,12}$	$\beta_{12,8}$
Percentage Errors	5.3%	22.14%	53.33%	2.08%

In Figure 4, the histogram at left depicts the entire sample of first period choices (1040 observations) and the histogram at right depicts the distribution when only one player in a pair chooses the lottery (644 observations). The panel at left in Figure 4 demonstrates that players respond to incentives: 95% of players in the $\{\$4, \$8\}$ treatment choose the risky option, while only 2% of those in the $\{\$12, \$8\}$ treatment choose the risky option. This pattern of behavior is consistent with Hypothesis 1. Furthermore, when we restrict our attention to cases where only one player in a pair chooses the risky option (i.e. the right panel in Figure 4), we see that the largest error rate is less than 4%. Thus, almost all deviations from theoretical predictions in period one are due to choices made by those in the $\{\$8, \$4\}$ and $\{\$8, \$12\}$ treatments. 22% of decisions in the $\{\$8, \$4\}$ treatment choose the

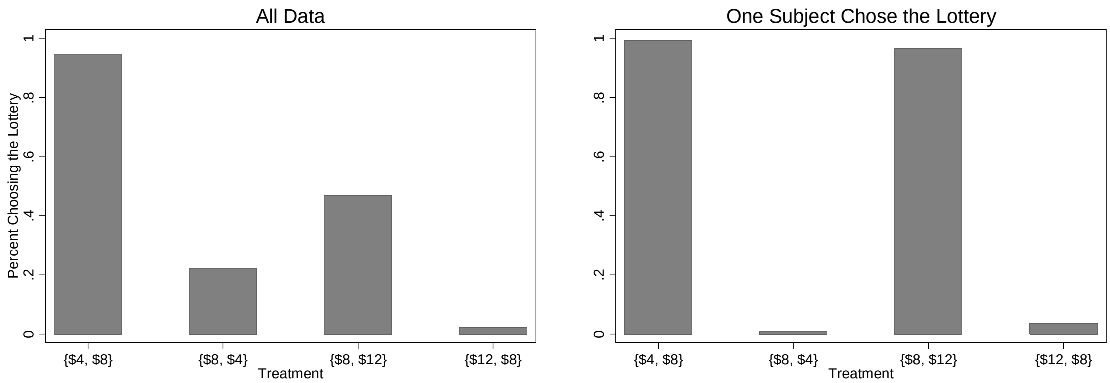

Figure 4: Frequency Distribution of First Period Risky Choices by Treatment.

risky option, even though the risky option is predicted to never be chosen in this treatment, and 47% of decisions in the $\{\$8, \$12\}$ treatment choose the risky option, even though the risky option is predicted to always be chosen in this treatment. While these frequencies differ from Nash equilibrium predictions, they differ from each other in the direction consistent with Hypothesis 2.

Choosing the lottery 47% of the time in the $\{\$8, \$12\}$ setting may appear to be the result of random choices. However, random choice would also generate a frequency of 50% of players choosing the risky option in the $\{\$8, \$4\}$ treatment. The lower error rate in the $\{\$8, \$4\}$ treatment than in the $\{\$8, \$12\}$ treatment is inconsistent with Hypothesis 4, that players' behavior is strategically irrational. However, this is consistent with Hypothesis 3, that players are behaving myopically. Myopia, which lowers the value of information, biases behavior toward the safe choice and lowers the return from the risky choice. But if players were behaving entirely myopically we would not expect a difference between the $\{\$8, \$4\}$ treatment and the $\{\$8, \$12\}$ treatment; both would produce a frequency of 0% risky choices. Thus, it appears that players undervalue information, yet value it "enough" to behave strategically.

While aggregate behavior appears to be consistent with Nash equilibrium predictions, it is important to determine whether the aggregate behavior is the result of many players deviating some of the time or a few deviating much of the time. We construct the error rate, defined to be the fraction of suboptimal choices, for each player in each treatment. The frequency distribution of player first period error rates across treatments are depicted in Figure 5. In 3 of 4 treatments the modal error rate is 0%. However, in the $\{\$8, \$12\}$

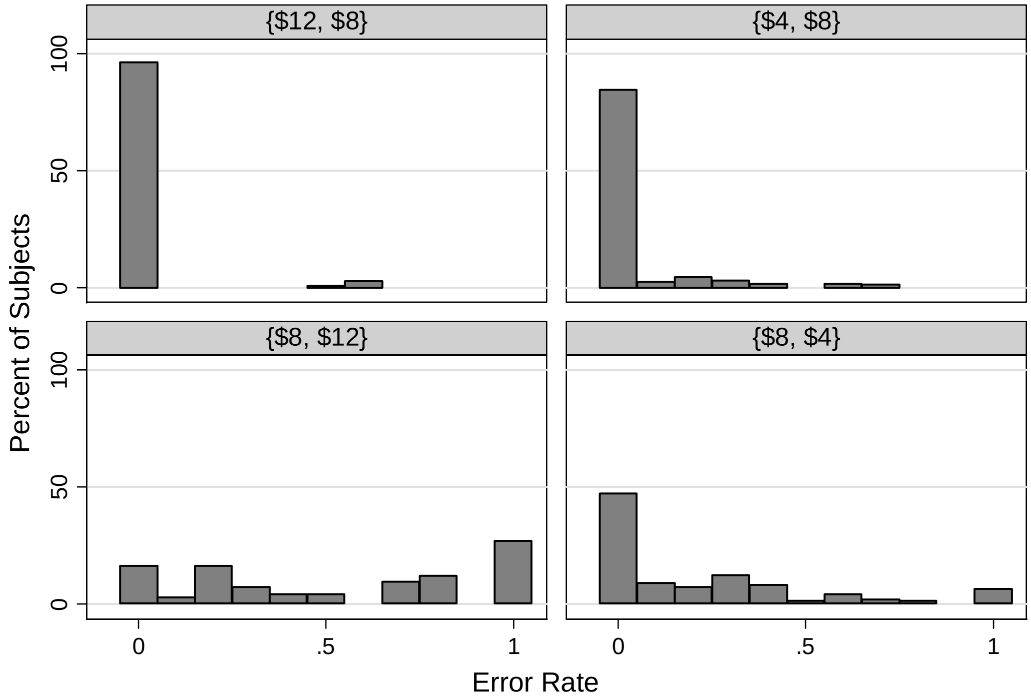

Figure 5: Frequency Distribution of First Period player Error Rate by Treatment.

treatment the modal error rate is 100%. Thus, in most treatments some players make some suboptimal choices, however, in the $\{\$8, \$12\}$ treatment there is a dramatically higher rate of players deviating frequently (e.g. 53% of decisions). Again, this pattern of error rates is consistent with undervaluation of information and lack of salience.

Since players repeat the game over twenty decision rounds, it is worthwhile to investigate the extent to which there are time trends (e.g. learning). Figure 6 shows the mean proportion of risky choices in each round by treatment. The mean proportions of players choosing the risky option in the $\{\$4, \$8\}$ treatment and the $\{\$12, \$8\}$ treatment are quite stable across rounds and are close to the theoretical predictions of one and zero, respectively. Treatments $\{\$8, \$4\}$ and $\{\$8, \$12\}$, however, reflect much more volatility and are quite far from the theoretical predictions of zero and one, respectively. Nevertheless, there are only two instances where the treatment time trends cross, and the vertical alignment is consistent with Hypotheses 1 and 2.

Analyzing the period one decisions, controlling for player-specific effects and round effects, yields useful insights. Table 2 reports the regression results for linear probability models estimated via ordinary least squares. Model (1) includes only the treatment dummy variables. Model (2) adds player and round fixed-effects. The results from the panel models

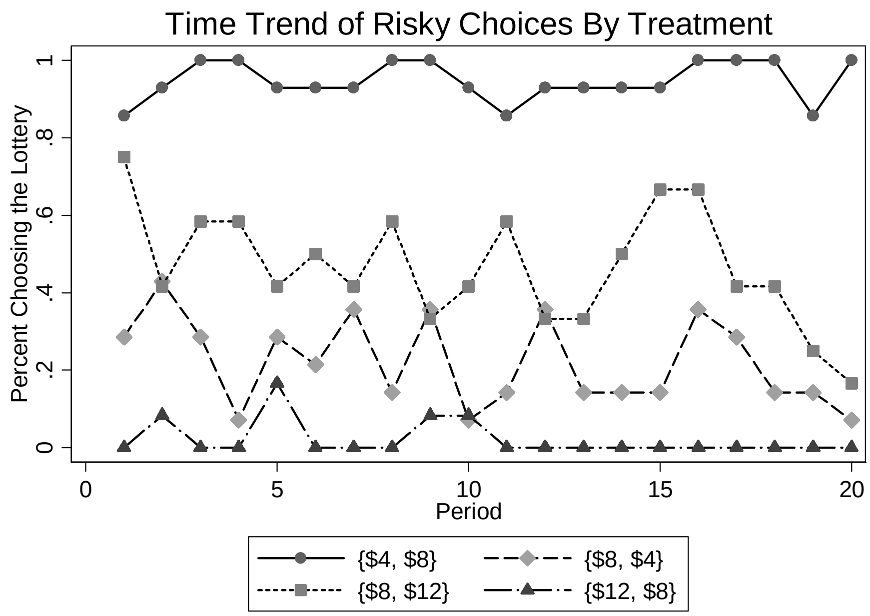

Figure 6: Mean Proportion of period one Risky Choices Across Periods by Treatment.

are consistent with the results from the pooled regression.

Table 3 reports tests of the hypotheses stated in the previous section using the regression models reported in Table 2. In general, the strict versions of the hypotheses are rejected in all cases. Players make errors and, by virtue of the experimental design, these errors are, one-sided. Our focus is on the weaker versions of the hypotheses that admit behavioral errors. We summarize the results of the hypothesis tests regarding period one decisions below.

Result 1. *Both models reject the hypothesis that players choose the risky option as often in the* $\{\$4, \$8\}$ *treatment as in the* $\{\$12, \$8\}$ *treatment, in the first period.*

Result 2. *Both models reject the hypothesis that players choose the risky option as often in the* $\{\$8, \$12\}$ *treatment as in treatment* $\{\$8, \$4\}$*, in the first period.*

Result 3. *Both models reject the hypothesis that errors in the* $\{\$8, \$4\}$ *treatment are larger than errors in the* $\{\$8, \$12\}$ *treatment in the first period.*

Result 4. *Neither model rejects the hypothesis that errors in the* $\{\$8, \$4\}$ *treatment are smaller than errors in the* $\{\$8, \$12\}$ *treatment in the first period.*

Table 2: Linear Probability Model Regression Results for First Period Choices.

	(1)	(2)
Treatment {\$4, \$8} Nash	0.946***	1.257***
	(0.019)	(0.055)
Treatment {\$8, \$4} Nash	0.221***	0.540***
	(0.040)	(0.055)
Treatment {\$8, \$12} Nash	0.467***	0.783***
	(0.055)	(0.065)
Treatment {\$12, \$8} Nash	0.021	0.334***
	(0.017)	(0.057)
Player Effects	No	Yes
Round Effects	No	Yes
R^2	0.714	0.785

Notes: The dataset consists of a panel of 52 players over 20 decision periods (1040 observations). Errors are clustered by player. Standard errors are reported in parentheses. Statistical significance of the estimated coefficients: "*" significant at the 10% level, "**"significant at the 5% level, and "***"significant at the 1% level.

Table 3: Hypothesis Test Results for First Period Choices.

	Hypothesis	(1)	(2)
1:	$H_O : \beta_{4,8} - \beta_{12,8} \leq 0$	36.74	34.1
	$H_A : \beta_{4,8} - \beta_{12,8} > 0$	(0.00)	(0.00)
2:	$H_O : \beta_{8,12} - \beta_{8,4} \leq 0$	4.7	4.82
	$H_A : \beta_{8,12} - \beta_{8,4} > 0$	(0.00)	(0.00)
3:	$H_O : 1 - \beta_{8,12} - \beta_{8,4} \leq 0$	3.85	2.94
	$H_A : 1 - \beta_{8,12} - \beta_{8,4} > 0$	(0.00)	(0.00)
4:	$H_O : 1 - \beta_{8,12} - \beta_{8,4} \geq 0$	3.85	2.94
	$H_A : 1 - \beta_{8,12} - \beta_{8,4} < 0$	(0.99)	(0.99)
5:	$H_O : 1 - \beta_{4,8} \geq 1 - \beta_{8,12}$	8.96	9.12
	$H_A : 1 - \beta_{4,8} < 1 - \beta_{8,12}$	(0.00)	(0.00)
6:	$H_O : \beta_{12,8} \geq \beta_{8,4}$	5.06	4.77
	$H_A : \beta_{12,8} < \beta_{8,4}$	(0.00)	(0.00)
	Player Fixed-Effects	N.A.	983.64
			(0.00)
	Round Fixed-Effects	N.A.	1.54
			(0.11)

Notes: Columns correspond to the models estimated in Table 2. The numbered hypothesis tests report the t-statistic. The demographic and round effects are F-statistics. The numbers in parentheses are the p-values. "N.A." means not applicable.

Result 5. *Both models reject the hypothesis that errors in the* $\{\$4, \$8\}$ *treatment are larger than errors in the* $\{\$8, \$12\}$ *treatment in the first period.*

Result 6. *Both models reject the hypothesis that errors in the* $\{\$12, \$8\}$ *treatment are larger than errors in the* $\{\$8, \$4\}$ *treatment in the first period.*

The results reported in Figures 4-6, as well as the regression analysis, are consistent with the error space being one-sided and the relative magnitude of decision cost versus decision reward. Andreoni (1995) noted that in a binary decision setting (his example was a public goods contribution game) errors can only be one-sided (i.e., contribute when it is rational to not). Thus, random decision error in a binary setting, such as ours, will only lower the sample proportion away from 1, as in treatment $\{\$8, \$12\}$, or raise the sample proportion above zero, as in treatment $\{\$8, \$4\}$. Smith and Walker (1993) demonstrate the importance of salience in the presence of random decision error. For a given cost of making a decision, here captured in the variance of the random error term, the probability of making the optimal decision is increasing. Hence, since treatments $\{\$8, \$4\}$ and $\{\$8, \$12\}$ are the least salient (i.e. they have the lowest opportunity cost of suboptimal behavior), these exhibit the largest error rates and the most volatile behavior. Still, we find support for Hypothesis 1 and 2, that players free-ride. Overall, the evidence suggests players respond to the incentives in the experimental treatments in a manner that is consistent with theory, but behavior is not nearly as responsive as theory predicts. In particular, players appear to be somewhat myopic (i.e., they undervalue information), nonetheless, they behave strategically based on what value they place on information. To further investigate whether players valued first period information, we analyze the extent to which they respond to the information signals. That is, do players actually use the first period information in a manner consistent with theory? Accordingly, we turn our attention to second period decisions.

4.2 Analysis of Second Period Decisions

Table 4 presents the error rates relative to the Bayesian predictions in the second period choices. When $N = 0$, players in treatment $\{\$4, \$8\}$ should choose the risky option, and players in treatments $\{\$8, \$4\}$, $\{\$8, \$12\}$, and $\{\$12, \$8\}$ should choose the safe option. When $N > 0$, players in each treatment should choose the risky option whenever $X \geq X_N^{S_i}$, and choose the safe option otherwise. When $N = 0$, players in the $\{\$4, \$8\}$, $\{\$8, \$12\}$, and $\{\$12, \$8\}$ treatments make few errors. But players in the $\{\$8, \$4\}$ make errors relative to the Bayesian predictions roughly $\frac{1}{4}$ of the time. When $N > 0$, error rates range from 10% to

Table 4: Second Period Error Rates by Treatment

Treatment	$\{\$4, \$8\}$	$\{\$8, \$4\}$	$\{\$8, \$12\}$	$\{\$12, \$8\}$
Error when $N = 0$	$1 - \beta_{4,8}$	$\beta_{8,4} - 0$	$\beta_{8,12} - 0$	$\beta_{12,8} - 0$
Percentage Errors	7.69%	23.08%	0.00%	3.23%
Error when $X < X_N^{S_i}$	$1 - \beta_{4,8}$	$1 - \beta_{8,4}$	$1 - \beta_{8,12}$	$1 - \beta_{12,8}$
Percentage Errors	38.36%	10.61%	24.53%	1.16%
Error when $X \geq X_N^{S_i}$	$\beta_{4,8} - 0$	$\beta_{8,4} - 0$	$\beta_{8,12} - 0$	$\beta_{12,8} - 0$
Percentage Errors	8.25%	33.33%	25.40%	70.00%

33% for players in the $\{\$8, \$4\}$ and $\{\$8, \$12\}$ treatments, and errors are roughly symmetric across positive and negative signals. Players in the $\{\$4, \$8\}$ and $\{\$12, \$8\}$ treatments respond differently to positive and negative signals. Players in the $\{\$4, \$8\}$ treatment make more errors when the signal is bad and players in the $\{\$12, \$8\}$ treatment make more errors when the signal is good. Given that these players are highly likely to have chosen the correct choice in the first period this suggests that their error rates are much higher when the signal does not reinforce their first period choices.[17]

To explore this more fully, Figure 7 depicts the proportion of players choosing the risky option according to the proportion of successes observed in the first period, for the cases where one or more players has chosen the risky option in period one. In the left panel, exactly one player in a pair chooses the risky option and in the right panel, one or both player(s) in a pair choose the risky option.[18] Figure 7 shows that across each treatment the propensity to choose the risky option is increasing in the number of observed successes. However, the model generates sharp predictions regarding responses to information.

Restricting our attention to the left panel, we see the largest increase in the propensity to choose the lottery in each treatment occurs after observing the critical number of successes. For example, given an uninformative prior, players in $\{\$4, \$8\}$ should choose the safe option only if they observe no successes out of three draws in the first period, otherwise they should choose the risky option. The largest increase in the propensity to choose the lottery in this treatment occurs after observing at least one success, an increase of 40%. The no uninformative prior prediction for the $\{\$8, \$4\}$ and $\{\$8, \$12\}$ treatments is to choose the safe option after observing less than two successes and to choose the risky option otherwise. In

[17]These error rates are influenced fairly significantly by a relatively small number of players. In the $\{\$4, \$8\}$ treatment 21 players always choose the risky option while none always choose the safe option. In the $\{\$8, \$4\}$ treatment 2 players always choose the risky option while 6 always choose the safe option. In the $\{\$8, \$12\}$ treatment 1 player always choose the risky option while 14 always choose the safe option. In the $\{\$12, \$8\}$ treatment 39 players always choose the safe option and none always choose the risky option.

[18]For 274 observations, N = 0; for 644 observations, N = 3; and for 122 observations, N = 6.

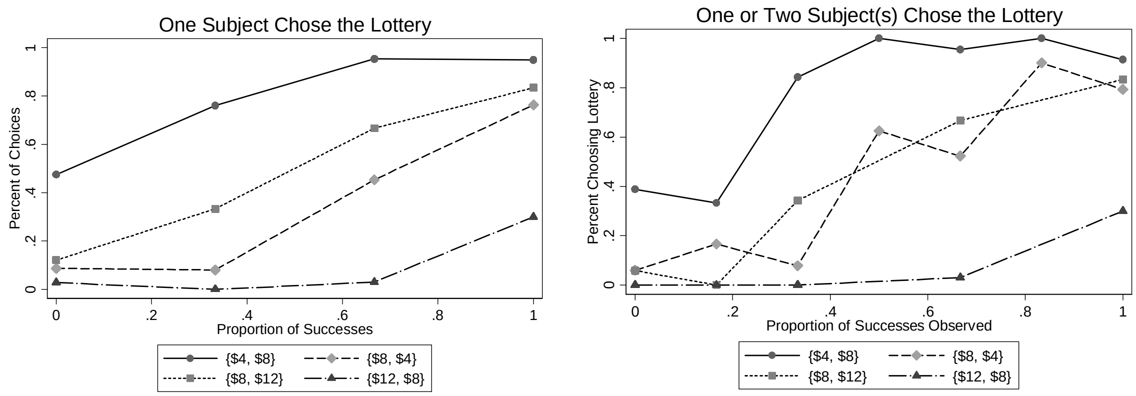

Figure 7: Percentage of Period Two Risky Choices by Proportion of Successes Observed.

both treatments we observe the largest increase in the propensity to choose the risky option occurring after observing at two successes, increases of 42% and 32% respectively. Finally, the no information prior prediction for the $\{\$12, \$8\}$ treatment is to choose the guaranteed amount after observing anything less than three successes. Indeed, the only significant increase in the propensity to choose the lottery occurs after observing three successes, an increase of 27%.

We saw from the first period analysis that the Nash equilibrium predicts that players are more likely to choose the risky option in the first period when they are in the $\{\$4, \$8\}$ and $\{\$8, \$12\}$ treatments.[19] It is in these treatments where cognitive dissonance would make a player more likely to continue to play their first period choice even in the face of contradictory evidence from the sample of first period draws. The left panel of Figure 7 shows that the propensity to choose the risky option is greatest in these two treatments.

Table 5 reports the regression results for linear probability model estimates of second period choices using both ordinary least squares. We estimate two equations using both treatment variables for hypothesis tests (model 1) as well as controlling for player and round fixed-effects (model 2). Models (1) and (2) include three types of variables: treatment effects interacted with the dummy variable D_{XN}, which is one when the critical value of X is observed, treatment effects interacted with the value $X - X_N^{S_i}$, and a dummy variable that is one if the first period choice was the risky option. The intercept in these models corresponds to the mean proportion of players who chose the risky option when there was

[19]Indeed, when only one player chooses the risky option (i.e. N = 3), the data in Figure 8 shows that 99% and 97% choose the risky option as $X/N \rightarrow 1$.

Table 5: Linear Probability Model Regression Results for Second Period Choices.

	(1)	(2)
Treatment $\{\$4, \$8\}$ Bayes	0.48***	0.57***
	(0.07)	(0.06)
Treatment $\{\$8, \$4\}$ Bayes	0.38***	0.41***
	(0.07)	(0.04)
Treatment $\{\$8, \$12\}$ Bayes	0.28***	0.31***
	(0.10)	(0.07)
Treatment $\{\$12, \$8\}$ Bayes	0.09	0.11
	(0.09)	(0.07)
Treatment $\{\$4, \$8\}$ Salience	0.02	0.01
	(0.02)	(0.02)
Treatment $\{\$8, \$4\}$ Salience	0.08***	0.06***
	(0.02)	(0.02)
Treatment $\{\$8, \$12\}$ Salience	0.14***	0.12***
	(0.05)	(0.04)
Treatment $\{\$12, \$8\}$ Salience	0.09***	0.08***
	(0.02)	(0.02)
Cognitive Dissonance	0.21***	0.19***
	(0.05)	(0.04)
Constant	0.20***	0.14
	(0.02)	(0.10)
Round Effects	No	Yes
Player Effects	No	Yes
R-squared	0.475	0.475

Notes: See the notes to Table 2.

no information generated from the first period decisions (i.e. $N = 0$).

The regression coefficients are interpreted as marginal effects given the linear specification. Twenty percent of players choose the risky option in period two when $N = 0$ draws are observed in period one. The coefficients on the treatments interacted with the dummy variables for the Bayesian criterion being satisfied are the increases in the proportion of players choosing the risky option when the Bayesian criterion that is satisfied. Thus, adding this to the constant yields the proportion of risky choices made by these players. The regressions indicate that the proportion who choose the risky option when the Bayesian criterion has been met varies from 0.68 (model 1) to 0.71 (model 2) for treatment $\{\$4, \$8\}$; from 0.55 (model 2) to 0.58 (model 1) for treatment $\{\$8, \$4\}$; from 0.45 (models 2) to 0.48 (model 1) for treatment $\{\$8, \$12\}$; and varies from 0.25 (model 2), which is not statistically different zero, to 0.29 (model 1) for treatment $\{\$12, \$8\}$. The results suggest an increase in the salience of the decision encourages players to choose the predicted choice, as all salience effects positive in sign and are statistically different from zero in all but the $\{\$4, \$8\}$ treatment. Players

Table 6: Hypothesis Test Results for Second Period Choices.

	Hypothesis	(1)	(2)
7:	$H_O : \beta_{4,8A} = \beta_{8,4A} = \beta_{8,12A} = \beta_{12,8A} = 0$	15.73	37.89
	$H_A : \beta_{i,jA} \neq 0$ for at least one i, j	(0.00)	(0.00)
8:	$H_O : \beta_{4,8B} = \beta_{8,4B} = \beta_{8,12B} = \beta_{12,8B} = 0$	7.44	7.39
	$H_A : \beta_{i,jB} \neq 0$ for at least one i, j	(0.00)	(0.00)
9:	$H_O : \beta_{CD} \leq 0$	17.04	25.73
	$H_A : \beta_{CD} > 0$	(0.00)	(0.00)
10:	$H_O : \beta_{8,4A} = \beta_{8,12A}$ and $\beta_{8,4B} = \beta_{8,12B}$	1.46 and 1.65	2.28 and 2.08
	$H_A : \beta_{8,4A} \neq \beta_{8,12A}$ or $\beta_{8,4B} \neq \beta_{8,12B}$, or both	(0.15) (0.17)	(0.13) (0.15)
	Player Fixed-Effects	N.A.	1.29
			(0.09)
	Round Fixed-Effects	N.A.	0.63
			(0.89)

Notes: The table reports F-statistics (p-values in parentheses) for the tests on the coefficients from Table 5.

also show a propensity to stick with their first period choices. The cognitive dissonance parameter shows that a player who chooses the risky option in period one is between 20% and 25% more likely to choose the risky option in period two.

The regression analysis permits formal testing of the hypotheses stated in Section 3.2. We report the results of the hypotheses tests in Table 6. We summarize the results of the hypotheses tests regarding period two decisions below.

Result 7. *Both models reject the hypothesis that the probability of choosing the risky option in the second period does not increase after observing the critical number of successes.*

Result 8. *Both models reject the hypothesis that salience has no effect upon deviations from the Nash predictions in the second period.*

Result 9. *Both models reject the hypothesis that cognitive dissonance has no effect upon deviations from the Nash predictions in the second period.*

Result 10. *Neither model rejects the hypothesis that players in the* $\{\$8, \$4\}$ *treatment behave the same as players in the* $\{\$8, \$12\}$ *treatment.*

We find support for the hypothesis that that the probability of choosing the risky option in the second period increases after observing the critical number of successes; lending support for Bayesian updating/reinforcement learning. We also find evidence that salience matters; namely, the more the number of successes observed exceeds the critical number of successes, the greater the probability of choosing the risky option in the second period. We also find evidence that cognitive dissonance influenced behavior. Specifically, if players' beliefs were such that they choose the lottery in the first period, they were more likely to do so in the second period. Note that this psychological effect is not inconsistent with the game theoretic model presented earlier. We do not explicitly induce prior beliefs. Hence, depending on the extent to which beliefs are skewed towards the lottery, players may very well update their beliefs in accordance with Bayes rule and still believe they were better off in the second period, regardless of the proportion of successes observed. Finally, players in treatments $\{\$8, \$4\}$ and $\{\$8, \$12\}$ should behave identically in the second period. We do not find a statistical difference in these players' second period behavior.

5 Discussion and Conclusions

We began with the question: how do information spillovers in a one-armed bandit game affect agents' decisions to take a risk, thereby providing information to others engaged in similar activities? This is a fairly common decision setting and an understanding of the information provided by players' decisions can inform policy debates concerning the public reporting of apprehension rates of various crimes, detection rates for tax evasion, success rates for innovators, performance of new technology, and so on. In such settings, some individuals will have private incentives to take risks but in doing so their behavior can inform others of the likelihood of a good (or bad) payoff. Regulatory policy would be more efficient if it incorporated the presence of such informational spillovers. The same can be said for policy directed toward encouraging risk taking - such as, policies encouraging research and development and/or the adoption of new technology. Future research will integrate our findings with the investment models of Dixit and Pindyck (1999).

Obviously, many real-world situations where agents engage in strategic experimentation are more complex than our laboratory environment. For example, the actions and outcomes of other agents may only be partially observable. Also, payoffs may not be independent. However, it is necessary to begin with a test of the simplest possible form of the game. After all, if behavior in the simplest setting is inconsistent with theoretical predictions, then we would not expect behavior in more complicated settings to conform with theory. Establishing behavioral regularities in the simplest possible scenario is a necessary prerequisite to investigations of behavior in more complicated environments. To the best of our knowledge, this is the first paper to test Nash equilibrium predictions in an armed-bandit problem with information spillovers.

Our laboratory results indicate that players in this setting exhibit behavior that is consistent with observations from public goods experiments. We find players free-ride too much when they should not, and too little when they should. Unlike most public goods settings, free-riding in the bandit problem is Pareto optimal. Hence, to the extent that they may be generalized, our results suggest that whether policy calls for a carrot or a stick depends upon the benefit-cost ratio of an activity and whether the activity itself is beneficial (e.g. technological innovation) or detrimental (e.g. noncompliance) to society. When a risky activity is beneficial to society then the incentives of individual agents and society are aligned. In the high benefit-cost ratio game (i.e. the \$4 type with a \$8 type) we observe too much risk-taking behavior and, by consequence, too much information provision. In such settings, a regulator could tax benefits in order to discourage risk-taking. Likewise, in the low benefit-cost ratio game (i.e. the \$12 type with a \$8 type), we observe less than optimal risk-taking behavior and information provision. In these situations a regulator could subsidize costs in order to encourage risk-taking. On the other hand, when a risky activity is detrimental to society, then the incentives of individual agents and society are in conflict. Hence, a regulator should do the exact opposite of the above policy recommendations.

Nonetheless, there are some encouraging results from the experiment that indicate policy intervention may not be entirely necessary. For example, when only one player pulls the arm, which accounts for about two-thirds of our observations, it is the predicted player at least 96% of the time; the others free-ride as predicted. Furthermore, players respond to information in a manner consistent with Bayesian updating/reinforcement with deviations attributable, in part, to reward salience and to cognitive dissonance. Our results should motivate future models of strategic experimentation to incorporate such factors.

References

Aghion, Philippe, Patrick Bolton, Christopher Harris, and Bruno Jullien, "Optimal Learning by Experimentation," *Review of Economic Studies*, 1991, *58*, 621 – 65.

Akerlof, George A. and William T. Dickens, "The Economic Consequences of Cognitive Dissonance," *American Economic Review*, 1982, *73*, 307 – 319.

Anderson, Christopher M., "Behavioral Models of Strategies in Multi-Armed Bandit Problems," *Ph.D Dissertation, California Institute of Technology*, 2001.

Andreoni, James, "Cooperation in Public Goods Experiments? Kindness or Confusion," *American Economic Review*, 1995, *85*, 891 – 904.

Banks, Jeffrey, Mark Olson, and David Porter, "An Experimental Analysis of the Bandit Problem," *Economic Theory*, 1997, *10*, 55 – 77.

Benveniste, Lawrence M., Alexander P. Ljungqvist, William J. Wilhelm Jr., and Xiaoyun Yu, "Evidence of Information Spillovers in the Production of Investment Banking Services," *The Journal of Finance*, 2003, *58*, 577 – 608.

Bergemann, Dirk and Juuso Välimäki, "Learning and Strategic Pricing," *Econometrica*, 1996, *64*, 1125 – 1149.

_ **and** _ , "Bandit Problems," *Helsinki Center of Economic Research Discussion Paper No. 93.*, 2006.

_ **and Ulrich Hege**, "Dynamic Venture Capital Financing, Learning and Moral Hazard," *Journal of Banking and Finance*, 1998, *22*, 703 – 735.

_ **and** _ , "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, 2005, *36*, 719 – 752.

Bolten, Patrick and Christopher Harris, "Strategic Experimentation," *Econometrica*, 1999, *67*, 349 – 374.

Caplin, Andrew and John Leahy, "Miracle on Sixth Avenue: Information Externalities and Search," *The Economic Journal*, 1998, *108*, 60 – 74.

Charness, Gary and Dan Levin, "When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity, and Affect," *American Economic Review*, 2005, *95*, 1300 – 1310.

DeGroot, Morris H., *Optimal Statistical Decisions*, New York: McGraw-Hill, 1970.

Dixit, Avanish and Robert Pindyck, *Investment Under Uncertainty*, Princeton, NJ: Princeton University Press, 1999.

El-Gamal, Mahmoud A. and David M. Grether, "Are People Bayesian? Uncovering Behavioral Strategies,," *Journal of the American Statistical Association*, 1995, *90*, 1137 – 1145.

Felli, Leonardo and Christopher Harris, "Job Matching, Learning and Firm-Specific Human Capital," *Journal of Political Economy*, 1996, *104*, 838 – 868.

Fischbacher, Urs, "Z-Tree: Zurich Toolbox for Readymade Economics Experiments," *Experimental Economics*, 2007, *10*, 171 – 178.

Gans, Noah, George Knox, and Rachel Croson, "Simple Models of Discrete Choice and Their Performance in Bandit Experiments," *Manufacturing & Service Operations Management*, 2007, *9*, 383 – 408.

Gittins, John C., "Bandit Processes and Dynamic Allocation Indices," *Journal of the Royal Statistical Society, Series B*, 1979, *41*, 148 – 177.

Greiner, Ben, "The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics," *Working Paper Series in Economics 10, University of Cologne.*, 2004.

Grether, David M., "Bayes Rule as a Descriptive Model: The Representativeness Heuristic," *Quarterly Journal of Economics*, 1980, *95*, 537 – 557.

_ , "Testing Bayes Rule and the Representativeness Heuristic: Some Experimental Evidence," *Journal of Economic Behavior and Organization*, 1992, *17*, 31 – 57.

Hendricks, Kenneth and Robert H. Porter, "The Timing and Incidence of Exploratory Drilling on Offshore Wildcat Tracts," *The American Economic Review*, 1996, *86*, 388 – 407.

Hirshleifer, Jack and John G. Riley, *The Analytics of Uncertainty and Information*, New York: Cambridge University Press, 1992.

Jovanovic, Boyan, "Job Search and the Theory of Turnover," *Journal of Political Economy*, 1979, *87*, 972 – 990.

Kahneman, Daniel and Amos Tversky, "Subjective Probability: A Judgment of Representativeness," *Cognitive Psychology*, 1972, *3*, 430 – 454.

_ **and** _ , "On the Psychology of Prediction," *Psychological Review*, 1973, *80*, 237 – 251.

Keller, Godfrey, Sven Rady, and Martin Cripps, "trategic Experimentation with Exponential Bandits," *Econometrica*, 2005, *73*, 39 – 68.

Kraemer, Carlo, Markus Nöth, and Martin Weber, "Information Aggregation with Costly Information and Random Ordering: Experimental Evidence," *Journal of Economic Behavior & Organization*, 2006, *59*, 423 – 432.

Lochner, Lance, "Individual Perceptions of the Criminal Justice System," *American Economic Review*, 2007, *97*, 444 – 460.

McKelvey, Richard D. and Talbot Page, "Public and Private Information: An Experimental Study of Information Pooling," *Econometrica*, 1990, *58*, 1321 – 1339.

Meyer, Robert J. and Yong Shi, "Sequential Choice under Ambiguity: Intuitive Solutions to the Armed-Bandit Problem," *Management Science*, 1995, *41*, 817 – 834.

Robbins, Herbert, "Some Aspects of the Sequential Design of Experiments," *Bulletin of the American Mathematical Society*, 1952, *55*, 527 – 535.

Roberts, Kevin and Martin L. Weitzman, "Funding Criteria for Research, Development and Exploration of Projects," *Econometrica*, 1981, *49*, 1261 – 1288.

Rothschild, Michael, "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 1974, *9*, 185 – 202.

Sah, Raaj K., "Social Osmosis and Patterns of Crime," *The Journal of Political Economy*, 1991, *99*, 1272 – 1295.

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

Smith, V.L., "Microeconomic Systems as an Experimental Science," *American Economic Review*, 1982, *72*, 923 – 995.

Tversky, Amos and Daniel Kahneman, "Beliefs in the Law of Small Numbers," *Psychological Bulletin*, 1971, *76*, 105 – 110.

_ **and** _ , "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology*, 1973, *5*, 207 – 232.

Weitzman, Martin L., "Optimal Search for the Best Alternative," *Econometrica*, 1979, *47*, 641 – 654.