Hindsight, Foresight, and Insight: An Experimental Study of a Small-Market Investment Game with Common and Private Values*†

(preliminary and incomplete)

Asen Ivanov
The Ohio State University

Dan Levin
The Ohio State University

James Peck
The Ohio State University

October 24, 2006

Abstract

We experimentally test an endogenous-timing investment model in which subjects observe their cost of investing and a signal correlated with the common investment return. Subjects overinvest, relative to the Nash benchmark. We can separately consider whether a subject draws inferences from the other subject's investment, in hindsight, and whether a subject has the foresight to delay profitable investment and learn from market activity. In contrast to Nash, cursed equilibrium, and level-k belief predictions, behavior remains the same across our experimental treatments. Our maximum likelihood estimates are inconsistent with belief-based theories, but are consistent with the notion that subjects use simple rules of thumb, based on insights about the game.

*This material is based upon work supported by the NSF under Grant No. SES-0417352. Any opinions, findings and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the NSF. We thank the John Glenn Institute at Ohio State for their support. We also thank Stephen Cosslett for helpful suggestions.

†ivanov.7@osu.edu, levin.36@osu.edu, peck.33@osu.edu.

1 Introduction

The theoretical literature on herding with endogenous timing, pioneered by Chamley and Gale (1994), explores the important issue of how market activity aggregates and transmits private information. Will firms with favorable information about the investment climate act on that information, thereby providing benefits to others, or will they postpone investment, to acquire information by observing other firms' investment activity? In Chamley and Gale (1994), firms receive a signal correlated with the unobserved investment return, which is common to all firms, and then face a sequence of decisions about whether or not to invest.[1] They find that the incentive to delay leads to inefficiency and the possibility of investment collapse. Indeed, firms are no better off than in a static game, in which firms must invest without learning anything about other firms' information. Levin and Peck (2006) introduce a second signal, about the cost of undertaking the investment, which is firm specific and independent of the costs faced by other firms. Observing the investment decisions of other firms could be used to improve inference about the aggregate state, but firms must disentangle whether another firm invests because it receives a favorable signal about investment returns or simply has a low cost.

Experimentally testing the theoretical implications of the endogenous timing herding literature is important for several reasons. First, we can compare the theoretical implications for aggregate investment activity on markets with what actually occurs in the lab. One intriguing possibility is that subjects with favorable information underestimate the option value of waiting and become more likely to invest immediately, thereby providing more information to the market and a more efficient outcome than the theoretical analysis would predict.[2] Second, interactions are purely informational

[1] The previous generation of herding models assumes that agents have one opportuninty to invest, and must decide in an exogenously specified sequence. See Banerjee (1992) or Bikhchandani, Hirshleifer, and Welch (1992). Also, Anderson and Holt (1997) provided the first experimental tests of herding models with exogenous timing.

[2] The other experimental work on endogenous timing herding models, by Sgroi (2003) and Ziegelmeyer et al (2006), are not well suited to address issues of aggregate investment, because in those games it is always optimal to "invest" eventually.

in our investment market, so decision making is not complicated by strategic concerns, as opposed to an auction environment where subjects must decide how much to bid, how others will respond, and so on. We can test separately (i) whether a subject understands the expected profits of investing, (ii) whether a subject draws inferences from the other's behavior (hindsight), and (iii) whether a subject looks ahead by considering the option value of delaying investment (foresight). On this score, we add to the work of Sgroi (2003) and Ziegelmeyer et al (2006).[3] Third, because interactions are purely informational, the setting offers a sharp test of various behavioral theories, such as *level-k beliefs* (see Crawford and Iriberri (2005)) and *cursed equilibrium* (see Eyster and Rabin (2005)).

Let us sketch our experimental treatments. Subjects are matched in two-player trials, and each subject observes her cost of investment and another signal that is correlated with the common investment return. Ranging from least favorable to most favorable, possible types are $(0, H)$, $(0, L)$, $(1, H)$, and $(1, L)$, where the first component is the common-value signal (0 is low and 1 is high) and the second component is the investment cost (L is low and H is high). At the beginning of each trial, the subjects observe their signals and simultaneously decide whether to invest in round 1. Then, subjects that do not invest in round 1 observe whether the other subject invested in round 1, and decide whether to invest in round 2, and so on. Investment is irreversible, and can be done at most once per trial per subject. In our Two-Cost Treatment, each subject's investment cost is high or low with probability one half, independent of the other subject's cost and the investment return. In our Alternating One-Cost Treatment, trials alternate between common knowledge of high cost and common knowledge of low cost.

We find that the typical subject is quite good at determining whether investment is profitable, based on her type. In particular, a type $(0, L)$ subject is far less likely to invest in round 1 than a type $(1, H)$ subject, even though it may not be obvious that a bad signal and low cost is slightly unprofitable and a good signal and high cost is

[3]Sgroi (2003) uses an "urn" setting, in which signals either strongly point to the (mostly) red urn, strongly point to the (mostly) white urn, or are completely neutral. The most profitable urn to pick and the proper inference from the other subjects' behavior are somewhat obvious. In Ziegelmeyer et al (2006), the most profitable investment choice is obvious for the first mover, and the only inference required is an understanding that the first mover has a stronger signal than someone who waits. Both papers exhibit an interesting tradeoff between the cost of delay and the information gained.

slightly profitable. Thus, this basic aspect of rationality is satisfied. Subjects usually treat investment by the other subject in the trial as good news about the investment return; they are more likely to invest in round 2 following investment in round 1 than following no investment in round 1. Thus, subjects show that they can draw inferences from the other subject's behavior, in hindsight. The evidence regarding foresight is nuanced. On the one hand, a significant fraction of type $(1, H)$ subjects, who would receive positive expected profits by investing in round 1, prefer costly waiting and invest in round 2 if and only if the other subject invested in round 1. It seems that these subjects have the foresight to understand that waiting will provide useful information. On the other hand, we argue below that it is unlikely that these subjects are computing the option value of waiting, which is then compared to the profits of investing immediately.

Here are our main results about aggregate investment and information flows. We find that the frequency of investment is higher and overall profits are lower than those predicted by the Nash equilibrium. Investment by the other subject generates an informational externality, which could be either positive or negative. There is a positive informational externality due to investment by subjects with a high common-value signal, and a negative informational externality due to investment by subjects with a low common-value signal. Not surprisingly, we find that the overall externality is positive in all our treatments. The more interesting comparison is to the Nash prediction. In our Two-Cost Treatment, the incremental positive informational externality (over Nash), due to overinvestment by subjects with a high common-value signal, is balanced by the negative externality created by unprofitable investment by subjects with a low common-value signal. The overall externality is as large as in the Nash equilibrium, in the sense that a subject best responding to the empirical distribution of strategies receives as high a profit as she would receive if everyone were playing Nash. In our Alternating One-Cost Treatment, the incremental positive externality due to overinvestment disappears, but it is the theoretical yardstick and not investment behavior that is different. We find that actual behavior is essentially the same across our two main treatments, even though the theory predicts different behavior.

For our two main treatments, there are essentially only three strategies that are consistent with Nash, level-k beliefs, or cursed equilibrium. The *self-contained* strategy, S, responds optimally to one's own information, but ignores the behavior of the other subject. The *myopic* strategy, M, is to invest in round 1 whenever investment is profitable (even if waiting is more profitable), but to properly infer (in hindsight) that investment by the other subject in round 1 is good news, and act accordingly. The *foresight* strategy, F, is to delay investment whenever valuable information can be revealed, but otherwise to invest when profitable. We perform maximum likelihood estimation on the proportion of S subjects, M subjects, and F subjects, allowing for errors. Even though the theory differs across our two main treatments, our maximum likelihood estimates are nearly identical. The proportion of F subjects is greater than one half, and the proportion of S subjects is greater than one third. These findings are inconsistent with any symmetric cursed equilibrium (see Eyster and Rabin (2005)), which allows for either F or S, depending on the level of cursedness, but not the coexistence of F and S. As far as level-k beliefs are concerned,[4] a level-1 subject plays S, and a level-2 subject plays F, so behavior across our two main treatments is broadly consistent with level-k beliefs.

Cursed equilibrium and level-k beliefs are belief-based theories, in which players form beliefs as specified in the theory and choose best responses accordingly. Our alternative interpretation of behavior is that subjects choose particular rules of thumb, based on various sorts of *insights* that subjects may experience. First, there is the insight that investment in round 1 is profitable for types $(1, H)$ and $(1, L)$, and not for the other types. Second, there is the insight that investment by the other subject signals the high common-value signal, calling for an updating of beliefs. Third, there is the insight that delaying investment may provide useful information. Thus, S, M, and F are interpreted as rules of thumb, reflecting the degree of insight acquired. To better separate our insight story from the belief-based explanations, we introduce a new treatment. In Treatment 3, a type $(1, H)$ subject receives higher expected profits by investing in round 1 than waiting, even if waiting revealed the other subject's type with probability one. Playing F is strictly dominated, so it is inconsistent with any theory of behavior involving a best response to beliefs about the other subject's strategy. On the other hand, it is plausible that our subject has the insight that waiting will yield useful information, but does not perform a calculation to see that

[4]See Stahl and Wilson (1995), Nagel (1995), or Crawford and Iriberri (2005).

F is dominated. Maximum likelihood estimates for Treatment 3 indicate that the proportion of F subjects remains above one half, providing support for the insight story and evidence against the belief-based theories.

2 Theoretical Framework

Our theoretical framework is based on the general model studied in Levin and Peck (2006). There are $n \geq 2$ risk-neutral players or potential investors. Let $Z \in \{0, 10\}$ denote the true investment return, common to all investors, with $pr(Z = 0) = pr(Z = 10) = \frac{1}{2}$. Each player $i = 1, ..., n$ observes a signal correlated with the investment return, $X_i \in \{0, 1\}$, which we call the "common-value" signal of player i. We assume that signals are independent, conditional on Z. The accuracy of the signal is given by the parameter, $\alpha \in [\frac{1}{2}, 1]$:

$$pr(Z = 0 \mid X_i = 0) = pr(Z = 10 \mid X_i = 1) = \alpha.$$

When we have $\alpha = \frac{1}{2}$, common-value signals have no information content at all, and when we have $\alpha = 1$, a common-value signal fully reveals the aggregate state. Thus, the parameter α effectively captures the informativeness of the common-value signal, X_i.

Each player i also privately observes a second signal, representing the idiosyncratic cost of undertaking the investment, c_i. We assume that c_i is independent of all other variables, and distributed according to a distribution function defined over the support, $[\underline{c}, \overline{c}]$. The structure of signals is common knowledge.

Impatience is measured by the discount factor, $\delta < 1$. If player i has cost c_i and the state is Z, her profits are zero if she does not invest, and $\delta^{t-1}(Z - c_i)$ if she invests in round t. We now describe the game. First, each player observes her signals, (X_i, c_i). For $t = 1, 2, ...$, each player observes the history of play through round $t - 1$, and players not yet invested simultaneously decide whether to invest in round t. A strategy for player i is a mapping from signal realizations and histories (including the null history) into a decision of whether to invest following that history. We require that a player can invest at most once.

Although Levin and Peck (2006) consider continuous cost distributions, our experimental design considers a discrete distribution containing either one or two points. This simplifies the decision making required of subjects and simplifies the data analysis. At the same time, it maintains the essential tradeoff between the incentive to delay and gain information by observing investment activity, versus the associated shrinkage of the (expected) pie due to discounting. We now define the three games relevant to our experiment, and solve for the Bayesian Nash equilibria. For the remainder of this section, we restrict attention to the parameter values, $n = 2$, $\delta = 0.9$, and $\alpha = 0.7$.

Game 1 (two costs):

There are two equally likely cost realizations, $L = 3.5$ and $H = 6.5$. Thus, we have four possible types of players, based on the common-value signal and the cost: $(0, H)$, $(0, L)$, $(1, H)$, and $(1, L)$. Equilibrium path play is uniquely determined, and involves pure strategies. A type $(0, H)$ player will never invest under any circumstances, because her expected profits from investing are negative even if she knows that the other player has the high common-value signal.[5] Similarly, a type $(1, L)$ player finds it profitable to invest even if she knows that the other player has the low common-value signal, and therefore invests in round 1. A type $(0, L)$ player will not invest in round 1, because the expected return is 3, while her cost is 3.5. However, since we have established that investment in round 1 must come from a player with the high common-value signal, a type $(0, L)$ player will invest in round 2 if the other player invests in round 1, because her conditional expected return becomes 5.

The remaining type, $(1, H)$, is the most interesting. Investment in round 1 yields positive expected profits of 0.5, but profits are slightly higher by taking advantage of the option value of waiting, investing in round 2 if and only if the other player invests in round 1. If all other type $(1, H)$ players wait, profits from waiting are 0.5085. If some other type $(1, H)$ players would invest in round 1, the advantage of waiting is even greater. Thus, we have characterized the equilibrium path, which is given in Table 1.[6] To simplify the discussion, denote the choice to invest in round 1 as “1”,

[5]In such a case, her conditional expected return would be 5, but her cost is 6.5.

[6]Several specifications of behavior and beliefs off the equilibrium path are consistent with sequential equilibrium, all yielding the same equilibrium path. After no one invests in round 1 and one player invests in round 2, beliefs about the investor’s common-value signal can affect the the

	Equilibrium Strategy	Expected Profits for each type
type $(0, H)$	N	0
type $(0, L)$	W	0.2835
type $(1, H)$	W	0.5085
type $(1, L)$	1	3.5

Table 1: Equilibrium Characterization for the Two-Cost Game

denote the choice never to invest as "N", and denote the choice to wait and invest immediately following investment by the other player as "W".

Game 2 (low cost):

There is only one possible cost realization, 3.5. Thus, we have two possible types of players, $(0, L)$ and $(1, L)$. It is easy to see that a type $(1, L)$ player would want to invest, no matter what it believes about the other player's type, so she invests in round 1. A type $(0, L)$ player finds it unprofitable to invest in round 1, but will invest in round 2 if the other player invests in round 1 (implying a high common-value signal).

Game 3 (high cost):

There is only one possible cost realization, 6.5. Thus, we have two possible types of players, $(0, H)$ and $(1, H)$. It is easy to see that a type $(0, H)$ player would not want to invest, no matter what it believes about the other player's type, so she never invests. A type $(1, H)$ player compares the profits of choices "1" and "W". If all other type $(1, H)$ players choose "1", then she can act with full information in round 2, and "W" is a best response. If all other type $(1, H)$ players choose "W", then nothing is learned from waiting, and "1" is a best response. Therefore, type $(1, H)$ players mix, choosing "1" with probability 0.4916 and choosing "W" with probability 0.5084.

In our Alternating One-Cost Treatment, the subjects alternate between Game 2 and Game 3. Because matching is random and anonymous, folk theorem issues do

remaining player's decision. However, the beliefs and subsequent decision do not affect the original investor's profit, so the equilibrium path is unaffected.

	Equilibrium Strategy	Expected Profits for each type
type $(0, H)$	N	0
type $(0, L)$	W	0.567
type $(1, H)$	1 with probability 0.4916	0.5
	W with probability 0.5084	
type $(1, L)$	1	3.5

Table 2: Equilibrium Characterization for the Alternating One-Cost Game

not arise, so the equilibrium characterization combines the equilibria of Game 2 and Game 3, as given in Table 2.

The Bayesian Nash equilibria characterized in Table 1 and Table 2 allow for inefficient delay. A type $(1, H)$ player does not take into account the positive informational externality that investing in round 1 provides to the other player. Thus, if our experimental subjects are more likely to invest than the theory predicts when they are type $(1, H)$, it is possible that profits are higher than the theory predicts. Of course, it is also possible that subjects invest in round 1 with the low common-value signal, leading to a negative informational externality and lower profits.

3 Behavioral Considerations

In the investment games we study here, interactions between players are purely informational, with no direct payoff consequences. This simple structure allows us to test separately whether subjects invest when investment is unprofitable; whether subjects draw inferences, from round 1 investment by the other subject (*hindsight*); and whether they delay profitable investment in order to observe the other subject's decision (*foresight*). These different aspects of rationality would be more difficult to disentangle in auctions or other games with strategic interaction.[7] In particular, we focus on three strategies, which will be relevant for the three behavioral theories we

[7]For example, Garratt and Keister (2006) experimentally test a model of bank runs, where a player's decision to withdraw deposits simultaneously involves Bayesian updating, considering the option of withdrawing in the future, and anticipating the strategies of the other players.

consider. A subject who disregards the behavior of the other subject will invest if and only if investment is profitable based upon her own signals. Such a subject chooses the type-dependent strategy $S \equiv (N, N, 1, 1)$, where S stands for *self-contained*.[8]

A subject who invests whenever investment is profitable, incorporating information from the history of play, chooses the type-dependent strategy $M \equiv (N, W, 1, 1)$, where M stands for *myopic*. Thus, an M subject is able to draw inferences from market activity, in *hindsight*, but does not look with *foresight* to the informational benefits of waiting. Note that it is possible that an M subject of type $(1, H)$ might understand that there is a benefit to waiting, but feel that investment in round 1 is the better choice.[9]

Consider a subject who updates beliefs, in hindsight, and looks with *foresight* to the useful information that can be gained by waiting when her type is $(0, L)$ or $(1, H)$.[10] Such a subject chooses the type-dependent strategy $F \equiv (N, W, W, 1)$. Note that this *foresight* does not necessarily mean that she actually calculates the value of waiting, or even that such a calculation would justify waiting.

The strategies, S, M, and F, have interpretations that depend on the theory used to justify behavior. We focus on two belief-based theories, cursed equilibrium and level-k beliefs, and a theory of boundedly rational rules of thumb, based on insights a subject might acquire about understanding the game. As it turns out, we can characterize the Nash equilibrium, level-k behavior, and the cursed equilibrium, all in terms of these three strategies, so it is natural to focus on them.[11]

There is considerable evidence that experimental subjects fail to pick actions in accordance with the relevant Nash equilibrium. Such discrepancies are particularly acute in tasks that require players to make inferences and update their priors based

[8]The type-dependent strategy, (A, B, C, D), means that type $(0, H)$ players choose A, type $(0, L)$ players choose B, type $(1, H)$ players choose C, and type $(1, L)$ players choose D.

[9]Another interpretation is that the subject is somehow more impatient than that implied by the discount factor, 0.9.

[10]A type $(0, H)$ subject does not receive useful information by waiting, because the expected profits from investment are always negative. A type $(1, L)$ subject also does not receive useful information by waiting, because the expected profits from investment are always positive.

[11]There is one other strategy in which a subject never makes an unprofitable investment, and invests immediately when nothing can be learned to make investment unprofitable. This strategy is $(N, N, W, 1)$. Maximum likelihood estimation, allowing for all four strategies, does not change our conclusions. Indeed, allowing this strategy would stack the deck against cursed equilibrium and level-k beliefs, in favor of our explanation of insight-based rules of thumb.

on other players' actions in games with incomplete information.[12] There are also several studies claiming such failure in real markets.[13] Faced with such overwhelming evidence, it is not surprisng that there were several attempts to explain these discrepancies. Stahl and Wilson (1995) and Nagel (1995) use a non-equilibrium model of "level-k" beliefs, where L_0 players behave in some pre-determined way (usually randomly), and for $k = 1, 2, ...$, the L_k players choose a best response to the strategy chosen by the L_{k-1} players. See Crawford and Iriberri (2005) for a survey and an explanation for the winner's curse in auctions, based on level-k beliefs. Eyster and Rabin (2006) propose an alternative theory, which they call "cursed equilibrium." Players are assumed to best respond to the other players' strategies in a certain sense. In a $\chi-$cursed equilibrium, players believe that with probability χ, each other player j chooses an action that is type-independent, and whose distribution is given by the ex ante distribution of player j's actions. Also, players believe that with probability $1-\chi$, each other player j chooses an action according to player j's type-dependent strategy. Thus, if $\chi = 0$, we have a standard Bayesian Nash equilibrium, and if $\chi = 1$, players draw no inferences about other players' types. Both level-k beliefs and cursed equilibrium weaken the "usual" requirements of correct beliefs regarding other players' strategies, while insisting on players rationally choosing best responses to the more flexible belief structures that are allowed.

For our Two-Cost Treatment (Game 1) and our Alternating One-Cost Treatment (Games 2 and 3), the Nash equilibrium, level-k behavior, and the cursed equilibrium are characterized in Table 3, all in terms of the three strategies, S, M, and F.[14]

[12] Leading examples from laboratory studies include failures in the *takeover game* (see Ball and Bazerman (1991) and Charness and Levin (2005)), and systematic overbidding and the winner's curse in common-value auctions (see Bazerman and Samuelson (1983), Kagel and Levin (1986), Levin, Kagel, and Richard (1996), Holt and Sherman (1994), and Kagel and Levin (2002)).

[13] Leading examples include oil and gas lease auctions (see Capen, Clapp, and Campbell (1971), Mead, Moseidjord, and Sorensen (1983) and (1984), and the opposite conclusions reached in Hendricks, Porter, and Boudreau (1987)), baseball's free agent market (see Cassing and Douglas (1980) and Blecherman and Camerer (1998)), book publishing (see Dessauer (1981)), initial public offerings (see Levis (1990) and Rock (1986)), and corporate takeovers (see Roll (1986)).

[14] We are being a little loose when we refer to F, M, and S as strategies, because behavior is not specified after unexpected contingencies that do not affect play in the NE, the cursed equilibrium, or under level-k beliefs. In our maximum likelihood estimation (Section 5), we are careful to specify the decisions following investment by the other subject in round 2 (after not investing in round 1), and following a tremble that is inconsistent with the strategy itself.

	Two-Cost Game	Alternating One-Cost Game
Nash Equilibrium	F	M with probability 0.4916
		F with probability 0.5084
Level-k		
L_1	S	S
L_2	F	F
L_3	F	M
L_4	F	F
Cursed Equilibrium		
$0 < \chi < \frac{17}{756}$	F	M with probability q^{alt}
		F with probability $1 - q^{alt}$
$\frac{17}{756} < \chi < \frac{517}{756}$	M with probability $q^{2\text{cost}}$	M with probability q^{alt}
	F with probability $1 - q^{2\text{cost}}$	F with probability $1 - q^{alt}$
$\frac{517}{756} < \chi < \frac{3}{4}$	M	M
$\frac{3}{4} < \chi < 1$	S	S

Table 3: Nash, Level-k, and Cursed Equilibrium

The probabilities given in Table 3, $q^{2\text{cost}}$ and q^{alt}, are the following.

$$q^{2\text{cost}} = \frac{756\chi - 17}{9(113 - 84\chi)} \quad \text{and} \quad q^{alt} = \frac{500}{9(113 - 84\chi)}.$$

For our Two-Cost Treatment (Game 1), the Nash equilibrium strategy is F,[15] and for our Alternating One-Cost Treatment (Games 2 and 3), the symmetric Nash equilibrium involves mixing over M and F according to the probabilities given. Now let us consider level-k beliefs, starting with the Two-Cost Treatment. A player choosing a best response to an L_0 player who plays randomly can learn nothing from the other player's behavior, so an L_1 player chooses the strategy S. An L_2 player chooses the best response to S, which is F. For $k = 3, 4, ...$, an L_k player chooses the best response to F, which is F. For the Alternating One-Cost Treatment, an L_1 player chooses the strategy S, and an L_2 player chooses the best response to S, which is F. However, the situation is different from the Two-Cost Treatment, where a type $(1, H)$ is better off waiting if others are playing F, based on the hope that the other player is type

[15] In fact, for the Two-Cost Treatment, suppose a player knows that everyone is playing one of the strategies, S, M, or F. Then the best response is F, no matter how many of the other players are choosing each of the three strategies.

$(1, L)$. In the Alternating One-Cost Treatment, a type $(1, H)$ is playing the high-cost game (Game 3), and knows that the other player also has high cost. Therefore, if the other player chooses F, this implies that the other player never invests in round 1, in which case the best response is to invest. It follows that the best response to F is M. On the other hand, the best response to M is F, because if the other player invests in round 1 when type $(1, H)$, a type $(1, H)$ is better off waiting.

Computing the symmetric cursed equilibria for our games is a bit more involved, and details are given in the Appendix. In the Two-Cost Treatment, when the cursedness parameter, χ, is small, the Nash equilibrium strategy, F, continues to be played. For $\frac{17}{756} < \chi < \frac{517}{756}$, if type $(1, H)$ players wait, the informativeness of round 1 investment is sufficiently weakened that it does not pay them to wait. However, if all type $(1, H)$ players invest in round 1, now it pays to wait. Therefore, the cursed equilibrium involves mixing by type $(1, H)$ players, but a type $(0, L)$ player will still want to invest if she sees the other player invest. Thus, players mix between M and F, with the probabililty of choosing the myopic strategy, $q^{2\text{cost}}$, increasing in χ. For $\frac{517}{756} < \chi < \frac{3}{4}$, the level of cursedness is high enough so that a type $(1, H)$ player is better off investing in round 1 than waiting, even if all other $(1, H)$ players are also investing in round 1. A type $(0, L)$ player will still want to invest if she sees the other player invest, so in the $\chi-$cursed equilibrium the players choose the pure strategy, M. Finally, for $\chi > \frac{3}{4}$, the level of cursedness is so high that no useful inferences can be made, and the equilibrium strategy is S.

For the Alternating One-Cost Treatment, when $\chi < \frac{517}{756}$, players mix between M and F in the cursed equilibrium, with the probabililty, q^{alt}, of choosing the myopic strategy M increasing in χ. The probability of choosing M ranges from the Nash equilibrium value, $q^{alt} = 0.4916$ for $\chi = 0$, to unity, $q^{alt} = 1$ for $\chi = \frac{517}{756}$. For $\frac{517}{756} < \chi < \frac{3}{4}$, the informativeness of round 1 investment is too low to justify waiting for a type $(1, H)$, but it is high enough for a type $(0, L)$ to be willing to invest after she sees the other player invest. Thus, the cursed equilibrium strategy is M. For $\chi > \frac{3}{4}$, even a type $(0, L)$ is unwilling to invest after she sees the other player invest, and the cursed equilibrium strategy is S.

Of course, there are many possibilities for bounded rationality that go beyond misspecification of beliefs.[16] For our investment games, a small but nonzero proportion of type $(0, H)$ subjects invests in round 1, which is a mistake indicating a lack of understanding of the game. If mistakes like this occur, perhaps less extreme departures from Nash behavior reflect a lack of insight about some of the fine points of the game.

Let us elaborate on the insights we have in mind. Investment in round 1 is profitable for types $(1, H)$ and $(1, L)$, and not for the other types. We start with the notion that subjects have the basic insight that the expected revenue is 7 when receiving the high common-value signal and 3 when receiving the low common value signal. The second level of sophistication occurs when a subject, upon seeing the other subject invest, revises upwards the expected investment return, because she realizes that the other subject is very likely to have the high common-value signal. Based on our parameters, any upward revision should be enough to induce a type $(0, L)$ subject to invest in round 2 after the other subject invests in round 1 (a numerical computation is not required). The third level of sophistication occurs when a subject has the insight that there is a tradeoff, between potentially higher profits of investing early versus the information gained by waiting. Since a type $(1, H)$ subject with this third level of sophistication does not compute the expected profits from waiting, she might invest in round 1 or wait, depending on how she evaluates this tradeoff.

We call these strategies rules of thumb because they are not based on computation of expected profits, by applying the discount factor and using Bayes' rule explicitly. Not only would such a computation be time consuming, but it would require deeper insights about the game that few subjects are capable of acquiring during the session.[17] Think of the subjects, either as the instructions are read or early in the session, as having a "Eureka moment" giving them a partial understanding of the game. Under this view, it is reasonable to suppose that the exact parameter values, or the difference between the Two-Cost Treatment and the Alternating One-Cost Treatment in how costs are determined, should not affect the generation of these insights. Fewer than 10% of our subjects make choices that are inconsistent with F, M, or S in more

[16] See, for example, Simon (1972) and Rubinstein (1998), including references therein.

[17] A subject with total command of the game would need to think about the other subject's type-contingent probability of investing in round 1, in order to compute the probability that the other subject has the high common-value signal, conditional on behavior and one's own type.

than half of their trials.[18]

4 Experimental Design

The experiment consisted of the Two-Cost Treatment, the Alternating One-Cost Treatment, and Treatment 3, which was adopted to clarify some of the behavioral issues. We conducted two sessions of the Two-Cost Treatment (18 participants in each session), two sessions of the Alternating One-Cost Treatment (26 and 20 participants, respectively), and one session of Treatment 3 (28 participants).[19]

In the Two-Cost Treatment, subjects played Game 1 (see Section 2) in each trial, so that each subject's private cost of investment was randomly selected to be either 3.5 or 6.5. In the Alternating One-Cost Treatment, subjects played Game 2 in odd numbered trials and Game 3 in even numbered trials, so that each subject's cost alternated between 3.5 and 6.5 from trial to trial (but was the same for all subjects within a trial). Treatment 3 was the same as our Alternating One-Cost Treatment, except that the discount factor was given by $\delta = 0.8$ and the costs alternated between 3.5 and 5.7. To guarantee that the trials ended, without changing the equilibria, subjects were told that the trial ended after either both subjects had invested or there were two consecutive rounds with no investment.

Each session in all treatments consisted of two practice trials and 24 trials in which subjects played for real money. At the start of each trial, subjects were randomly and anonymously matched in pairs to form separate two-player markets which bore no relation to each other. There was a new random matching from trial to trial. Subjects were given an initial cash balance of 20 experimental currency units (ECU). In addition, they could gain (lose) ECU in each trial, which were added to (subtracted from) their cash balances. At the end of the session, ECU were converted into dollars at a rate of 0.6 $/ECU in our two main treatments, and 0.5 $/ECU in Treatment 3. Subjects were paid the resulting dollar amount or $5, whichever was greater. If a subject's cash balances fell below 0 at any point during the session, that subject was

[18] Excluding these subjects from the sample has almost no effect on our estimates for the proportion of F, M, and S subjects, although the error term is reduced.

[19] We also conducted a pilot session for the Two-Cost Treatment (14 participants).

paid \$5 and was asked to leave.[20] Average earnings for the Two-Cost Treatment, the Alternating One-Cost Treatment, and Treatment 3 were \$22.82, \$23.53, and \$20.27 respectively. Including the reading of instructions, sessions lasted between 1 hour 45 minutes and 2 hours.

Subjects in the experiment were students at The Ohio State University who were enrolled in undergraduate classes in Economics. The sessions were held at the Experimental Economics Lab at OSU. Before starting the trials, the experimenter read the instructions aloud as subjects read along, seated at their computer terminals. Subjects were invited to ask questions, including after the practice trials. Once the real trials began, no more questions were allowed. See the Appendix for our Instructions and a printout of the screen seen by a player with cost $c_i = 6.5$ and signal $X_i = 1$, who is deciding whether to invest in round 2 after the other player has invested in round 1.

The experiment was programmed and conducted with the software z-Tree (Fischbacher (1999)).

5 Results

5.1 Aggregate-Level Analysis

Table 4 presents aggregate-level decisions in the Two-Cost and the Alternating One-Cost Treatments. There are only six possible histories after which a subject can invest: the null history, $\{\}$; the history following no investment in round 1, $\{0\}$; the history following one subject investing in round 1, $\{1\}$; the history following no investment in round 1 and one subject investing in round 2, $\{0, 1\}$; and so on. The first six rows of the table show, for each treatment, type, and history, how many times a subject facing a decision in that situation invested. There are only three possible histories after which no investment would end the game: $\{0\}$, $(1, 0\}$, and $\{0, 1, 0\}$. Rows 7-9 show, for each treatment, type, and history for which no investment would

[20]This occurred for two subjects in the Two-Cost Treatment, for one subject in the Alternating One-Cost Treatment, and for one subject in Treatment 3. After a subject goes bankrupt, if the number of subjects in a session is no longer even, one subject was randomly selected to sit out during each trial.

History	Two-Cost				Alternating One-Cost			
	(0,H)	**(0,L)**	**(1,H)**	**(1,L)**	**(0,H)**	**(0,L)**	**(1,H)**	**(1,L)**
{}	20 (9%)	27 (14%)	67 (35%)	148 (78%)	16 (6%)	35 (13%)	109 (42%)	206 (74%)
{0}	12 (9%)	12 (11%)	22 (27%)	11 (46%)	19 (9%)	14 (10%)	30 (27%)	13 (46%)
{1}	9 (15%)	22 (37%)	26 (59%)	18 (100%)	9 (15%)	54 (53%)	24 (62%)	34 (79%)
{0,1}	2 (9%)	5 (45%)	7 (58%)	4 (100%)	5 (14%)	5 (28%)	7 (88%)	1 (100%)
{1,0}	6 (12%)	17 (46%)	6 (33%)	0 N/A	8 (16%)	16 (33%)	9 (60%)	4 (44%)
{0,1,0}	0 (0%)	1 (17%)	0 (0%)	0 N/A	3 (10%)	4 (31%)	1 (100%)	0 N/A
no {0}	101	83	49	9	160	106	74	14
no {1,0}	44	20	12	0	43	32	6	5
no {0,1,0}	20	5	5	0	27	9	0	0
Total	214	192	194	190	290	275	260	277

Table 4: Aggregate Actions and Frequency of Investment at each History

end the game, how many times a subject in that situation made the final decision *not* to invest. The numbers in parenthesis show what percentage of decisions made at a particular history were decisions to invest. For example, in the Two-Cost Treatment, the 22 times a type $(0, L)$ subject invested after history $\{1\}$ represent 37% of all decisions taken at that history.[21] The last line of the table shows the total number of realizations of each type. Note that a subject's play in a given trial is counted in one, and only one, cell of the table. Hence the last line is simply the sum of all previous lines.

We are interested in performing statistical tests, based on investment decisions and realized profits in our experiment. However, given that there is dependence between the investment decisions and profits of the two subjects in any trial, we will arbitrarily

[21]To see how the 37% figure is computed, note that the number of times a type $(0, L)$ subject passed through history $\{1\}$ is the sum of the 22 times a type $(0, L)$ subject invested after $\{1\}$, plus the 17 times a type $(0, L)$ subject invested after $\{1, 0\}$, plus the 20 times a type $(0, L)$ subject who experienced history $\{1, 0\}$ ended up not investing. Thus, $\frac{22}{(22+17+20)} = 0.37$.

call one subject in each market an A subject (the one with the lower subject ID) and the other subject a B subject. Then we will perform the relevant test separately for A subjects and B subjects. Note that now each test will be based on observations, no two of which were from the same market trial.[22]

Let us start by checking if subjects' aggregate behavior satisfies some basic requirements on rationality. First note that, in the aggregate, subjects in both the Two-Cost and Alternating One-Cost Treatments respond to their own information (common value signal and investment cost). This can be seen when we consider investment in round 1, and when we consider investment in any round.

The higher the expected profit from investment, given a subject's type, the more likely she is to invest in round 1. In particular, the frequencies with which subjects of type $(0, H)/(0, L)/(1, H)/(1, L)$ invest in round 1 are 9%/14%/35%/78% in the Two-Cost and 6%/13%/42%/74% in the Alternating One-Cost Treatment. The increase in the probability of investing in round 1, per ECU increase in the expected profit of investing in round 1, is 0.134 ($se = 0.017; p < 0.001$) in the Two-Cost Treatment and 0.125 ($se = 0.01; p < 0.001$) in the Alternating One-Cost Treatment. The estimating procedure used was a random effects probit.

The higher the expected payoff from investment, given a subject's type, the more likely she is to invest during some round. In particular, the frequencies with which subjects of type $(0, H)/(0, L)/(1, H)/(1, L)$ invest are 23%/44%/66%/95% in the Two-Cost and 21%/47%/69%/93% in the Alternating One-Cost Treatment.[23] The increase in the probability of investing in some round, per ECU increase in the expected profit of investing in round 1, is 0.158 ($se = 0.016; p < 0.001$) for A subjects and 0.132 ($se = 0.014; p < 0.001$) for B subjects in the Two-Cost Treatment. The corresponding effect is 0.131 ($se = 0.011; p < 0.001$) for A subjects and 0.133 ($se = 0.011; p < 0.001$) for B subjects in the Alternating One-Cost Treatment (estimated via random effects probit).

The data suggest that subjects understand when investment is profitable and when it is not profitable, based on their signals. The fact that slightly over one in five type

[22]There is no dependence between both subjects' decisions to invest conditional on any particular history, so we do not have to split the sample when we test behavior after any particular history.

[23]These numbers are obtained from Table 4 by dividing the sum of the first six numbers in a column (the number of times that type invested) by the last number in the column (the total number of times that type occurred).

$(0, H)$ subjects eventually invests is somewhat high, because one should realize that even when the other subject has the high signal, the two signals would cancel and expected revenue is 5. However, nearly four in five type $(0, H)$ subjects get it right and never invest. It is quite impressive that less than 15% of type $(0, L)$ subjects invest in round 1, even though expected profits are only slightly negative.[24] Also impressive is that a type $(1, H)$ subject is far more likely to invest than a type $(0, L)$, even though the expected profits are only slightly positive.[25]

We now move on to the question of whether subjects respond to the behavior of the other subject in their trial. We are interested in determining whether subjects were more likely to invest after seeing the other subject invest in round 1, as compared to seeing the other subject not invest in round 1. In the Two-Cost Treatment, the frequency with which subjects invest immediately after the history $\{1\}$ is 42%, and the frequency with which subjects invest immediately after the history $\{0\}$ is 16%. In the Alternating One-Cost Treatment, the corresponding frequencies are 50% and 15%. Using a random effects probit, the incremental effect of the other subject investing in round 1, on the probability that a subject invests in round 2 (controlling for subjects' types) is 0.292 ($se = 0.051; p < 0.001$) in the Two-Cost Treatment and 0.306 ($se = 0.039; p < 0.001$) in the Alternating One-Cost Treatment.

Let us summarize our results about the ability of subjects to invest only when profitable and to draw inferences from the investment of the other subject.

Result 1 *In the aggregate, for both treatments, (i) types with higher expected profits are more likely to invest in round 1 and are more likely to invest eventually, and (ii) subjects are more likely to invest in round 2 after the other subject invests in round 1 than after the other subject does not invest in round 1. Both results are statistically significant.*

Now that we have established that behavior satisfies some basic requirements on rationality, let us turn to the question of how well it complies with Nash equilibrium

[24]Of course, the percentage of type $(0, L)$ subjects who eventually invest is higher, but much of that investment is following an investment by the other subject in the trial, which is consistent with equilibrium.

[25]It is not a problem that the percentages of type $(1, H)$ subjects who invest in round 1 are only 35% in the Two-Cost Treatment and 42% in the Alternating One-Cost Treatment. Many subjects who do not invest understand that investment is profitable, but are waiting to obtain more information.

(NE). Table 5 shows the percentage of decisions in each treatment which comply with NE, broken down by type and overall.[26]

	(0,H)	(0,L)	(1,H)	(1,L)	Overall
Two-Cost	77%	60%	45%	78%	65%
Alternating One-Cost	79%	63%	82%	74%	75%

Table 5: Compliance with Nash equilibrium

Compliance with NE for $(0, H)$ and $(1, L)$ types is high in both treatments. Note that these are the types for which there is no need for strategic interaction with the other subject. A type $(0, H)$ subject should not invest, even if she knew that the other subject's common-value signal were 1, and a type $(1, L)$ subject should invest, even if she knew that the other subject's common-value signal were 0. The rate of compliance with NE for $(0, L)$ and $(1, H)$ types is more interesting, since decisions are less clear cut, and these types must draw inferences from the other subject's behavior in the NE. The table shows that for type $(0, L)$ subjects in both treatments, and for type $(1, H)$ subjects in the Two-Cost Treatment, a substantial percentage of decisions deviates from the NE. Compliance with NE for type $(1, H)$ subjects in the Alternating One-Cost Treatment is high, but this is due to the fact that the NE involves mixing, so it is consistent with NE to invest in round 1, and also to wait and invest in round 2 if the other subject invests in round 1.

One might conjecture that behavior moves closer to NE in later trials. In the Two-Cost Treatment, this is indeed the case. Using random effects probit estimation, the marginal effect of the trial number on the probability that a decision is consistent with NE (controlling for subjects' types) is 0.0057 ($se = 0.0026; p = 0.03$). Thus, the probability of a subject playing her NE strategy increases by $24 \times 0.0057 = 0.137$ over the 24 trials, which is not a small effect. However, the same marginal effect in the Alternating One-Cost Treatment is actually negative (although insignificantly so), suggesting that there is no movement towards NE in that treatment.

Result 2 *When only one strategy is consistent with NE and the choice is not clear cut (i.e., types (0,L) and (1,H) in the Two-Cost Treatment and type (0,L) in the*

[26] In computing these numbers, we treat any decisions as consistent with NE, following a history of no investment in round 1 and the other subject investing in round 2. See footnote 6.

Alternating One-Cost Treatment), then a substantial percentage of decisions (37%-55%) deviates from the NE. The probability that a decision is consistent with NE increases with time in the Two-Cost Treatment, but not in the Alternating One-Cost Treatment.

Let us now turn to the question of how actual investment and profit outcomes compare with those in the NE.

	Two-Cost					**Alternating One-Cost**				
	(0,H)	**(0,L)**	**(1,H)**	**(1,L)**	**Overall**	**(0,H)**	**(0,L)**	**(1,H)**	**(1,L)**	**Overall**
Actual	23%	44%	66%	95%	56%	21%	47%	69%	93%	57%
NE	0%	21%	29%	100%	38%	0%	42%	64%	100%	51%

Table 6: Investment

Table 6 shows, for the Two-Cost and Alternating One-Cost Treatments, the actual frequency of investment as well as the ex ante expected frequency of investment in the NE (broken down by type as well as overall). In both treatments, the actual frequency of investment exceeds the NE frequency for all types, except type $(1, L)$, where the actual frequency comes close to the NE frequency of 100%. In both treatments, the actual overall frequency of investment is higher than the NE frequency. This overinvestment is both economically (especially in the Two-Cost Treatment) and statistically significant ($p < 0.001$ for both A and B subjects in the Two-Cost Treatment and $p = 0.011/p = 0.003$ for A/B subjects in the Alternating One-Cost Treatment[27]).

Result 3 *In both treatments, the actual overall frequency of investment is significantly higher than the NE frequency of investment. In the Two-Cost Treatment, overinvestment is especially pronounced (actual frequency of investment is 56% vs. 38% in the NE).*

Table 7 shows, for the Two-Cost and Alternating One-Cost Treatments, the average actual profits per period as well as the NE expected profits per period (broken

[27]These p-values are obtained in the following way. In order to account for possible dependence between investment decisions made by the same player, we perform a random effects probit estimation with only a constant as a right-hand-side variable. Then we test the hypothesis that the estimated constant is the same as the constant which leads to a predicted probability of investment equal to the NE ex ante probability of investment.

	Two-Cost					Alternating One-Cost				
	(0,H)	**(0,L)**	**(1,H)**	**(1,L)**	**Overall**	**(0,H)**	**(0,L)**	**(1,H)**	**(1,L)**	**Overall**
Actual	-0.56	0.07	0.21	3.61	0.79	-0.34	0.03	0.44	3.05	0.79
NE	0	0.28	0.51	3.50	1.07	0	0.57	0.50	3.50	1.14

Table 7: Average Profits per Period (in ECU)

down by type as well as overall). In both treatments, actual average profits are lower than the NE expected profits. Due to the high variance of realized profits, this difference is not statistically significant at the 5% level ($p = 0.215/p = 0.092$ for A/B subjects in the Two-Cost Treatment and $p = 0.078/p = 0.064$ for A/B subjects in the Alternating One-Cost Treatment).[28] The reason that profits are lower than the NE prediction is primarily due to unprofitable investment by subjects with the low common-value signal, and the ensuing unprofitable investment by subjects drawing the wrong inference. Result 4 summarizes our findings about aggregate profits. Determinants of profits at the individual level will be considered later.

Result 4 *Actual average profits per trial are 73% of expected NE profits in the Two-Cost Treatment, and 69% of expected NE profits in the Alternating One-Cost Treatment.*

Central to our study is the informational interaction between subjects. A positive informational externality exists when subjects are likely to invest in round 1 with type $(1, H)$ or $(1, L)$, but unlikely to invest in round 1 with type $(0, H)$ or $(0, L)$.[29] Not surprisingly, this is indeed the case: in the Two-Cost Treatment, subjects with common-value signal 1 invest in round 1 55% of the time, and subjects with common-value signal 0 invest in round 1 only 12% of the time. In the Alternating Two-Cost Treatment, the corresponding percentages are 59% and 9%.[30] In the NE, these

[28]These p-values are obtained in the following way. In order to account for possible dependence between profits earned by the same player in different periods, we perform a random effects regression with only a constant as a right-hand-side variable. Then we test the hypothesis that the estimated constant is the same as NE ex ante profits.

[29]The informational externality created by investment after the history $\{0\}$ is more tricky, because this behavior is off the equilibrium path, and it is unclear that one should infer that the investor has the high common-value signal. This behavior is relatively rare in our experiment, and we ignore it in our analysis of the informational externality.

[30]The marginal effect of a subject having common-value signal 1, on the probability that she invests in round 1, is highly significant in both treatments ($p < 0.001$), estimated via random effects probit.

percentages would be 50% and 0% in the Two-Cost Treatment, and 75% and 0% in the Alternating Two-Cost Treatment.[31]

The difference between the actual informational externality and the NE prediction involves three effects. First, the difference is increased if type $(1, H)$ subjects invest in round 1 more often than in the NE. Second, the difference is decreased if type $(1, L)$ subjects fail to invest in round 1. Third, the difference is decreased if subjects with common-value signal 0 invest in round 1. The first effect is present in the Two-Cost Treatment, but not in the Alternating One-Cost Treatment. The second and third effects are present in both treatments (see tables 1, 2 and 4).

How do we combine these three effects to compare the actual informational externality with the NE prediction? To answer this question we compare the ex ante profits that a subject would receive, in the NE, with the ex ante profits that a subject would receive, based on best responding to the empirical frequencies of strategies chosen in each treatment. We will refer to this hypothetical best responder as a BR subject.[32] Thus, we compare the long run profits earned by choosing the optimal strategy in a market where others play the NE strategy, and the long run profits earned by choosing the optimal strategy in a market where others play according to the empirical distribution of strategies chosen in our experiment. The ex ante profits in the NE and for a BR subject are 1.07 ECU in the Two-Cost Treatment. In the Alternating One-Cost Treatment, the ex ante profits in the NE are 1.14 ECU, and the profits for a BR subject are 1.08 ECU. The difference in profits for the Alternating One-Cost Treatment is not negligible, given that a subject who decides whether to invest based solely on her own information (and ignores any information provided by the other subject) earns expected profits of 1 ECU. Our findings about information flows are summarized in Result 5.

Result 5 *Behavior in both the Two-Cost and Alternating One-Cost Treatments creates (in the aggregate) a positive informational externality. This externality is as*

[31] Even in the NE, the informational externality is inefficiently low, because a type $(1, H)$ player ignores the benefit that investing provides to the other player. Thus, it is possible that the actual play could be more informationally efficient than the NE play. We are currently running large market experiments ($n = 10$) to explore this possibility.

[32] A BR player plays F (and invests after $\{0, 1\}$ when she is type $(1, H)$) in the Two-Cost Treatment and plays M (and invests after $\{0, 1\}$ when she is type $(0, L)$ or $(1, H)$) in the Alternating One-Cost Treatment.

large as the NE externality in the Two-Cost Treatment, and is smaller than the NE externality in the Alternating One-Cost Treatment (by 5%).

Inspection of Tables 4, 6 and 7 indicates that behavior and outcomes in the Two-Cost and Alternating One-Cost Treatments are remarkably similar.[33] To quantify this similarity, we employ random effects probit estimation with treatment dummies as right hand side variables, to test the hypothesis that there is no treatment effect on (i) key history and type-contingent investment choices, (ii) aggregate investment, and (iii) aggregate profits. First, let us compare behavior across treatments for some key types, after the history, $\{\}$, and after the history, $\{1\}$. In the Two-Cost/Alternating One-Cost Treatment, type $(1, L)$ subjects invest in round 1 78%/74% of the time ($p = 0.373$). In the Two-Cost/Alternating One-Cost Treatment, type $(1, H)$ subjects invest in round 1 35%/42% of the time ($p = 0.397$). In the Two-Cost/Alternating One-Cost Treatment, type $(1, H)$ subjects invest after the history, $\{1\}$, 59%/62% of the time ($p = 1$). In the Two-Cost/Alternating One-Cost Treatment, type $(0, L)$ subjects invest after the history, $\{1\}$, 37%/53% of the time ($p = 0.305$).

As can be seen from table 6, frequency of investment is very similar between treatments: 56% vs. 57%. Testing the hypothesis that there is no treatment effect on the probability of investing (controlling for subjects' types) yields a p-value of 0.885/0.554 for A/B subjects (within a random effects probit model). From table 7, average profits per period are very similar between treatments: 0.79 ECU in both treatments. Testing the hypothesis that there is no treatment effect on profits (by regressing profits on a treatment dummy within a random effects model, controlling for subjects' types) yields a p-value of 0.813/0.982 for A/B subjects. Summarizing these comparisons across treatments, we have:

Result 6 *We cannot reject the hypothesis that there is no treatment effect on (i) key history and type-contingent investment choices, (ii) aggregate investment, and (iii) profits.*

[33]Table 5 shows the compliance with NE, and differs across the two treatments. This is because the NE themselves differ across the treatments, not the actual behavior.

5.2 What is Driving Behavior?

Given the substantial departure from NE behavior and the finding that there is no clear tendency towards NE behavior as the trials progress (Result 2), it remains to study what drives the actual behavior. Our task is simplified by the fact that there are essentially only three strategies that are consistent with level-k beliefs or cursed equilibrium: F, M and S. We say "essentially" because there are variations on these strategies, based on how decisions are made following a mistake, or when the other subject invests in round 2 after no one invests in round 1. Furthermore, these same three strategies can also be interpreted as rules of thumb, based on insights a subject might have about the value of observing the other subject's behavior. We will proceed as follows. First, we consider a model for each treatment, in which subjects are drawn from a population of subjects who play one of the strategies F, M or S, fully specified off the equilibrium path. For $j \in \{F, M, S\}$, the probability of a subject being drawn from class j is denoted by p_j. At each decision node a subject faces, we assume that the subject chooses as dictated by her strategy class with probability $(1 - \varepsilon)$, and makes the other decision with probability ε. We then estimate, via maximum likelihood, the parameters $(p_F, p_M, p_S, \varepsilon)$.[34] Next, armed with our estimates of the proportions of the population in each strategy class for the Two-Cost Treatment, the Alternating One-Cost Treatment, and Treatment 3 (described below), we evaluate three behavioral theories: cursed equilibrium, level-k beliefs, and the insight-based rules of thumb described above.

Before proceeding to the maximum likelihood estimation, we must fully specify the strategies, F, M and S. For now, we specify the behavior of a subject playing according to her strategy class, and we will introduce errors later. The basic principle we use is that a subject corrects her own departures, and chooses each action with probability $\frac{1}{2}$ following an unexpected choice by the other subject.[35] Consider first a subject of class F. When type $(1, L)$, she invests with probability 1 following all histories. When type $(1, H)$, she invests with probability 1 following the histories $\{1\}$ and $\{1, 0\}$, invests with probability $\frac{1}{2}$ following the histories $\{0, 1\}$ and $\{0, 1, 0\}$,

[34] This statistical framework is similar to that in many experimental papers, especially Camerer and Harless (1994) and Costa-Gomes, Crawford and Broseta (2001). The main difference is that, in our context, subjects are making a sequence of decisions in each game.

[35] Other specifications yield similar results, because these departures are relatively rare.

and does not invest following the histories $\{\}$ and $\{0\}$. When type $(0, L)$, she invests with probability 1 following the histories $\{1\}$ and $\{1, 0\}$, invests with probability $\frac{1}{2}$ following the histories $\{0, 1\}$ and $\{0, 1, 0\}$, and does not invest following the histories $\{\}$ and $\{0\}$. When type $(0, H)$ she does not invest following all histories.

Consider next a subject of class M. When type $(1, L)$, she invests with probability 1 following all histories. When type $(1, H)$, she invests with probability 1 following the histories $\{\}$, $\{1\}$, and $\{1, 0\}$, invests with probability $\frac{1}{2}$ following the histories $\{0, 1\}$ and $\{0, 1, 0\}$, and does not invest following the history $\{0\}$. When type $(0, L)$, she invests with probability 1 following the histories $\{1\}$ and $\{1, 0\}$, invests with probability $\frac{1}{2}$ following the histories $\{0, 1\}$ and $\{0, 1, 0\}$, and does not invest following the histories $\{\}$ and $\{0\}$. When type $(0, H)$ she does not invest following all histories.

Finally, consider a subject of class S. When type $(1, L)$ or type $(1, H)$, she invests with probability 1 following all histories. When type $(0, L)$ or type $(0, H)$, she does not invest following all histories.

Our model of behavior is that, each time a subject of class $j \in \{F, M, S\}$ has to make a decision to invest or not invest, given her type and the observed history of the other subject's behavior, she makes the decision prescribed by strategy j with probability $1 - \varepsilon$ and makes an "error" with probability $\varepsilon \in [0, 0.5]$. Errors are assumed to be i.i.d. across types and histories, trials, and subjects. Table 8 shows the probability that a class F subject invests, or makes a final decision never to invest, conditional on her type, conditional on the history, and conditional on the fact that the other subject's behavior allows the history to occur. To understand how the table is constructed, consider the following examples. The probability that an F subject of type $(0, L)$ invests after the history $\{1\}$ is $(1 - \varepsilon)^2$, since she acted according to her strategy class twice: by not investing in round 1 and then by investing after the other subject invests in round 1. The probability that an F subject of type $(0, L)$ invests after the history $\{1, 0\}$ is $\varepsilon(1 - \varepsilon)^2$, since she acted according to her class by not investing in round 1, then she made an "error" by not investing after $\{1\}$, and finally she acted according to her class by recovering from her error and investing after $\{1, 0\}$. The probability that an F subject of type $(0, L)$ ends up not investing after experiencing history $\{0\}$ is $(1 - \varepsilon)^2$, since she made two type-consistent decisions: she did not invest either after $\{\}$ or after $\{0\}$.

History	(0,H)	(0,L)	(1,H)	(1,L)
{}	ε	ε	ε	$(1-\varepsilon)$
{0}	$\varepsilon(1-\varepsilon)$	$\varepsilon(1-\varepsilon)$	$\varepsilon(1-\varepsilon)$	$\varepsilon(1-\varepsilon)$
{1}	$\varepsilon(1-\varepsilon)$	$(1-\varepsilon)^2$	$(1-\varepsilon)^2$	$\varepsilon(1-\varepsilon)$
{0,1}	$\varepsilon(1-\varepsilon)^2$	$\frac{1}{2}(1-\varepsilon)^2$	$\frac{1}{2}(1-\varepsilon)^2$	$\varepsilon^2(1-\varepsilon)$
{1,0}	$\varepsilon(1-\varepsilon)^2$	$\varepsilon(1-\varepsilon)^2$	$\varepsilon(1-\varepsilon)^2$	$\varepsilon^2(1-\varepsilon)$
{0,1,0}	$\varepsilon(1-\varepsilon)^3$	$\frac{1}{4}(1-\varepsilon)^2$	$\frac{1}{4}(1-\varepsilon)^2$	$\varepsilon^3(1-\varepsilon)$
no {0}	$(1-\varepsilon)^2$	$(1-\varepsilon)^2$	$(1-\varepsilon)^2$	ε^2
no {1,0}	$(1-\varepsilon)^3$	$\varepsilon^2(1-\varepsilon)$	$\varepsilon^2(1-\varepsilon)$	ε^3
no {0,1,0}	$(1-\varepsilon)^4$	$\frac{1}{4}(1-\varepsilon)^2$	$\frac{1}{4}(1-\varepsilon)^2$	ε^4

Table 8: Probability of an F subject's behavior

For M subjects and S subjects, we can construct tables analogous to Table 8. An M subject behaves in the same way as an F subject, except when her type is $(1, H)$. Thus, columns 1, 2, and 4 are as in Table 8, but the nine entries in column 3 should instead be: $(1-\varepsilon), \varepsilon^2, \varepsilon(1-\varepsilon), \frac{1}{2}\varepsilon(1-\varepsilon), \varepsilon^2(1-\varepsilon), \frac{1}{4}\varepsilon(1-\varepsilon), \varepsilon(1-\varepsilon), \varepsilon^3$, and $\frac{1}{4}\varepsilon(1-\varepsilon)$. An S subject behaves in the same way as an F subject when her type is $(0, H)$ or $(1, L)$. Furthermore, her behavior when her type is $(0, L)$ is identical to her behavior when her type is $(0, H)$, and her behavior when her type is $(1, H)$ is identical to her behavior when her type is $(1, L)$. Therefore, columns 1 and 2 are identical to column 1 in Table 8, and columns 3 and 4 are identical to column 4 in Table 8.

Before constructing the likelihood function, we need some more notation. We number all of a subject's trials by $t = 1, 2, ..., 24$. Let B_i^t denote the full behavior of subject i during (i's) trial t. By full behavior, we mean the round in which she invests, if at all. We formalize B_i^t as a four dimensional vector of zeros and ones. $B_i^t = (0, 0, 0, 0)$ signifies that the subject did not invest, the vector $B_i^t = (0, 0, 1, 0)$ signifies that she invested in round 3, and so on. Let $-i_t$ denote the subject matched with subject i during trial t, and let B_{-i}^t denote the full behavior of subject $-i_t$ during trial t.[36] Denote the behavior of subject i over all trials as B_i, where we have $B_i = (B_i^1, ..., B_i^t, ..., B_i^{24})$, and denote the behavior of all the subjects matched with subject i (during the trials they are matched with i) as B_{-i}, where we have $B_{-i} = (B_{-i}^1, ..., B_{-i}^t, ..., B_{-i}^{24})$. Finally, let $T_i^t \in \{(0, H), (0, L), (1, H), (1, L), -1\}$ de-

[36]Define $B_i^t = B_{-i}^t = (-1, -1, -1, -1)$ if i did not participate in trial t (either because she sat out or because she went bankrupt).

note subject i's type during trial t, where type -1 means that subject i was sitting out or bankrupt, let $T_i = (T_i^1, ..., T_i^t, ..., T_i^{24})$, and let $T = (T_1, ..., T_i, ...T_n)$.

From Table 8, or the analogous tables corresponding to M subjects and S subjects, the probability of B_i^t is determined, given that the subject is of strategy class j, type T_i^t, and given that the other subject's behavior is B_{-i}^t.[37] We denote this probability, which also depends on the parameter ε, as

$$\Pr(B_i^t | j, T_i^t, B_{-i}^t; \varepsilon).$$

For example, suppose that in trial t, subject i is type $(1, H)$ and the other subject invests in round 1, $T_i^t = (1, H)$ and $B_{-i}^t = (1, 0, 0, 0)$. If subject i is an F subject, then the probability of $B_i^t = (1, 0, 0, 0)$ is ε (row 1, column 3 of Table 8), the probability of $B_i^t = (0, 1, 0, 0)$ is $(1-\varepsilon)^2$ (row 3, column 3 of Table 8), the probability of $B_i^t = (0, 0, 1, 0)$ is $\varepsilon(1-\varepsilon)^2$ (row 5, column 3 of Table 8), and the probability of $B_i^t = (0, 0, 0, 0)$ is $(1-\varepsilon)\varepsilon^2$ (row 8, column 3 of Table 8). Given that the other subject invests in round 1, investing in round 4 is impossible, so the probability of $B_i^t = (0, 0, 0, 1)$ is zero.

The probability that subject i chooses behavior B_i, given that her strategy class is j, given her type realizations T_i, and given the behavior of the other subjects she faces is B_{-i}, is given by

$$\Pr(B_i | j, T_i, B_{-i}; \varepsilon) = \prod_{t=1}^{24} \Pr(B_i^t | j, T_i^t, B_{-i}^t; \varepsilon).$$

Thus, we can compute the probability that subject i chooses behavior B_i, given her type realizations T_i, and given that the behavior of the other subjects she faces is B_{-i},[38]

$$\Pr(B_i | T_i, B_{-i}; p_F, p_M, p_S, \varepsilon) = \sum_{j \in \{F,M,S\}} p_j \Pr(B_i | j, T_i, B_{-i}; \varepsilon).$$

[37] If a subject is sitting out trial t or has gone bankrupt, then $B_i^t = B_{-i}^t = (-1, -1, -1, -1)$ with probability one.

[38] The other subjects' behavior does not affect the probability of being in class F, M, or S, except for the unlikely event that a subject has gone bankrupt, so we have $B_{-i}^t = (-1, -1, -1, -1)$. Not only were bankruptcies exceedingly rare, but the probabilities of each strategy class (conditional on bankruptcy) would not change very much. The stronger inference is that the subject made many "errors." We ignore this complication.

	Two-Cost	**Alternating One-Cost**
p_F	0.607 (0.102)	0.567 (0.083)
p_M	0 N/A	0.092 (0.058)
p_S	0.393 (0.102)	0.341 (0.078)
ε	0.192 (0.011)	0.176 (0.01)
log-likelihood	-757.5402	-1002.5052

Table 9: Maximum Likelihood Estimates

In the Appendix, we show that the likelihood function is given by[39]

$$\Pr(B|T; p_F, p_M, p_S, \varepsilon) = \prod_{i=1}^{n} \Pr(B_i|T_i, B_{-i}; p_F, p_M, p_S, \varepsilon). \tag{1}$$

For each treatment, we estimate the vector of parameters $\theta = (p_F, p_M, p_S, \varepsilon)$, by maximizing the likelihood function (1).[40] Table 9 shows our estimates, along with estimated standard errors, for the Two-Cost and the Alternating One-Cost Treatments.[41]

As can be seen from Table 9, in both of our main treatments, the population frequency of class F is estimated to be more than one half; the population frequency of class S is estimated to be more than one third; and the population frequency of class M is estimated to be very small (0% in the Two-Cost Treatment).[42] Error rates

[39]The likelihood function is also implicitly conditional on the realized matching of subjects.

[40]This function is continuous and differentiable in θ for every B and T and has a strict maximum at the true θ. The parameter space is obviously compact. All other technical requirements (as given in theorems 13.1 and theorem 13.2 in Wooldridge (2001)) hold so that the ML estimator is consistent and asymptotically normal (for asymptotic normality the true θ also needs to be interior).

[41]In the Two-Cost Treatment, the estimate of p_M is on the boundary of the parameter space. We do not compute the standard error for this estimate since the standard error does not have the usual interpretation in terms of confidence intervals. The standard errors for the elements of θ which are not on the boundary are approximate, since they are computed by estimating a restricted model in which p_M is set equal to 0. Estimated standard errors in the Alternating One-Cost Treatment should be treated with caution, since the estimate of p_M is only 1.58 (rather than at least 1.96) estimated standard errors from the boundary of the parameter space.

[42]Our estimates are robust to the specification of errors following an error. We estimated a model in which, after a subject makes a mistake, she chooses to invest or not invest in each subsequent round with probability one half. The estimate of θ in the Alternating One-Cost Treatment changed

are not very high, and parameters are nearly the same across treatments.

We wish to demonstrate that the population contains both class F and class S subjects, but formal testing is complicated by boundary issues and the possibility that test statistics are not asymptotically normal. However, we are able to test, using a likelihood-ratio test, the hypotheses that (i) $p_F = 0.25$ in the Two-Cost Treatment, (ii) $p_S = 0.2$ in the Two-Cost Treatment, (iii) $p_F = 0.25$ in the Alternating One-Cost Treatment, and (iv) $p_S = 0.16$ in the Alternating One-Cost Treatment.[43] The corresponding p-values are 0.001 or less for (i), (iii) and (iv); the p-value for (ii) is 0.071. Note that the p-value for (ii) would be lower if we could test the hypothesis that $p_S = 0$ in the Two-Cost Treatment. The estimates of the parameter vector θ are very similar in both treatments. We cannot reject (using a likelihood-ratio test) any of the hypotheses that $p_F/p_M/p_S/\varepsilon/p_F\&p_M\&p_S/p_F\&p_M\&p_S\&\varepsilon$ are equal in both treatments (p=0.754/0.341/0.675/0.235/0.635/0.532).

Let us summarize:

Result 7 *(i) More than half the population in each treatment is estimated to be F; slightly more than one third is estimated to be S; and only a small minority is estimated to be M. The estimates of p_F and p_S are both statistically different from 0.*[44]

(ii) The estimates of θ are very similar across treatments. Any differences are statistically insignificant.

How well do the various behavioral theories explain our estimation results? First consider the cursed equilibrium framework. There are several reasons why symmetric cursed equilibrium (or NE, which is a special case) is probably not driving our maximum likelihood estimates.[45] Referring back to Table 3, we see that no level of the cursedness parameter, χ, can explain the simultaneous presence of F and S subjects. Thus, symmetric cursed equilibrium is inconsistent with result 7. Also,

only negligibly. In the Two-Cost Treatment, the estimate of p_F decreased by 0.085, and the estimate of p_M increased by 0.073.

[43] These values are chosen so that they are at least 1.96 (estimated) standard errors from the boundary of the parameter space.

[44] At the 10% level in the case of the estimate of p_S in the Two-Cost Treatment.

[45] We consider here a common cursedness parameter for all subjects. When we introduce our estimates for Treatment 3, we will discuss the possibility of heterogeneity.

in the Alternating One-Cost Treatment, the strategy, M, is played with probability at least 0.4916, for any level of the cursedness parameter in which F is also played. This restriction is inconsistent with our estimate of p_M, which is significantly different from 0.496 ($p < .01$). Furthermore, the low level of χ required to explain the high frequency of F requires different behavior across treatments.[46] In particular, for fixed small χ we would expect a higher fraction of M subjects in the Alternating One-Cost Treatment.

Next, consider the level-k framework. Recall that, in both treatments, L_1 plays S and L_2 plays F. According to our estimates, the majority of the population is indeed F or S. The estimates of p_F and p_S are nearly the same across treatments, which is what one would expect if the proportions of L_1 and L_2 players in the population are stable across the treatments. Therefore, our estimates based on the Two-Cost and Alternating One-Cost Treatments are consistent with the level-k framework.

Finally, consider the framework in which each subject uses a rule of thumb prescribing either F, M or S. Because the Two-Cost and Alternating One-Cost Treatments are essentially the same, in terms of the nature and difficulty of the insights discussed above, one would expect to see nearly identical behavior across the two treatments. This is indeed the case. Therefore, our estimates based on the Two-Cost and Alternating One-Cost Treatments are consistent with the framework in which subjects use a rule of thumb, based on the various insights discussed above.

5.3 Treatment 3

To distinguish better between the possible explanations of behavior, and especially between the belief-based theories and insight-based rules of thumb, we conducted an additional treatment (one session, 28 participants). This treatment, which we call Treatment 3, is almost the same as the Alternating One-Cost Treatment, with the sole differences being that the high investment cost is 5.7, rather than 6.5, and the

[46] We performed separate maximum likelihood estimations of χ, for a model in which subjects play a symmetric cursed equilibrium with errors. In particular, each time a subject has to make a decision, she chooses the strategy prescribed by the cursed equilibrium with probability $1 - \varepsilon$ and makes an error with probability $\varepsilon \in [0, 0.5]$. The estimate of χ is 0 in the Alternating One-Cost Treatment, and 0.282 in the Two-Cost Treatment. The difference across treatments is troubling, and the low levels suggests that cursed equilibrium adds little explanatory power over NE.

discount factor is 0.8, rather than 0.9.[47] With the new parameters, a type $(1, H)$ subject has a dominant strategy to invest in round 1. The expected profits from investing in round 1 are greater than the expected profits of waiting, even if waiting would reveal the other subject's type. Therefore, the strategy F is never the optimal strategy for a risk-neutral subject within the expected utility framework, regardless of her beliefs. If behavior in our Two-Cost and Alternating One-Cost Treatments is driven by cursed equilibrium (even allowing each subject to have her own χ), level-k beliefs, or some other belief-based behavioral theory such as Quantal Response Equilibrium (see McKelvey and Palfrey (1995) and (1998)), the resulting maximum likelihood estimation for Treatment 3 should show a collapse of p_F.[48]

Suppose behavior is driven by rules of thumb. Some subjects play S because they do not acquire the insight that investment by the other subject is good news for them. Some subjects play M because either (i) they acquire the insight that investment by the other subject is good news, but not the insight that there is a tradeoff between the cost of waiting and the information gained by waiting, or (ii) they acquire both of the above insights, but simply resolve the tradeoff in favor of investing in round 1. Some subjects play F, because they acquire both of the above insights, but resolve the tradeoff in favor of gathering information by waiting. The new parameters in Treatment 3 should not affect the difficulty of acquiring these insights. Due to the reductions of the high investment cost and the discount factor, there may be some subjects who would resolve the tradeoff in favor of M in Treatment 3 but F in the Two-Cost and Alternating One-Cost Treatments. However, one would expect to see a significant proportion of subjects who continue to play F, because they do not explicitly perform the computation to determine that F is dominated.

Table 10 shows the estimates of θ for all three treatments along with estimated standard errors.[49] Amazingly, the estimate of p_F in Treatment 3 is very similar to that

[47] To maintain roughly the same expected earnings as before, the exchange rate was changed to \$0.50/ECU.

[48] The symmetric cursed equilibrium is to play M for $\chi < 0.75$, and S for $\chi > 0.75$. In any cursed equilibrium, symmetric or not, F is never played. For all $k > 0$, a subject with level-k beliefs plays M.

[49] Estimated standard errors in Treatment 3 should be treated with caution, since the estimate of p_M is 1.49 (rather than at least 1.96) estimated standard errors from the boundary of the parameter space. See footnote 41.

	Two-Cost	**Alternating One-Cost**	**Treatment 3**
p_F	0.607 (0.102)	0.567 (0.083)	0.574 (0.115)
p_M	0 N/A	0.092 (0.058)	0.287 (0.118)
p_S	0.393 (0.102)	0.341 (0.078)	0.139 (0.093)
ε	0.192 (0.011)	0.176 (0.01)	0.198 (0.013)
log-likelihood	-757.5402	-1002.5052	-600.378

Table 10: Maximum Likelihood Estimates - All Treatments

	Two-Cost & Alternating One-Cost	Two-Cost & Treatment 3	Alternating One-Cost & Treatment 3	All three treatments
p_F	0.754	0.824	0.959	0.95
p_M	0.341	0.04	0.098	0.086
p_S	0.675	0.074	0.089	0.143
ε	0.235	0.714	0.139	0.268
p_F, p_M, p_S	0.635	0.082	0.124	0.173
p_F, p_M, p_S, ε	0.532	0.159	0.091	0.17

Table 11: Hypotheses tests - p-values.

in the initial two treatments. The point estimate is actually slightly higher than in the Alternating One-Cost Treatment. The estimate of p_M is higher than in the initial two treatments and the estimate p_S is correspondingly lower.[50] Using likelihood-ratio tests, we can test hypotheses about whether certain elements of θ are the same in different pairs of treatments, as well as across all three treatments. Entry (i, j) in Table 11 shows the resulting p-value, under the null hypothesis that the parameters in row i are restricted to be the same across all treatments in column j.

Any belief-based theory would imply large differences across treatments, and in particular, no F subjects in Treatment 3. Table 10 shows that the estimates of p_F are virtually indistinguishable across treatments (high p-values), and we can reject the hypothesis that $p_F = 0.25$. This supports the view that behavior in our experiment is, to a large extent, driven by boundedly rational rules of thumb, rather than explicit

[50]We performed the same robustness check that we did for the initial two treatments (see footnote 42), where we assume that after a subject makes a mistake, she chooses to invest or not invest in each subsequent round with probability one half. The estimate of θ changed only negligibly.

belief formation about the behavior of the other subject. Moreover, among type $(1, H)$ subjects who perceive a tradeoff between the costs and benefits of waiting, the proportion deciding to wait does not change as we vary the high-cost parameter.[51]

Result 8 *Any differences in the estimates of p_F in all three treatments are both economically and statistically insignificant. In Treatment 3, we can reject the hypothesis that p_F is as low as 0.25.*

Overall, there is no strong evidence of statistically significant differences in behavior across treatments. The p-values below 0.10 are due to higher estimates of p_M and lower estimates of p_S in Treatment 3. These higher estimates for p_M and lower estimates for p_S are primarily driven by the fact that type $(0, L)$ subjects in Treatment 3 invested 70% of the time after history $\{1\}$, versus 37% in the Two-Cost Treatment and 53% in the Alternating One-Cost Treatment. These differences seem anomolous to us. It is particularly hard to see why changing the high investment cost would affect behavior when the investment cost is low (Alternating One-Cost Treatment vs. Treatment 3), so we do not attach much significance to the different estimates of p_M and p_S.

Our maximum likelihood estimation assumes that the strategy classes F, M, and S are drawn from the population at the beginning of the experiment, and do not evolve as the trials progress. This specification was made for simplicity, and because learning issues are not our main focus. However, there seems to be some learning going on, which sheds light on our interpretation of behavior as rules of thumb. Random effects probit estimation is used to study the effect of the trial number on the probability that a type $(1, H)$ subject invests in round 1. In the Two-Cost Treatment, the marginal effect is -0.018 (standard error $= 0.006$, $p = 0.002$). In the Alternating One-Cost Treatment, the marginal effect is -0.018 (standard error $= 0.005$, $p = 0.001$). In Treatment 3, the marginal effect is 0.001 (standard error

[51]Our organizing explanation of subjects' rules of thumb is based on some subjects acquiring the insight of a tradeoff between the costs and benefits of waiting. If so, their propensity to resolve this tradeoff in favor of waiting, when their type is $(1, H)$, seems to be "hard wired" in the sense that the high-cost and discount parameters do not seem to matter. We cannot rule out the explanation that subjects who play F do so based on some other rule of thumb, such as only investing if that is the obvious choice.

$= 0.007$, $p = 0.89$). The negative marginal effect in the Two-Cost and Alternating One-Cost Treatments indicates that subjects are "learning" to wait and observe the behavior of the other subject.[52] This learning moves behavior towards the NE in the Two-Cost Treatment, but moves behavior beyond and away from NE in the Alternating One-Cost Treatment. This learning could be due to a "Eureka" effect, where some subjects suddenly acquire the insight that waiting gives them useful information about the other subject. Why is learning absent in Treatment 3 (the estimated marginal effect is insignificant and of the wrong sign)? Perhaps while some subjects acquire the insight that waiting gives them useful information, other subjects acquire the additional insight that the benefits are not adequate to compensate for the discounting.[53] These two opposing effects may offset each other.

Treatment 3 addresses the potential criticism that the incentives of a type $(1, H)$ subject are weak, so that drawing conclusions about behavior is problematic. In the Two-Cost and Alternating One-Cost Treatments, the differences in the ex ante expected payoff of playing F, M, and S are quite small (both in the NE and given the empirical frequencies of play).[54] In Treatment 3, the expected payoff of playing $F/M/S$, given NE beliefs, is 1.32/1.326/1 ECU. However, the expected payoff of playing $F/M/S$, given the empirical frequencies is 1.141/1.29/1 ECU. Therefore, the advantage of M over F in Treatment 3 is quite substantial. Over 24 trials, the expected profit gain of playing M rather than F is 3.576 ECU or \$1.79.

5.4 Personal Characteristics as Determinants of Behavior

In this section we investigate whether subject's personal characteristics (GPA and SATs in particular), affect their earnings in the experiment.

[52] This effect is large in the Two-Cost and Alternating One-Cost Treatments. The predicted probability of investment in round 1 by a type $(1, H)$ player is higher in trial 1 than trial 24, by about 0.43.

[53] After all, even if the other subject is revealed to have the low common-value signal, expected losses are only 0.7 ECU in Treatment 3, while it is 1.5 ECU in the other treatments. Also the other subject may choose to wait with the high common-value signal, thereby weakening the inference.

[54] The expected payoff of playing $F/M/S$, given the empirical frequencies of play, is 1.071/1.05/1 ECU in the Two-Cost Treatment and 1.084/1.085/1 ECU in the Alternating One-Cost Treatment. The expected payoff of playing $F/M/S$, given NE beliefs, is 1.073/1.071/1 ECU in the Two-Cost Treatment and 1.142/1.142/1 ECU in the Alternating One-Cost Treatment.

GPA	**SAT**	$\boldsymbol{\pi}_{BR}$	$\mathbf{D}_{\textbf{Two-Cost}}$	$\mathbf{D}_{\textbf{Alt}}$	$\mathbf{D}_{\textbf{Treatment 3}}$
0.1032**	-	0.8337***	-0.4011***	-0.3336**	-0.4176***
(0.0438)	-	(0.0113)	(0.1408)	(0.141)	(0.1452)
-0.03	0.0008139***	0.8402***	-0.9679***	-0.8884***	-0.9311***
(0.072)	(0.0002856)	(0.0148)	(0.2888)	(0.2969)	(0.3003)

Table 12: Regression Results for Earnings (*/**/*** indicates significance at the 10%/5%/1% level.)

Table 12 shows the results of two random effects regressions. The first one regresses earnings in each period on a subject's GPA, the earnings a subject would have made had she played the best response to the empirical frequencies (π_{BR}), and treatment dummies. The second regression is the same as the first with the sole difference that it also includes subjects' SAT scores as a right-hand-side variable. The earnings a subject would have made had she played the best response to the empirical frequencies are included in the regressions in order to eliminate any noise in earnings due to a favorable combination of a subject's cost signal and behavior of the other subject.[55]

The first regression seems to suggest that a subject's GPA is positively correlated with her earnings in our experiment. However, this effect is wiped out when one includes SATs in the regression. The effect of a higher SATs score on a subject's earnings comes out strongly significant. This effect is also economically significant: it implies a difference in earnings per period of \$0.244 between a subject who is one standard deviation above the mean SAT score and a subject who is one standard deviation below it. This translates into \$5.849 over 24 periods.

6 Concluding Remarks

To summarize our main results, we find that subjects are more likely to invest as their signals become more favorable, even for the subtle comparison between type $(0, L)$ and type $(1, H)$. Subjects overinvest relative to the Nash benchmark. To the extent that subjects with the low common-value signal invest in round 1, a negative informational

[55]We also used another, more direct, method to control for these factors. In particular, we included a dummy variable for each possible combination of a player's cost, signal, whether the other person invested in round 1 or not and the treatment she is in (48 dummies in total). This yields very similar results regarding the effect of GPA and SATs.

externality is created, and to the extent that subjects with the high common-value signal invest in round 1, a positive informational externality is created. When we compare behavior with the theoretical predictions in the Two-Cost Treatment, the negative externality is balanced by a "theoretically excessive" positive externality, so a subject best responding to actual play receives the same profit that would be received if everyone were playing Nash. In the Alternating One-Cost Treatment, the positive externality is no longer excessive, so best responding to actual play yields lower profits than what would be received if everyone were playing Nash. This difference across treatments is due entirely to difference in the theoretical predictions, because we cannot reject the hypothesis that there is no treatment effect on behavior or profits.

Maximum likelihood estimates for our two main treatments indicate strong evidence of both S and F subjects in the population, which is inconsistent with symmetric cursed equilibrium. Level-k beliefs can account for these estimates, due to the flexibility to allow for subject heterogeneity.[56] We can also account for these estimates if subjects choose rules of thumb, based on insights about how to understand the game (hindsight and foresight). To separate these explanations, we introduce Treatment 3, in which F is strictly dominated and is inconsistent with any theory of best responding to beliefs. We find that the proportion of F subjects does not decline significantly, and remains above 50%.

In conclusion, we do not want to discredit the belief-based theories. Cursed equilibrium formalizes the notion that subjects do not fully draw inferences about others' types from their behavior. Level-k beliefs allow for heterogeneous beliefs about how sophisticated the other players are, but requires best responding to those beliefs. These theories provide important generalizations to Bayesian Nash equilibrium, and can account for many behavioral anomalies. They are particularly plausible when the primary task is to figure out what strategies the other players will choose, and figuring out a best response is fairly easy. In our context, the difficulty in forming best responses adds a new element.

[56]We feel that a better comparison would be to some notion of asymmetric cursed equilibrium that allows for heterogeneous subjects. We did not go to the considerable trouble of considering such a concept, because Treatment 3 would rule it out in any case.

7 Appendix: Derivation of Symmetric Cursed Equilibrium

In any cursed equilibrium, no matter what a player believes about the strategies being played by others, a type $(0, H)$ never invests, and a type $(1, L)$ invests in round 1. It is easy to see that the only viable possibilities for a type $(0, L)$ are N and W, and that the only viable possibilities for a type $(1, H)$ are W and 1. Denote the probability that a type $(0, L)$ chooses N as r, and denote the probability that a type $(1, H)$ chooses 1 as q.

Two-Cost Game: Given the above probabilities, one can compute the probability that the other player chooses 1, conditional conditional on having the high common-value signal, as follows

$$pr(1|X_i = 1) = \frac{.7^2 + .3^2}{2}(1 + q) = .29(1 + q).$$

Similarly, we have

$$\begin{aligned} pr(W|X_i &= 1) = .5 - .29q - .21r, \\ pr(N|X_i &= 1) = .21(1 + r), \\ pr(1|X_i &= 0) = .21(1 + q), \\ pr(W|X_i &= 0) = .5 - .21q - .29r, \\ pr(N|X_i &= 0) = .29(1 + r). \end{aligned}$$

With a cursedness parameter χ, the objective of a type $(1, H)$ is 0.5 if she invests in round 1, and is

$$.9\chi[.29(1 + q)(.5)] + .9(1 - \chi)[(.7)(.7)\frac{1 + q}{2}(3.5) + (.3)(.3)\frac{1 + q}{2}(-6.5)] \quad (2)$$

if she chooses W. Expression (2) is strictly increasing in q and decreasing in χ. For $\chi < \frac{17}{756}$, expression (2) is greater than 0.5 for all q, so we are at a corner solution with $q = 0$. For $\frac{517}{756} < \chi$, expression (2) is less than 0.5 for all q, so we are at a corner solution with $q = 1$. For $\frac{17}{756} < \chi < \frac{517}{756}$, we solve for q by setting expression (2) equal to 0.5, yielding $q^{2\text{cost}}$.

The objective of a type $(0, L)$ is 0 if she chooses N, and is

$$.9\chi[.21(1+q)(-.5)] + .9(1-\chi)[(.3)(.7)\frac{1+q}{2}(6.5) + (.7)(.3)\frac{1+q}{2}(-3.5)] \quad (3)$$

if she chooses W. For all q and r, expression (3) is positive for $\chi < \frac{3}{4}$, and negative for $\chi > \frac{3}{4}$. Therefore, except for the knife-edge case, $\chi = \frac{3}{4}$, we are always at a corner solution. A type $(0, L)$ must choose W for $\chi < \frac{3}{4}$, and N for $\chi > \frac{3}{4}$. Combining these choices for each type yields the type-dependent strategies for the Two-Cost Game given in Table 3.

Alternating One-Cost Game: Consider first the low-cost game, Game 2. We know that a type $(1, L)$ invests in round 1. For a type $(0, L)$, the probabilities that the other player chooses 1, W, and N are

$$\begin{aligned} pr(1|X_i &= 0) = .42, \\ pr(W|X_i &= 0) = .58(1-r), \\ pr(N|X_i &= 0) = .58r. \end{aligned}$$

The objective of a type $(0, L)$ is 0 if she chooses N, and is

$$.9\chi[.42(-.5)] + .9(1-\chi)[(.3)(.7)(6.5) + (.7)(.3)(-3.5)] \quad (4)$$

if she chooses W. Therefore, except for the knife-edge case, $\chi = \frac{3}{4}$, the cursed equilibrium of Game 2 is in pure strategies. A type $(0, L)$ must choose W for $\chi < \frac{3}{4}$, and N for $\chi > \frac{3}{4}$.

Now consider the high-cost game, Game 3. We know that a type $(0, H)$ chooses N. For a type $(1, H)$, the probabilities that the other player chooses 1, W, and N are

$$\begin{aligned} pr(1|X_i &= 1) = .58q, \\ pr(W|X_i &= 1) = .58(1-q), \\ pr(N|X_i &= 1) = .42. \end{aligned}$$

The objective of a type $(1, H)$ is 0.5 if she invests in round 1, and is

$$.9\chi[.58q(.5)] + .9(1-\chi)[(.7)(.7)q(3.5) + (.3)(.3)q(-6.5)] \tag{5}$$

if she chooses W. Expression (5) is strictly increasing in q and decreasing in χ. For $\frac{517}{756} < \chi$, expression (5) is less than 0.5 for all q, so we are at a corner solution with $q = 1$. For $\chi < \frac{517}{756}$, we solve for q by setting expression (5) equal to 0.5, yielding q^{alt}. Combining these choices for each type across Game 2 and Game 3 yields the type-dependent strategies for the Alternating One-Cost Game given in Table 3.

8 Appendix: Derivation of the Likelihood Function

Let $j_i \in \{F, M, S\}$ denote subject i's strategy class. Label all trials, $m = 1, ..., M$, and let $m(1)$ and $m(2)$ be the two subjects in trial m, where $m(1)$ is the subject with the lower identification number. The likelihood function (suppressing the dependence on the realized types, realized matchings, and θ) is given by

$$\Pr(B) = \sum_{j_1,...,j_n} p_{j_1} \cdots p_{j_n} \Pr(B|j_1, ..., j_n). \tag{6}$$

Errors are independent, so behavior in one trial, conditional on the strategy classes of the subjects in that trial, is independent of behavior in any other trial. Therefore, we have

$$\Pr(B|j_1, ..., j_n) = \prod_{m=1}^{M} \Pr(B_{m(1)}, B_{m(2)}|j_{m(1)}, j_{m(2)}). \tag{7}$$

Now, we claim that, for subjects 1 and 2 in a particular trial (this is without loss of generality), we can write

$$\Pr(B_1, B_2|j_1, j_2) = \Pr(B_1|j_1, B_2) \Pr(B_2|j_2, B_1). \tag{8}$$

To verify the claim, let B_i^r denote the behavior of subject i during round r (one if the subject invests during that round, zero otherwise). Then we have (sometimes suppressing the dependence on j_1 and j_2)

$$\begin{aligned}
\Pr(B_1, B_2|j_1, j_2) &= \Pr(B_1^1, B_2^1)\Pr(B_1^2, B_2^2|B_1^1, B_2^1)\Pr(B_1^3, B_2^3|B_1^1, B_2^1, B_1^2, B_2^2)\cdot \\
&\quad \Pr(B_1^4, B_2^4|B_1^1, B_2^1, B_1^2, B_2^2, B_1^3, B_2^3) \\
&= \Pr(B_1^1)\Pr(B_2^1)\Pr(B_1^2|B_1^1, B_2^1)\Pr(B_2^2|B_1^1, B_2^1)\Pr(B_1^3|B_1^1, B_2^1, B_1^2, B_2^2)\cdot \\
&\quad \Pr(B_2^3|B_1^1, B_2^1, B_1^2, B_2^2)\Pr(B_1^4|B_1^1, B_2^1, B_1^2, B_2^2, B_1^3, B_2^3)\cdot \\
&\quad \Pr(B_2^4|B_1^1, B_2^1, B_1^2, B_2^2, B_1^3, B_2^3) \\
&= \Pr(B_1|j_1, j_2, B_2)\Pr(B_2|j_1, j_2, B_1).
\end{aligned}$$

The behavior of one subject in a trial depends on the other subject's behavior but not on the other subject's strategy class (given the other's behavior), so the claim follows.

From (7) and (8), we have

$$\begin{aligned}
\Pr(B|j_1, ..., j_n) &= \prod_{m=1}^{M} \Pr(B_{m(1)}|j_{m(1)}, B_{m(2)})\Pr(B_{m(2)}|j_{m(2)}, B_{m(1)}) \\
&= \prod_{i=1}^{n} \Pr(B_i|j_i, B_{-i}). \qquad (9)
\end{aligned}$$

Substituting (9) into (6), we have

$$\begin{aligned}
\Pr(B) &= \sum_{j_1,\dots,j_n} \left[p_{j_1}\cdots p_{j_n}\prod_{i=1}^{n}\Pr(B_i|j_i, B_{-i})\right] \\
&= \sum_{j_1,\dots,j_n} \left[\prod_{i=1}^{n} p_{j_i}\Pr(B_i|j_i, B_{-i})\right] \\
&= \prod_{i=1}^{n}\left[\sum_{j\in F,M,S} p_j \Pr(B_i|j, B_{-i})\right] \\
&= \prod_{i=1}^{n}\Pr(B_i|B_{-i}),
\end{aligned}$$

which is what we wanted to show.

9 Appendix: Instructions for the Two-Cost Treatment

INSTRUCTIONS – Two Player Trials

This is an experiment on decision-making in investment markets. The National Science Foundation has provided funds for conducting this research. The instructions are simple, and if you follow them carefully and make good decisions, you may earn a CONSIDERABLE AMOUNT OF MONEY, which will be PAID TO YOU IN CASH at the end of the experiment.

Every participant in the experiment is guaranteed a payment of at least $5, independent of their performance in the experiment. All monetary values in the experiment, such as investment costs, investment returns, and account balances, are written in experimental currency units (EC). Your balance of ECs at the end of the experiment will be converted to US dollars at the exchange rate of $0.60 for each EC. Because your decisions may involve losses, we will endow you with a starting cash balance of 20 ECs. Your gains (losses) during the experiment will be added to (subtracted from) your cash balance. However, if your cash balance falls below zero, you will no longer be allowed to continue. At the end of the experiment you will receive in cash your end of experiment balance of ECs converted to US dollars, or $5, whichever is greater.

1. In this experiment we will create a sequence of market trials. In each given market trial, the participants will act as potential investors. Each potential investor will have to decide whether, and when, s/he wishes to invest, based on the information s/he is provided (and which we will explain later).

2. In the experimental session today we will have between 20-25 market trials. Each market trial has several rounds. The initial round is round 1, the next is round 2, and so on. In each round you and the other potential investor in your market trial will have to decide (simultaneously) whether to invest in that round or not. The decision to invest is irreversible. Any potential investor who has not yet invested will be told whether the other potential investor has invested, and if so, during which round of that trial.

3. In each trial, the market in which you are a potential investor has just one

more potential investor besides yourself. In a typical session we will recruit (about) 20 students. The computer will randomly match 10 pairs out of the 20 students. Each such matched pair, including the one that you are in, constitutes a separate market trial that has no relation to the other nine markets. A given market trial keeps the same matched students over the several rounds of that market trial. However, after the market trial is over, the computer randomly rematches students to form a new set of market trials. This matching procedure makes it very unlikely that you will be matched with the same student from one trial to the next.

4. **The structure of information**.

Information about investment cost: Each potential investor will know, before each market trial starts, his/her investment cost for that trial. There are two possible levels of investment cost: low cost, CL=3.5 and high cost, CH =6.5. Each potential investor will be assigned one of the two cost levels with equal probability (1/2). In other words, in your market trial, you will know your investment cost, and that the investment cost of the other potential investor is equally likely to be either 3.5 or 6.5.

Information about investment gross returns: The computer assigns a gross return to every market trial. The gross return remains the same for all rounds of the same market trial, and is completely uncorrelated with your investment cost. The computer randomly picks the gross return to be either 10 or 0, with equal probabilities. Once the gross return is picked, high or low, it is the same for both potential investors, and it remains the same for all rounds of the same trial. You will NOT observe whether the gross return for that trial is high or low. Instead, each potential investor will be given his/her own signal, which takes the value of either 0 or 1. Signals are 70% accurate, in the following sense:

If the gross return is 10, you have a 70% chance of observing signal 1 and a 30% chance of observing signal 0. If the gross return is 0, you have a 70% chance of observing signal 0 and a 30% chance of observing signal 1.

Each potential investor's signal is related to the gross return, but the computer randomizes separately for each potential investor, so the two signals can be different. For example, if the gross return is 10, there is a 49% chance that both signals are 1, there is a 42% chance that one signal is 1 and the other signal is 0, and there is a 9%

chance that both signals are 0.

The signal for each potential investor is chosen at the beginning of the trial and remains the same for all rounds of that trial. Each potential investor observes his/her own signal, but not the signal of the other potential investor in that trial. Observing your signal may help you better predict the likelihood that the gross return in your market trial is high or low.

Information about other investors in your market: You will NOT be told the signal of the other potential investor in your market trial. However, you will be informed about whether the other potential investor has invested, and if so, during which round. If this information reveals something about his/her signal, it could improve your decision about if and when to invest.

You are not allowed to reveal or discuss your information with other students or look at another student's screen (this will be strictly monitored and violators will be removed from the experiment).

5. **The structure of the game.**

The computer randomly matches you with another potential investor to form a market trial. Once you are assigned to a market trial, you privately observe your cost and your signal, which remain constant for that market trial. The other potential investor observes his/her cost and signal. In round 1, you are asked to decide if you wish to invest. If you do not invest in round 1, you are informed about whether the other potential investor invested in round 1, and you are asked if you wish to invest in round 2. If you have not invested by round 2, we move to round 3, and so on. Once you have decided to invest, there are no more decisions to make in that market trial. That is, an investment decision in a given trial is irreversible. You cannot disinvest or invest a second time. After two consecutive rounds in which no one in your trial invests, that trial is over.

In order to make good decisions, you must understand how your gains and losses are determined. This will be carefully explained below.

Once a market trial is over, the whole process starts again. The computer matches you to another potential investor to form a new market trial, you will be assigned an investment cost and a signal, etc.

Your screen will inform you of the trial number, and the round number within the trial.

How your gains (discounted net returns) or losses are determined.

If you invest, your gains from that trial are the discounted difference between the gross return and your investment cost. Let us illustrate what this means by using a simple example. Suppose that in the current market trial your investment cost is 3.5. If you decide to invest in round 1, then your gains are: 6.5 if the gross return is 10 (10 – 3.5 = 6.5) or – 3.5, a loss of 3.5, if gross return is 0 (0 – 3.5 = -3.5). Note that gains or losses in round 1 are not discounted; they are just the difference between the market gross return and your investment cost. For each round that you wait, your gains or losses are discounted by a factor of 0.9, as shown in the following table.

Discounted Net Returns when Cost is 3.5

Round that you Invest	If return is 10 (high)	If return is 0 (low)
1	6.5	−3.5
2	5.85	−3.15
3	5.26	−2.84

There are several important things to note here:

(i) If for whatever reason you have decided not to invest at all in a particular market trial, you will earn zero for that market trial.

(ii) You will not be told the actual gross return during a market trial. After each trial is over, the gross return is revealed and you will learn your discounted net gains or losses, which will be added to, or subtracted from, your cash balances.

(iii) It is up to you to decide if and when to invest. Clearly, your investment cost and your signal can affect your decision. Observing the activity of the other potential investor in your trial might indirectly yield useful information about your gross return, because his/her behavior might tell you something about his/her signal.

6. **Information on the computer screen.** Throughout the experimental session, the computer screen will show your ID number and current cash balances, in the upper left corner. The upper left corner of the screen will also remind you of the number of potential investors in each trial (2), the discount factor (0.9), and the "accuracy parameter" of your signal (70%).

At the beginning of each round of each market trial, you will see the number of the market trial, your cost of investment (either 3.5 or 6.5), and your signal (0 or

1). This information stays the same during the trial. In the middle of the screen, you will see the current round number. At the bottom of the screen, you will see a "history" of investment in previous rounds of that trial. (If it shows all zeros, no one has invested; if it shows a 1 under some round, the other potential investor invested during that round.)

You will have 25 seconds to think about whether to invest in that round. At that time, boxes marked "YES" and "NO" will appear, and you should mark a box to indicate whether you want to invest or not. Please make your choice within 5 seconds.

At the end of the market trial, you will see a screen that tells you the market trial number, your investment cost, your signal, the actual gross return, and your net discounted gains or losses from that trial. You will also see your current cash balance and your personal statistics from your previous trials. (If you are listed as investing in round -1, this means that you never invested during that trial.)

7. We will start the session with two practice "dry runs" that do not count towards your earnings, at which point we will stop and answer additional questions. At the end of the experiment, while we are calculating your earnings, we ask that you answer the short questionnaire on your computer.

8. Are there any questions?

References

[1] Anderson, L. R. and C. A. Holt (1997), "Information Cascades in the Laboratory," *American Economic Review*, 87, 847-862.

[2] Ball, S. B., M.H. Bazerman, and J.S. Carroll (1991), "An Evaluation of Learning in the Bilateral Winner's Curse," *Organizational Behavior and Human Decision Processes*, 48, 1-22

[3] Banerjee, A. V., "A Simple Model of Herd Behavior," *Quarterly Journal of Economics* 107 (3), August 1992, 797-817.

[4] Bazerman, M.H. and W.F. Samuelson (1983), "I Won the Auction But Don't Want the Prize," *Journal of Conflict Resolution*, 27, 618-634.

[5] Blecherman, B. and C.F. Camerer (1998), "Is There a Winner's Curse in the Market for Baseball Players?" mimeograph, Brooklyn Polytechnic University, Brooklyn, NY.

[6] Bikhchandani, S., D. Hirshleifer, and I. Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, Vol. 100 (1992), pp. 992-1026.

[7] Camerer, C. F. and D. W. Harless (1994), "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, Vol. 62, 1251-1289.

[8] Capen, E.C., R.V. Clapp, and W.M. Campbell (1971), "Competitive Bidding in High-Risk Situations," *Journal of Petroleum Technology*, 23, 641-653.

[9] Cassing, James and Richard W. Douglas (1980), "Implications of the Auction Mechanism in Baseball's Free agent Draft," *Southern Economic Review*, 47, 110-121.

[10] Chamley, C. and D. Gale (1994), "Information Revelation and Strategic Delay in a Model of Investment," *Econometrica*, Volume 62, Issue 5, 1065-1085.

[11] Charness, Gary and dan Levin, 2005, "When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity, and Psychological Affect", *American Economic Review,* Vol. 95, No. 4 (September 2005), 1300-1309.

[12] Costa-Gomes, M., V. Crawford, and B. Brosetta (2001), "Cognition and Behavior in Normal-Form Games: An Experimental Study," *Econometrica*, Vol. 69, 1193-1235.

[13] Crawford, V. P. and N. Iriberri (2005), "Level-k Auctions: Can a Non-Equilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?" UCSD Working Paper.

[14] Dessauer, John P., (1981) Book Publishing, New York: Bowker.

[15] Eyster, E. and M. Rabin (2005), "Cursed Equilibrium," *Econometrica*, Vol. 73, No. 5, 1623-1672.

[16] Fischbacher, U. (1999), "z-Tree–Zurich Toolbox for Readymade Economic Experiments–Experimenter's Manual," Working Paper Nr. 21, Institute for Empirical Research in Economics, University of Zurich.

[17] Garratt, R. and T. Keister (2006), "Bank Runs as Coordination Failures: An Experimental Study," UCSB and ITAM Working Paper.

[18] Hendricks, K., R. Porter, and B. Boudreau (1987), "Information and Returns in OCS Auctions, 1954-1969," *Journal of Industrial Economics*, 35, 517-42.

[19] Holt, C.A. Jr., and R. Sherman (1994), "The Loser's Curse and Bidder's Bias," *American Economic Review*, 84, 642-652.

[20] Kagel, J. H. and D. Levin (1986), "The Winner's Curse and Public Information in Common Value Auctions," *American Economic Review*, 76, 894-920.

[21] ______ and ______. (1991), "The Winner's Curse and Public Information in Common Value Auctions: Reply," *American Economic Review*, 81, 362-69.

[22] ______and______ (1999), "Common Value Auctions with Insider Information," *Econometrica*, 67, 1219-1238.

[23] ______ and _______ (2002), Common Value Auctions and the Winner's Curse, Princeton: Princeton University Press.

[24] Levin D., Kagel J. H. and J. F. Richard (1996), "Revenue Effects and Information Processing in English Common Value Auctions" *American Economic Review*, 86, 442-460.

[25] Levin, D. and J. Peck (2006), "Investment Dynamics with Common and Private Values," Ohio State University Working Paper.

[26] Levis, Mario (1990), "The Winner's Curse Problem, Interest Costs and the Underpricing of Initial Public Offerings," *Economic Journal*, 100, 76-89.

[27] McKelvey, R. D. and T. R. Palfrey (1995), "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior* 10, 6-38.

[28] McKelvey, R. D. and T. R. Palfrey (1998), "Quantal Response Equilibria in Extensive Form Games," *Experimental Economics*, 9-41.

[29] Mead, W. J., A. Moseidjord, and P.E. Sorensen (1983), "The Rate of Return Earned by Leases Under Cash Bonus Bidding in OCS Oil and Gas Leases," *Energy Journal*, 4, 37-52.

[30] ______, ______, and ______ (1984), "Competitive Bidding Under Asymmetrical Information: Behavior and Performance in Gulf of Mexico Drainage Lease Sales, 1954-1969," *Review of Economics and Statistics*, 66, 505-08.

[31] Nagel, R. (1995), "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 85, 1313-1326.

[32] Rock, Kevin (1986), "Why New Issues are Underpriced," *Journal of Financial Economics*, 15, 187-212.

[33] Roll, Richard (1986), "The Hubris Hypothesis of Corporate Takeovers," *Journal of Business*, 59, 197-216.

[34] Rubinstein, A. (1998), *Modeling Bounded Rationality*, The MIT Press (Cambridge, MA and London, England).

[35] Samuelson, W.F. and M.H. Bazerman (1985), "The winner's curse in bilateral negotiations," In V. Smith (Ed.), Research in Experimental Economics (Vol. 3), Greenwich, CT: JAI Press.

[36] Sgroi, D. (2003), "The Right Choice at the Right Time: A Herding Experiment in Endogenous Time," Experimental Economics 6, 159-180.

[37] Simon, H. A. (1972), "Theories of Bounded Rationality," in C. B. McGuire and R. Radner, eds., *Decisions and Organization*, 161-176, Amsterdam: North Holland.

[38] Stahl, D. and P. Wilson (1995), "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior* 10, 218-254.

[39] Wooldridge, J. (2001), *Econometric Analysis of Cross-Section and Panel Data*, The MIT Press, 1st Edition.

[40] Ziegelmeyer, A., K. B. My, J.-C. Vergnaud, and M. Willinger (2005), "Strategic Delay and Rational Imitation in the Laboratory," Working Paper.