Strategies and Beliefs in Sequential Games of Two-Sided Incomplete Information: An Experimental Study[1]

Dustin H. Tingley[2] and Stephanie W. Wang[3]

January 2008

[1]We gratefully acknowledge the financial support from Princeton's Research Program in Political Economy. We thank Scott Ashworth, Raymond Hicks, Jens Grosser, Kristopher Ramsay, Adam Meirowitz and Thomas Palfrey for helpful discussions. Paper prepared for presentation at the CESS-NYU 2008 Experimental Political Science Conference, and previously presented at the Princeton Lunchtime Political Methodology Meeting, and 2007 Economic Science Association Annual Meeting, Tuscon, Arizona.

[2]Department of Politics, Princeton University, Princeton, NJ 08540.

[3]Department of Economics, Princeton University, Princeton, NJ 08540.

Abstract

We investigate experimentally and theoretically a dynamic game of two-sided incomplete information where player types are drawn from a commonly known continuous distribution. Within the experiment we also elicited beliefs from players about their opponent's type using the strictly proper quadratic scoring rule. We implement a 2x2 crossed design: varying whether the game is played a single time versus repeated 3 times, and whether the cost of reaching one of the terminal nodes is high or low. Thus our data allows us to examine the relationship between beliefs and actions that are still not well understood in the empirical literature. We find that most beliefs and strategies are responsive to several of our treatments in the way predicted by theory, and that beliefs track departures from our theoretical expectations about strategy choice relatively closely. We also find that subjects by and large behave as Bayesians, but in certain instances appear to rationalize decisions with non-Bayesian beliefs. While our experiment and analysis is guided by predictions from a formal game-theoretic model, we also focus on behavioral themes such as the extent to which rational expectations form, motivated beliefs, and reputation building.

1 Introduction

Strategic interactions rarely take place in an information rich environment. Instead, decision makers not only have incomplete information about each other, but they often have strong incentives to keep this information private or have trouble credibly communicating their private information. This is the case for players that range from competing firms to countries and rebel factions. Accordingly, the game theoretic literature has developed rich modeling approaches that directly incorporate this uncertainty into equilibrium concepts such as the Perfect Bayesian Equilibrium (PBE). PBE specifies equilibrium strategy/belief pairs that highlight the interplay of both expectations and actions.

The theoretical literature has made great strides in understanding strategic interaction in situations with incomplete information. Experimentally we know relatively less. As discussed below, the vast majority of experimental work has been done with game forms that have at most only one-sided incomplete information. Yet if decision makers have incentives to not reveal their type or have trouble doing so, it seems likely that many dyadic strategic situations would be better characterized by two-sided incomplete information. In order to address this short coming in the experimental literature, we use incentivized laboratory experiments to test equilibrium predictions from a simple sequential game of two-sided incomplete information.

Our goal is to suggest a simple framework for analyzing updating within games, explore the intersection between strategies and beliefs in games of incomplete information, and characterize specific ways belief updating departs from equilibrium predictions. In order to investigate these topics we explore treatments along two dimensions in this study. First of all, we vary the payoff at one particular node in the game so as to generate very sharp differences in equilibrium behavior and beliefs. Secondly, we conduct both single play and repeat play (with three periods) sessions.

Repetition of a stage game can influence behavior, and provides us with a richer set of opportunities to explore updating within a dyad. Conducting these treatments generates a richer set of cases to observe beliefs and strategies.

Our framework for analyzing updating is to elicit beliefs about the private value of an opponent using a strictly proper quadratic scoring rule (where the score reduction is proportional to the square distance between the elicited expectation and the actual private value). We elicit these beliefs within the play of the game. This method allows us to test the strong assumptions that are often imposed upon the players' beliefs such as Bayesian updating (which includes no updating when no action is observed), rational expectations, and best response given beliefs. We find strong evidence that players are best responding given their beliefs, but also find suggestive evidence of conservatism in beliefs where players update less than what the equilibrium would predict. We also find some evidence that beliefs differ systematically from the prior when there was no opportunity for updating, and that these beliefs may help rationalize accompanying strategy choices. Furthermore, we find that beliefs do correspond well for the most part to the actual private values but we also document two interesting cases of departure from such rational expectations.

We find empirical support for the majority of comparative static predictions across the two payoff treatments. We do not observe the full realization of extreme equilibrium predictions (such as always Entry or Always Fight in one or the other payoff condition), but we observe behavioral shifts in the right direction for the most part. More importantly for the focus of this paper, beliefs appear to respond to strategy changes across our treatments. We explore the more significant departures from the PBE predictions in the results section. The small departures from equilibrium such as a very low level of non-Entry when 100% entry is predicted suggests that we might adopt a Quantal Response Equilibrium framework in the future to fit the data more precisely.

We also find some suggestive evidence of reputation building when comparing the behavior in the single play vs. repeat play treatments.

We survey the existing literature on experimental sequential games as well as the relevance of beliefs and expectations in Section 2. In Section 3, we describe the game form and solve for the Perfect Bayesian Equilibrium in the single play case. We derive the equilibrium cutpoints and expectations for both sets of payoff parameters and put forth some hypotheses based on the comparative statics. We present the experimental design as well as some summary statistics by treatment in Section 4. Section 5, the results section, provides tests of each of the hypotheses we enumerated in Section 3 as well as comparisons of behavior across the single vs. repeat play treatments. We conclude with findings and future work in Section 6.

2 Background

2.1 Experimental Sequential Games

Although there has been a wealth of theoretical work on sequential games with one-sided and two-sided incomplete information (See Fudenberg and Tirole, Ch. 9 for an overview), the corresponding experimental studies have lagged behind with a few exceptions (Camerer and Weigelt 1988, Jung et al. 1994, Bolton and Ockenfels 2007, Tingley and Walter 2007, Tingley and Walter 2008). In Camerer and Weigelt's lending game, there was only one-sided incomplete information about the borrower's type. They found support for the sequential equilibrium of reputation building given some restrictions on the belief that the lender held about the borrower's type. Jung et al. ran a chain store game with one-sided incomplete information about the monopolist's type (strong vs. weak) and found some deviations from sequential equilibrium predictions, but that reputations were developed by weak monopolists. Tingley and Walter investigate a repeated entry-deterrance game with one sided incomplete information by varying the number of total entrants across experimental treatments and by introducing a cheap talk component. To our knowledge, Carillo and Palfrey (C-T 2006) are the first to conduct a lab experiment on a sequential game with two-sided incomplete information.

In the C-T's compromise game, both players have a "strength" value drawn from an uniform distribution on $[0, 1]$. Player 1 first chooses to Fight or Retreat. If Player 1 chooses to Fight, the game ends and the player with the higher "strength" gets the win payoff (H) while the player with the lower "strength" gets the lose payoff (L). If Player 1 chooses to Retreat, then Player 2 must now choose between Fighting and Retreating. If Player 2 chooses to Fight, then again the player with the higher "strength" gets the win payoff (H) while the player with the lower "strength" gets the lose payoff (L). However, if Player 2 also chooses to Retreat, then both players receive the compromise payoff M ($L < M < H$). While the compromise game they study and the one we focus on do share the element of conflict, the extensive forms do differ significantly. In our experiment, another decision node for player 1 could be reached after player 2 chooses an action whereas the compromise game ends after player 2's decision. Furthermore, Carillo and Palfrey focus on the one-shot case while we look at repeated game/reputation effects in addition to the single play case.

2.2 Applications in Economics and Political Science

The specification of this sequential game reflects a class of strategic interactions between firms in the Industrial Organization literature. Neither firm has full information about the other firm's 'type' which deviates from the more common assumption of one-sided incomplete information in which the payoff to one firm in the different contingencies are known with certainty. This assumption would likely not hold in applications such as competition in a new industry for example. Firm 1 has to decide whether or not to enter an industry knowing that it may face resistance from Firm 2 who already has a foothold in the industry. Firm 2 then has to decide whether or not it wants to maintain the claim to the industry after Firm 1 enters knowing that it would face competition from Firm 1. Finally, if Firm 2 does decide to stay, Firm 1 has to decide whether the fight for the industry will be worth it. At each step along the way, both firms are faced with uncertainty about the other's 'type' or 'strength' which crucially impacts their payoffs from the competition for the industry should such a fight break out. It is natural then in this context to try to measure the beliefs the players would hold about each other's 'strength'. Also, unlike in the classic chain store setup, two long-lived firms are involved in the repeated game context rather than one incumbent (long-lived) and one entrant (short-lived). With the incorporation of belief elicitation into our experimental design, we are able to observe how the beliefs of *both* players evolve with regards to the uncertainty about the other player's 'strength'.

The structure of the game we study also features prominently in the international relations literature in political science. Here, countries have private types which are thought of as how much they value the outcome where "war" is reached. Because countries exist within an anarchic system, they may have a strong incentive to keep their private information from other states or be unable to credibly communicate such information. Thus conflict games with two-sided incomplete information are an important piece of understanding relations between countries. Several recent publications in this literature have gone on to develop statistical estimators that are derived from the game form itself, better linking the theoretical work to large-N analysis using dyadic data sets about conflict outbreak (Lewis and Schultz 2003, Wand 2006). A separate paper considers the behavior of these statistical estimators when applied to our experimental data set (Mukharjee et al. 2007). Thus our focus in this paper is analyzing the effects of our experimental treatments. While international relations scholars have used the type of model we consider directly, this type of strategic interaction has also been considered by legal scholars in considering decision to sue.

2.3 Belief Elicitation

Savage (1971) specified the general class of proper scoring rules for the elicited expectations of continuously distributed variables. We use one of the most common scoring rules within this class, an analog of the quadratic scoring rule for probabilistic forecasts. The squared distance between the expectation and the realized value of the variable is the deduction for this scoring method. A scoring rule is considered incentive compatible, or proper, if a forecaster cannot attain a higher expected score by reporting an expectation different from her true expectation. This rule is shown to be incentive compatible and has the added advantage of being relatively easy to understand for experimental subjects.

Previous experimental studies have elicited beliefs that can be categorized into three types: probabilistic beliefs about opponents' action in matrix games (Nyarko and Schotter 2002, Costa-Gomes and Weiszaecker 2006), probabilistic beliefs about the state of the world (McKelvey and Page 1990, Dominitz and Hung 2004), and beliefs about other players' actions where the action space is continuous. Examples of the third type include contributions in public goods games (Offerman et al. 1996; Croson, 2000) and investment in trust games (Dufwenberg and Gneezy 2000). We refer readers to Manski's (2004) comprehensive review on the theoretical importance of measuring beliefs as well as empirical attempts to do so in the survey and experimental context.

To our knowledge, ours is the first study to elicit beliefs using a strictly proper scoring rule from players about their opponents' private value as they are playing a sequential game with two-sided incomplete information. Our approach offers several innovations to the existing methods and gives us sharp tests of the assumptions and predictions about beliefs in sequential games without the common confounds.

The previous belief elicitation studies in which players played the game and stated beliefs about the opponent's action at the same time violate the "no-stakes" condition of Kadane and Winkler (1988), a necessary condition for truthful belief revelation. In our experiment, we are eliciting beliefs about the private value of each player's opponent rather than the action. The "no-stakes" condition is therefore satisfied because the player's payoff in the game depends on the action of the opponent and not on the opponent's private value except as it affects the action choice.[1] Secondly, by eliciting beliefs about the opponent's private rather than chosen action, we are gathering richer data that allows us to test equilibrium predictions of beliefs about types that previous studies have not been able to address. We are able to uncover behavioral deviations from Bayesian updating and rational expectation assumptions that coarse beliefs about actions would not have revealed.

Our experiment also uses a strictly proper scoring rule to elicit expectations about private values drawn from a continuous distribution whereas previous experiments that elicited beliefs about actions chosen from a continuous action space rewarded belief accuracy with non-incentive-compatible mechanisms or ones that were not transparently so for the sake of simplicity. Methodological studies have demonstrated the importance of using an incentive compatible belief elicitation mechanism (Palfrey and Wang 2007). We paid based a scoring rule that was both incentive compatible and understood by our subjects to encourage truthful belief revelation. We also elicit beliefs about each player's specific opponent rather than about the population of players in a particular role as some previous studies have done. Beliefs about behavior on the aggregate level often does not recognize the relevant individual heterogeneity in beliefs or beliefs that reflect

[1] We note that Player 1 may have an incentive to choose Enter to observe Player 2's action choice in order to state a more accurate expectation of Player 2's private value. However, this uncertain improvement in earnings from beliefs is likely not be comparable to the game payoffs and should not distort game play.

heterogeneity in behavior.

Another gap in the literature arises from the way beliefs are elicited in sequential games. Beliefs are elicited simultaneously from the players even in these games (Dufwenberg and Gneezy 2000, Charness and Dufwenberg 2006) so that the response of beliefs to observed actions cannot be ascertained. We fill this gap by eliciting the expectation each player holds about the other player's private value after observing the other player's action at the previous decision node.

3 Model and Hypotheses

We present the sequential game of two-sided incomplete information that we study experimentally and solve for the Perfect Bayesian Equilibrium (PBE) strategies and beliefs in the one-shot case. Although the PBE results apply generally to this class of games, we focus on the crisis bargaining interpretation (Lewis and Schultz 2003, Easery et al. 2006) in our exposition. There are two players, 1 and 2, where 1 hopes to possess a good currently held by 2. The extensive form of this game is depicted in Figure 1. The sequence of moves in this game is as follows.

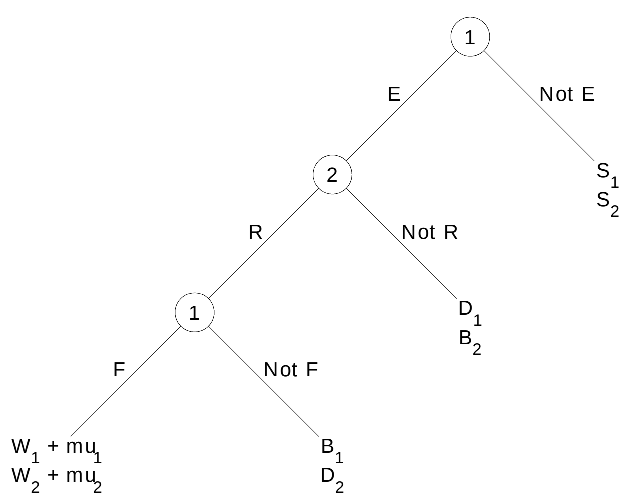

Figure 1: Sequential Game with Two-Sided Incomplete Information

First, Nature draws each player's type, μ_i, independently from the same continuous distribution with cdf $G(x)$. The distribution from which the types are drawn is common knowledge while the actual types of the respective players are private information. 1 then chooses whether or not to enter and challenge state 2 for the good (E or not E). If 1 chooses not to enter (not E), the status quo prevails and the players receive S_1 and S_2 respectively. If 1 does enter (E), then 2 decides whether or not to resist 1's demand for the good (R or Not R). If 2 does not resist, the crisis ends peacefully with 2 conceding the good to 1 and the players receive D_1 and B_2. If state 2 resists, then 1 must decide whether or not to fight (F or Not F). If 1 chooses not to fight (Not F), then it backs down from its threat and state 2 continues to retain its control over the good. The players receive B_1 and D_2 in this case. If 1 decides to fight and goes ahead with its demand for the good possessed by 2 (F), then the game ends with a war and the players receive $W_1 + \mu_1$ and $W_2 + \mu_2$ respectively. Players' types only affect their payoffs in the case of war so μ_i can be thought to capture each state's aversion towards (or propensity for) war outside of the payoffs already modeled within the game structure.

3.1 Perfect Bayesian Equilibrium

We solve for Perfect Bayesian Equilibria in both of our cost conditions using the standard backward induction method assuming risk neutrality of the players. Without loss of generality, we normalize the status quo payoffs to be zero for both players. At the last node, 1 must choose between F and not F. 1 will fight only if the payoff from F is greater than that from not fighting:

$$W_1 + \mu_1 > B_1 \tag{1}$$

Let $\mu_1^F = B_1 - W_1$ be the cutoff private value above which Player 1 would choose Fight. Define $\Pr(F|E) = 1 - G(\mu_1^F)$ as the probability Player 1 Fights given that she has already Entered.

Next consider Player 2's decision node where he must choose between R and Not R. He will only choose R if the expected payoff from Resisting, $\Pr(F|E)[W_2 + \mu_2] + [1 - \Pr(F|E)]D_2$, is greater than the payoff from Not Resisting, B_2:

$$\Pr(F|E)W_2 + \Pr(F|E)\mu_2 + D_2 - \Pr(F|E)D_2 > B_2 \tag{2}$$

Let $\mu_2^R = \frac{B_2 - \Pr(F|E)W_2 - D_2 + \Pr(F|E)D_2}{\Pr(F|E)}$ be the cutoff private value above which Player 2 would choose Resist. Define $\Pr(R|E) = 1 - G(\mu_2^*)$ as the probability that 2 will resist given that 1 has chosen to Enter.

Now consider 1's initial decision to Enter or Not Enter (E or Not E). She will choose E if the expected payoff from Entering, $[1 - \Pr(R|E)]D_1 + \Pr(R|E)\max[W_1 + \mu_1, B_1]$, is greater than the payoff from the status quo (Not Entering), $S_1 = 0$:

$$\max[W_1 + \mu_1, B_1] > \frac{-D_1}{\Pr(R|E)} + D_1 \tag{3}$$

If $B_1 > \frac{-D_1}{\Pr(R|E)} + D_1$, Player 1 would choose to Enter regardless of her private value according to (3). Let $\mu_1^E = \frac{-D_1}{\Pr(R|E)} + D_1 - W_1$ be the cutoff private value above which Player 1 would choose to Enter if it is not the case that Player 1 chooses to Enter regardless of private value. Define $\Pr(E) = 1 - G(\mu_1^E)I(B_1 < \frac{-D_1}{\Pr(R|E)} + D_1)$ as the probability that Player 1 enters at the initial decision node (where I is an indicator function).

We are now able to fully specify $\Pr(F|E) = \Pr(\mu_1 > \mu_1^F \,|\, \max[W_1 + \mu_1, B_1] > \frac{-D_1}{\Pr(R|E)} + D_1)$. Using the definition of conditional probability then simplifying, we can rewrite this as $\Pr(F|E) = \frac{\Pr[(\mu_1 > \mu_1^F) \cap (\mu_1 > \mu_1^E)]}{\Pr(\mu_1 > \mu_1^E)}$.

Equations 1, 2, and 3 allow us to solve for the equilibrium cutpoints players 1 and 2 use at each of the decision nodes. Given the cutpoints, we can also find the equilibrium belief that each player has about the other player's type after observing each possible action. $\Pr(E)$, $\Pr(R|E)$, and $\Pr(F|E)$ yield the equilibrium probabilities of each node in the extensive game being reached. We now solve for the equilibrium cutpoints, beliefs, and probabilities in the specific games we implement in the experimental sessions. Both players' strategies hinge crucially on the value of B_1, the cost for Player 1 of Not Fighting and this is the parameter we vary in our experimental treatments.

The players' type are drawn from a logistic distribution with mean 0 and variance 1. The cdf is $G(x) = \frac{e^x}{1+e^x}$ with a corresponding pdf of $g(x) = \frac{e^x}{(1+e^x)^2}$.[2] In all sessions, the payoffs were $S_1 = S_2 = 0$, $D_1 = D_2 = W_1 = W_2 = 0.5$, and $B_2 = -0.3$.[3] In the low cost of Player 1 backing down treatment, $B_1 = -0.1$ while $B_1 = -0.3$ in the high cost treatment. We will look for Perfect Bayesian Equilibria in the two cost treatments that have very sharp and different predictions for outcomes.

3.1.1 Low Cost: $B_1 = -0.1$

In the low cost condition, we look for the Perfect Bayesian Equilibrium in which Player 1 always enters then decides whether or not to Fight subject to the incentive compatibility constraint.

From (1), we know that 1 chooses F if $\mu_1 > B_1 - W_1 = -0.6$. With this cutpoint,

$$\Pr(F|E) = 1 - G(\mu_1^F) = \int_{-0.6}^{\infty} \frac{e^x}{(1+e^x)^2} dx = 0.65$$

[2] We use this distribution for several reasons. First, it allowed us to generate data that matches on to assumptions made by a statistical estimator we test in a seperate paper and is slightly easier to work with analytically than exponential class distributions. Second, like the normal distribution, it has easily understood properties of centering around 0 and symmetry. Finally, its continuity makes the application of our scoring rule straightforward. We are preparing to run a set of experiments with a uniform distribution, and are generally interested in the intersection between the structure of information in the game and distributional properties of the information source.

[3] These payoffs represent a slight departure from the standard crisis bargaining interpretation, in that the payoff to actor 2 of retaining the good is higher if they have been challenged by an actor 1 that subsequently backs down, compared to if the status quo is maintained. I.e., $S_2 < D_2$. While there are plausible stories one could tell for why this might be the case, as our interest is more in observing behavior at as many nodes as possible, we set D_2 to be positive in order to induce more resistance decisions.

Similarly, 2 chooses R if $\mu_2 > \frac{-.8}{\Pr(F|E)} = -1.23$ according to (2). Given this cutpoint,

$$\Pr(R|E) = \Pr(\mu_2 > \mu_2^*) = \int\limits_{-1.23}^{\infty} \frac{e^x}{(1+e^x)^2} dx = 0.77$$

Since $B_1 > \frac{-0.5}{\Pr(R|E)} + 0.5 = -0.20$, (3) always holds regardless of the μ_1 value so Player 1 always chooses E.

$$\Pr(E) = 1 - G(\mu_1^E) I(B_1 < \frac{-D_1}{\Pr(R|E)} + D_1) = 1$$

We can now calculate the equilibrium beliefs that each player holds after observing each possible action on the equilibrium path. 2's expectation of 1's type does not change from the prior of 0 when 2 observes E because in equilibrium, 1 always challenges regardless of μ_1.

$$E_1(\mu_1|E) = 0$$

The equilibrium solution is silent on 2's expectation of 1's type if 1 chooses Not E because this action is off the equilibrium path. We discuss possible restrictions on the expectations using different equilibrium refinement concepts when we examine empirically the expectations elicited when 1 did choose Not E.

In equilibrium, 2 chooses R if $\mu_2 > -1.23$ so 1's expectation of 2's type if 2 chooses R is

$$E_2(\mu_2|R) = \frac{\int\limits_{-1.23}^{\infty} \frac{xe^x}{(1+e^x)^2} dx}{\int\limits_{-1.23}^{\infty} \frac{e^x}{(1+e^x)^2} dx} = 0.69$$

Similarly, 1's expectation of 2's type if 2 chooses Not R is

$$E_2(\mu_2|\text{Not R}) = \frac{\int\limits_{-\infty}^{-1.23} \frac{xe^x}{(1+e^x)^2}dx}{\int\limits_{-\infty}^{-1.23} \frac{e^x}{(1+e^x)^2}dx} = -2.36$$

3.1.2 High Cost: $B_1 = -0.3$

In the high cost condition, Player 1 incurs a greater loss if she Enters then decides to Not Fight. Therefore, we look for a Perfect Bayesian Equilibrium where this strategy profile is not chosen at all. That is, all Player 1s who Enter decide to Fight regardless of private value so that the probability of Fight given Enter is 1:

$$\Pr(F|E) = 1$$

From (2), we have that $\mu_2 > \frac{-0.8}{\Pr(F|E)} = -0.8$ and

$$\Pr(R|E) = \Pr(\mu_2 > \mu_2^R) = \int\limits_{-.8}^{\infty} \frac{e^x}{(1+e^x)^2}dx = 0.69$$

Since $B_1 < \frac{-0.5}{\Pr(R|E)} + 0.5 = -0.22$, 1 chooses E if $\mu_1 > \mu_1^E = -0.5 - 0.22$ according to (3), so that

$$\Pr(E) = 1 - G(\mu_1^E)I(B_1 < \frac{-D_1}{\Pr(R|E)} + D_1) = \int\limits_{-.72}^{\infty} \frac{e^x}{(1+e^x)^2}dx = 0.67$$

Now we show that it is indeed the case all Player 1s will Fight given Entry. All Player 1s who Entered have a $\mu_1 > -.72$ which is greater than $B_1 - W_1 = -.8$ so she will always choose to Fight given the incentive compatibility constraint at this terminal node in this equilibrium.

The equilibrium belief calculations in the high cost case are analogous to those in the low cost case. In equilibrium, 2 chooses R if $\mu_2 > -0.8$ so 1's expectation of 2's type if 2 chooses R is

$$E_2(\mu_2|R) = \frac{\int\limits_{-0.8}^{\infty} \frac{xe^x}{(1+e^x)^2} dx}{\int\limits_{-0.8}^{\infty} \frac{e^x}{(1+e^x)^2} dx} = 0.90$$

Similarly, 1's expectation of 2's type if 2 chooses Not R is

$$E_2(\mu_2|\text{Not R}) = \frac{\int\limits_{-\infty}^{-0.8} \frac{xe^x}{(1+e^x)^2} dx}{\int\limits_{-\infty}^{-0.8} \frac{e^x}{(1+e^x)^2} dx} = -2.00$$

In equilibrium, 1 chooses E if $\mu_1 > -0.72$ so 2's expectation of 1's type if 1 chooses E is

$$E_2(\mu_2|E) = \frac{\int\limits_{-0.72}^{\infty} \frac{xe^x}{(1+e^x)^2} dx}{\int\limits_{-0.72}^{\infty} \frac{e^x}{(1+e^x)^2} dx} = 0.94$$

Similarly, 2's expectation of 1's type if 1 chooses Not E is

$$E_2(\mu_2|\text{Not E}) = \frac{\int\limits_{-\infty}^{-0.72} \frac{xe^x}{(1+e^x)^2} dx}{\int\limits_{-\infty}^{-0.72} \frac{e^x}{(1+e^x)^2} dx} = -1.93$$

For both the Low Cost and High Cost treatments, we report the expectations one player has (E) about the other player's type after every possible action in Table 1 along with the cutpoints (μ_1^F, μ_2^*, and μ_1^E) at each decision node for both cost treatments. Note that we also elicit expectations of player 1 about player 2's type when player 1 chooses Not E and player 2 does not take an action (N/A).

	$B_1 = -0.1$	$B_1 = -0.3$
μ_1^E	$\forall \mu_1$	-0.72
μ_2^R	-1.23	-0.80
μ_1^F	-0.60	$\forall \mu_1$
E_2: E	0	0.94
E_2: Not E	N/A	-1.93
E_1: R	0.69	.9
E_1: Not R	-2.36	-2
E_1: N/A	0	0

Table 1: Equilibrium cutpoints and expectations

In the Appendix, we characterize an equilibrium solution for the entire range of payoffs to player 1 for not Fighting (B_1). Furthermore, we show that the equilibria we found for the low cost treatment and the high cost treatment are indeed unique.

3.2 Hypotheses

We derive a number of testable comparative statics for the low vs. high cost treatment from the equilibrium cutpoints and beliefs/expectations. Our experimental design does not allow us to directly observe players using cutpoints. Thus we craft our empirical analysis to tease out differences across treatments in several different ways. We consider three approaches. First, we can ask whether the average private value of those making a particular decision differs across treatments. Second, we can ask questions about the proportion of all subjects at a particular node making some choice. Finally, we estimate cutpoints by using an optimal classification strategy used by Carillo and Palfrey. This approach selects some value that correctly predicts the most decisions. If there are multiple such values we take the average.[4] We test differences in these quantities by

[4]Earlier versions of this paper implemented a fourth strategy, which was to use parametric techniques to estimate probit models using only the subject's private value: $\Pr(Enter) = \Phi(value)$, where Φ denotes the probit link function. Using the parameter estimates from estimating this model on the seperate treatments, we can plot the predicted probability of a decision given a subject's private value. The intersection of the predicted probability curve with the horizontal line equal to .5 represents the estimated cutpoint. Calculating the 95% confidence intervals around these predicted probability curves, in order to account for uncertainty in our parameter estimates, allows us to estimate whether differences in the estimate of a cutpoint is statistically different. We also used Bayesian methods to compare the slope coefficient on subject private values. This is a measure of the degree to which the probit model

using difference in means, differences in proportions, and nonparametric Mann-Whitney rank sum statistics.

First note from Table 1 that in equilibrium, we should only observe Entry in the low cost case. Furthermore, we should only observe Fighting given Entry and Resistance in high cost case.

Hypothesis 1: Player 1 should always choose E when $B_1 = -0.1$ and always choose F if that decision node is reached when $B_1 = -0.3$.

Hypothesis 1 is a knife edge result. We also consider two related hypotheses about Player 1s who do choose E or F given the equilibrium cutpoints, given that it may be that not all in the $B_1 = -0.1$ treatment enter, or all at the Fight node choose F in the $B_1 = -0.3$ treatment.

Hypothesis 2: The mean private value and estimated cutpoint of those who chose E when $B_1 = -0.1$ should be lower than the mean private value of those who chose E when $B_1 = -0.3$. The proportion of subjects choosing E should be higher in the $B_1 = -0.1$ treatment.

Hypothesis 3: The mean private value and estimated cutpoint of those who chose F when $B_1 = -0.1$ should be higher than the mean private value of those who chose F when $B_1 = -0.3$. The proportion of subjects choosing F should be higher in the $B_1 = -0.3$ treatment.

We state a similar implication for the private value of Player 2s who choose to resist.

Hypothesis 4: The mean private value and estimated cutpoint of those who chose R when $B_1 = -0.1$ should be lower than the mean private value of those who chose R when $B_1 = -0.3$. A higher proportion of player 2's should choose to resist in the low cost treatment.

discriminates between those that enter and those that do not enter. We find interesting differences across our treatments. Results from these approaches are largely consistent with our other approaches, and we omit them here.

The equilibrium beliefs we derived gives us a similar set of hypotheses about the expectation of the other player's private value that we elicit after the first two decision nodes.

Hypothesis 5: The expectation of Player 1's private value that we elicit from Player 2 after Player 1 has chosen E should be higher when $B_1 = -0.3$.

Hypothesis 6: The expectation of Player 2's private value that we elicit from Player 1 after Player 2 has chosen either R or not R should be higher when $B_1 = -0.3$.

Hypothesis 7: The expectation of Player 2's private value that we elicit from Player 1 after Player 1 has chosen Not E should remain at the prior of 0 because Player 1 has observed no action from Player 2 on which to update her prior.

4 Experimental Design and Procedures

We conducted 14 sessions with a total of 112 subjects in the Spring and Fall of 2007. The subjects were registered students at Princeton university who were recruited by email solicitation. Sessions were conducted in a computer lab and all interaction was computerized. No subject participated in more than one session. The primary treatment variables were single versus repeated play, and cost to player 1 of choosing to not Fight.

After entering the laboratory subjects were seated at workstations with dividers and were given several instructional worksheets to read over. During the instructional period a set of instructions were presented verbally with the assistance of a Powerpoint presentation. Subjects were informed of the structure of the game (strategies/payoffs), re-matching procedure, and a description of the properties of the distribution that private values were drawn from. During this period subjects were guided through an example series of interactions using their computers, and then were asked to answer a series of questions about the experiment. They could not proceed unless all answers were correct. A complete set of instructions are available from the authors.

The game form is a sequential game of two-sided incomplete information. In the experiment, two subjects would be paired together and randomly assigned a position (position 1 or position 2) and a private value drawn from a standard logistic distribution. After receiving these assignments subjects would then play the game once or three times before being paired with someone else, depending on the repetition treatment. Within a pair all position assignments and private values were fixed, across pairs these were randomly reassigned. In the game the subject in position 1 would choose between "E" and "not E". If not E was chosen both subjects earned 0 points. If E was chosen then the position 2 subject was asked to report what they thought the position 1 subject's private value was by moving a slider on their computer interface. This allows them to report their belief about the type of player they face. To incentivize this belief elicitation we utilized a quadratic scoring rule in order to ensure that subjects had no incentive to report anything other than what they truly thought the other subjects private value was (Savage 1971). The quadratic scoring rule is of the form

$$S = \alpha - \beta(E - X)^2$$

where E is the elicited expectation of the other player's private value and X is the other player's actual private value. In all experimental sessions, we set $\alpha = 0.5$ and $\beta = 0.1$.

After reporting their belief the position 2 subject would choose between "R" and "not R". If not R was chosen subject 1 earned .5 points and subject 2 earned $-.3$ points. If R was chosen then subject 1 first reported that they thought player 2's private value was. Next they chose between F and not F. If not F was chosen then player 1 earned either $-.3$ or $-.1$ (depending on the payoff treatment) and player 2 earned .5 points. If F was chosen, then both subjects earned the sum of .5 plus their private value. Thus in this case, "type" is interpreted as a player's valuation of the outcome (E, R, F).

If not E was chosen, player 2 was asked for their belief about player 1's type. Then player 1 was asked about player 2's type. Similarly, if not R was chosen by player 2 then player 1 would be asked for their belief about player 2, but would not subsequently be able to make a decision between "F" and "not F". Thus, we observe beliefs along all possible paths of play.

In the repeat play design subjects would then repeat the same set of decisions two more times, for a total of three times per pair. Afterwards they were repaired and receive randomly assigned private values and positions. When all subjects had been paired with all other subjects once, we repeated the experiment, randomly matching subjects and repeating the process described above.

At the bottom of the screen all subjects were provided with a tabbed history screen with their decisions, the decisions of the person they were paired with, their point earnings from the reached terminal nodes, and their reported belief about the other subject's private value. They were not told their earnings from beliefs, as this would then allow them to infer the private value of the other subject, until after they had completed all interactions with that subject.

Within a repetition of the experiment (each subject playing each other subject once) subjects earned points that could be converted to dollars from 1) all rounds in the single play case or 2) a randomly chosen round within each pairing for the repeat play case. Thus because there were 8 total subjects, there were seven rounds (randomly) chosen to pay them with. We selected the random payment round selection rule for the repeat play case so as to minimize repeat play effects, and thus making finding differences between single and repeat play harder. After finishing a pairing subjects were told which of the three rounds were chosen for payment. We repeat the experiment 9 times in the single play case, and 3 times in the repeat play case, so as to generate the same number of total observations. We randomly selected one repetition of the experiment to pay subjects with, and reminded subjects of this procedure so as to maintain their focus. This payment procedure was selected to minimize repeat play effects across repetitions of the experiment that we do not explicitly model. After completing the experiment subjects completed a short survey about themselves and the experiment. All subjects were paid one at a time in private.

4.1 Descriptive Statistics

As is evident in Table 2, the average private values and standard deviation of the private values are close to logit distribution we used as our random number generator. The gender ratio across treatments was close to 50/50 in all but our low cost single play treatment, where a majority of subjects were female. Because we find relatively few gender differences we are not concerned about this imbalance. One striking difference across the single play and repeat play experiments is the average payoff for the belief elicitation component. In both the high and lo cost treatments, subjects earned significantly more from the belief payoffs on average in the repeat play design (t-statistics from difference in means tests with unequal variances were -5.2 for the high cost and -3.7 for the low cost treatments). Restricting these tests to the first round of play, thus all single play rounds and only the first round of the three iterations played in the repeat play treatment, also reveal significant differences (though the t-ratios are lower in absolute magnitude). These differences in belief payoffs are a rough indicator that subjects update their beliefs closer to the real value of their opponents' private value in the repeat play design. This may be indicative of a number of different things. We intend to explore this finding in more detail in future research, and instead turn to tests of our hypotheses.

	Private Value		Belief		Game Payoff		Belief Payoff		Gender	N
Treatment	Mean	Std	Mean	Std	Mean	Std	Mean	Std		
Single, Hi Cost	0.05	1.82	0.16	1.75	0.68	1.14	0.08	1.44	0.29	1512
Single, Lo Cost	-0.07	1.86	0.13	1.28	0.66	1.27	0.13	0.84	0.5	1512
Repeat, Hi Cost	-0.08	1.77	0.09	1	0.63	1.18	0.28	0.4	0.56	2016
Repeat, Lo Cost	0	1.78	0.25	1.12	0.66	1.1	0.22	0.58	0.44	2016

Table 2: Summary Statistics for all treatments

5 Results

We structure our analysis by focusing on the elicited beliefs in our game. Other parts of our larger project focus on strategies with greater empirical attention. However, because beliefs and strategies jointly constitute a PBE, we cannot simply ignore the strategy choices preceding the elicitation of beliefs. We are interested in 1) whether players update, 2) whether updating occurs according to the predictions and comparative statics of the PBE 3) whether updating and differences in updating across treatments are rational given the actual private values of subjects choosing specific strategies in the experiment, 4) whether beliefs other than those elicited would have yielded higher payoffs, 5) how updating works within a dyad in a repeat play treatment and 6) whether strategies are best responses to beliefs.

We begin first with the simple question of whether players update. Do players report something other than the expectation of the prior? The answer is an unequivocal yes. Upon observing a decision to enter or not enter, player 2 updates their belief to a value other than the mean of the prior, 0. Upon observing resistance or not resistance, player 1's update their beliefs. Even in the case where player 1 does not enter, these subjects regularly report a value greater than 0 for player 2's type, though this can hardly be called updating in that they did not observe a strategy choice by player 2. We return to this interesting phenomenon below.

B_1= −0.1	All Repetitions			Repetitions 1-4			Repetitions 5-9			Equilibrium
Action	N	Mean	Std	N	Mean	Std	N	Mean	Std	
1-Not E	163	-1.24	1.14	70	-1.26	1.17	93	-1.22	1.13	N/A
2-NotApplicNoEntry	163	**0.12**	0.93	70	**0.05**	0.58	93	**0.17**	1.13	0
1-Enter, 2 Backs Down	142	0.53	0.61	58	0.49	0.59	84	0.56	0.63	0
2-No Resistance	142	-1.23	0.96	58	-1.17	0.92	84	-1.27	1	-2.36
2-Resistance	451	0.58	1.15	208	0.45	0.97	243	0.69	1.27	0.69
1-Enter, Not Fight	73	0.15	1.26	31	0.32	1.1	42	0.03	1.37	0
1-Enter and Fight	378	0.56	1.12	177	0.39	0.89	201	0.72	1.26	0
Total N			1512			672			840	

Table 3a: Mean Elicited Expectation of Private Value by Actions (Single Play Low Cost)

B_1= −0.3	All Repetitions			Repetitions 1-4			Repetitions 5-9			Equilibrium
Action	N	Mean	Std	N	Mean	Sd	N	Mean	Sd	
1-Not E	205	-1.43	1.64	99	-1.35	1.55	106	-1.5	1.73	-1.98
2-NotApplicNoEntry	205	**0.34**	1.33	99	**0.48**	1.77	106	**0.22**	0.7	0
1-Enter, 2 Backs Down	177	0.75	1.32	68	0.93	1.32	109	0.63	1.31	0.91
2-No Resistance	177	-1.35	1.43	68	-1.37	0.96	109	-1.34	1.66	-1.83
2-Resistance	374	0.66	1.61	169	0.83	1.47	205	0.53	1.71	1.01
1-Enter, Not Fight	28	0.55	0.64	12	0.47	0.84	16	0.61	0.45	0.91
1-Enter and Fight	346	0.87	1.55	157	1.05	1.66	189	0.72	1.44	0.91
Total N			1512			672			840	

Table 3b: Mean Elicited Expectation of Private Value by Actions (Single Play High Cost)

Next we ask whether players update in a way that is consistent with the PBE predictions. We begin first with player 2's expectation about player 1's private upon observing entry. We focus on the single play treatment. In the low cost case all player 1's should enter, and thus player 2's beliefs should not be updated. We observe that player 2's in the low cost treatment regularly update their beliefs upon observing entry to a positive value. Thus we do not find support for this prediction of the PBE. Below we ask whether updating of this nature is rational given actual behavior in the laboratory because not all player 1's choose to enter. In the high cost case only player 1's with a private value above -0.72 should enter, and thus player 2's expectation should be 0.94. Conversely decisions to not enter should generate an expectation of -1.93. We observe updating in the correct direction, but we find that on average subjects reported beliefs that were smaller in magnitude from both of these theoretical predictions. Such conservatism in beliefs (Edwards 1968) has been documented for probabilistic stated beliefs in experiments (Huck and Weizsäcker 2002).

To look at this more closely, and to see how close a player's belief is to their opponent's actual private value (rational expectations), we plot belief/private value pairs for several sets of moves in our single play design in Figure 2. Our theoretical model predicts that, in the high cost case, Player 2 should update their belief if they see Entry or no Entry by Player 1. We draw these equilibrium prediction lines in our graphs. In equilibrium, all Player 1's should enter in the low cost case thus there is no predicted updating from the prior if Entry is observed and no predictions about beliefs after observing the off-equilibrium path action of no Entering. As discussed above and shown in the bottom two plots of Figure 2 player 2 indeed updates. In the top two plots of Figure 2 it appears that beliefs cluster slightly closer to 0 than what the equilibrium prediction for beliefs would be. We can statistically test whether the difference between observed beliefs and equilibrium beliefs is significantly different from zero. In the case of Entry, this difference is negative and significantly different from zero, suggesting conservative belief updating ($t=-1.56$, $\Pr(T < t) = 0.056$). If we exclude outlier cases that gave beliefs of 10 or beliefs below -8, the t-statistic increases to -3.35 . We observe similar conservative updating in the no Entry case. Here, the difference between the elicited belief and the equilibrium belief is positive and significantly different from zero ($t=4.8$).

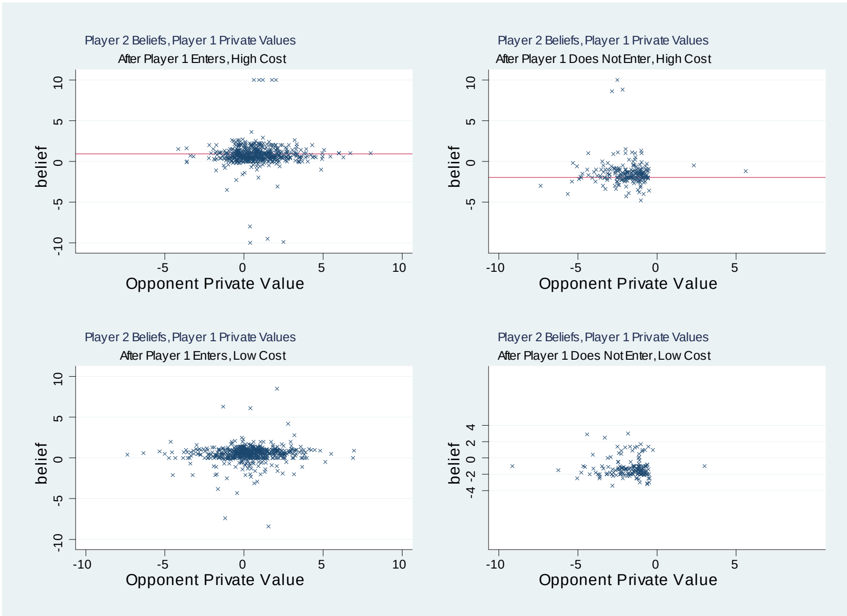

Figure 2: Elicited and Equilibrium Beliefs after Enter or not Enter is Observed

Next we look at updating by player 1 about the type of player 2 they face (Figure 3). While this is not directly relevant to the decision of player 1 were they to reach the fight node (i.e., after they have entered and player 2 has resisted), our equilibrium model still makes predictions about the beliefs they should hold given the decisions made by player 2. Unlike in the Entry case, we have equilibrium predictions about beliefs in both payoff treatments in the single play case. We begin with the high cost case. Here, when player 1 observed Resistance, the difference between the elicited belief and the equilibrium prediction was significantly negative (t=-4.2). When player 1 observed no Resistance, beliefs were also updated in a conservative manner (t=4.46). Similar results obtain if we consider the low cost case, generating corresponding t-statistics of t=-3.38 & t=15.9. We find very strong support for the idea that elicited beliefs are significantly closer to the prior of 0 than the equilibrium predictions. Updating in our experiment, as measured by a strictly proper scoring rule, appears to be more conservative than predicted in the theory.

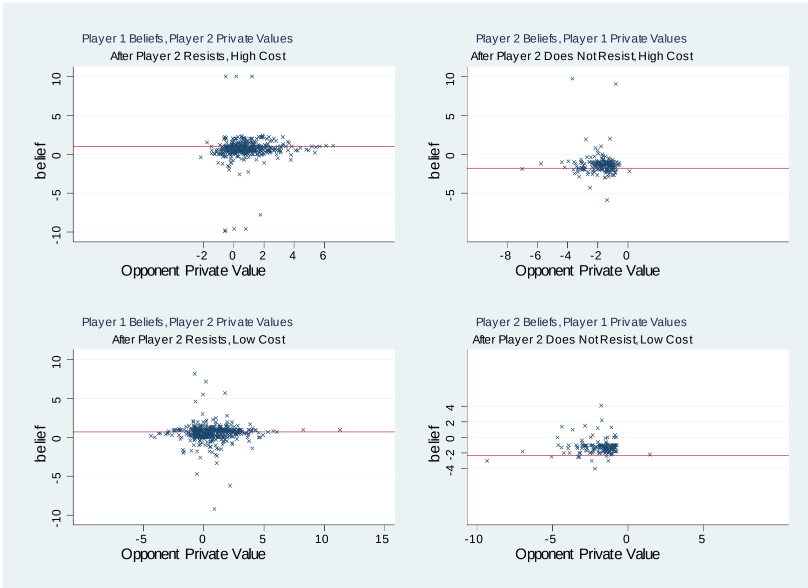

Figure 3: Elicited and Equilibrium Beliefs after Resist or not Resist is Observed

A final prediction from the theoretical model involves when player 1 subjects choose not to enter, they receive no information about player 2 because they do not observe their strategy choice. The prediction, stated as Hypothesis 7, is that Player 1 does not update her prior of 0 for Player 2's private value. But do we observe this in the experiment? We actually find that the expectations are signficantly higher than 0 in all of our experimental treatments. To help see this we plot in Figure 2 player 1's private value against their *belief* about player 2's private value in figure 4. We do this for each of our treatments, and only include the first period of the repeat play treatment. We list the t-statistic for a test whether the mean of the beliefs is significantly different from 0. These t-statistics are calculated without a handful of outlier observations not included in the graphs, and only increase if we include these subjects. The difference from 0 is strongest in the high cost single play case.

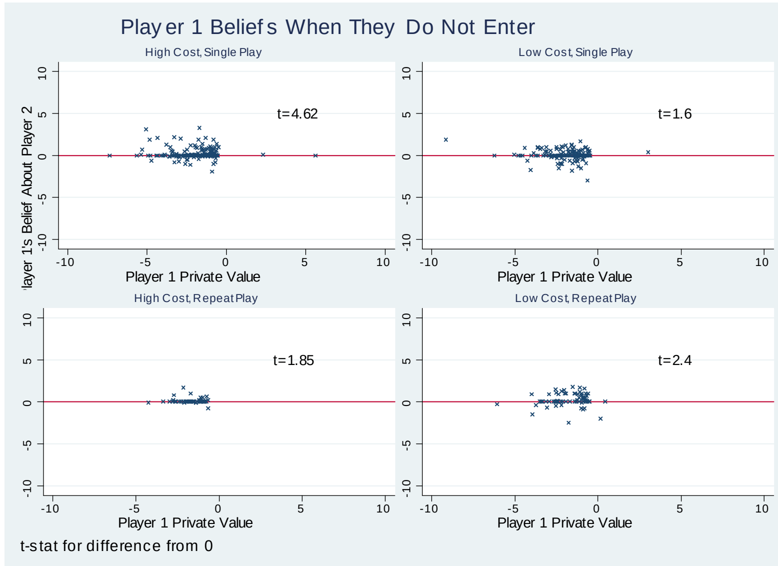

Figure 4: Motivate Beliefs and No-Entry

In each treatment condition it appears that subjects are not being Bayesians given the way privately values are generated. There are several possible explanations for this. First, subjects in our experiment may be forgetting/ignoring the iid assumption on how private values are generated. Our instructional period emphasized this fact, and subjects had to indicate on a quiz that they understood this part of the experiment. Thus to the extent possible this influence was minimized. A second possibility is that subjects in this position may hold "motivated beliefs" (Benabou and Tirole 2006) so as to rationalize their decision to not Enter. These subjects tended to report expectations of Player 2 having a positive private value, consistent with the belief that Player 2 would likely resist if they were to choose Enter. Thus, the decision to not enter could be rationalized in terms of these beliefs. There is a related strand in the psychological literature on choice-supportive bias which finds that people often selectively remember past events so as to lend support to the actions they chose (Mather et al. 2000). This rationalizing belief that deviates from the prior can be seen as a case of resolving "cognitive dissonance" where beliefs are molded so as to be consistent with the actions chosen (Akerlof and Dickens 1982). A final interpretation is that of regret avoidance.

Here, subjects not entering believed that they would be likely to face a strong opponent and thus would not want to enter and regret the outcome. This mechanism has been used, for example, to explain why people are reluctant to exchange lottery tickets (Risen and Gilovich 2007).

Space does not permit us to investigate these different possible mechanisms, and we expect that new experimental designs would be necessary to study the operation of this bias more closely in strategic settings. One limitation of our current design is that many of the subjects not entering in the single play high cost case should not have entered given the private value cutpoint equilibrium we consider. This makes teasing out the effect of beliefs and private values on strategy choices more difficult. This is not, on the other hand, the case for our low cost treatment, where all subjects should enter. Thus, if we assume behavior is governed by the equilibrium we consider, the low cost treatment provides a more revealing picture of this bias. Interestingly, we found less evidence of this bias when we looked at player 2 beliefs depending on whether they chose to resist or not resist. Only in the repeat play high cost treatment was there a significant difference in beliefs between player 2's that resisted and player 2's that did not resist (t=-1.71). In this case, players not resisting reported higher beliefs on average. Because both sets of subjects observed the same action (a single entry), in theory there should be no difference in their beliefs.

Next we address questions about how beliefs relate to actual behavior in the experiment. This section lets us also consider the comparative static predictions made about how beliefs should be different across treatments, insofar as strategy selection differs across treatments in the laboratory. Thus prior to discussing our results about beliefs we discuss briefly whether our comparative predictions about strategies are borne out in the laboratory.

5.1 Empirical Strategies and Rational Expectations

Actual beliefs, of course, should not necessarily be held against the exact predictions of a theoretical model that prescribes strategies different from those that players actually use. Beliefs may form given the particular set of strategies present in the laboratory. We test whether rational expectations holds, that is whether the actual mean private values broken down by moves are equal to the corresponding elicited expectations of those private values. We find that many of the expectations are qualitatively very similar to the actual private values with some interesting exceptions that we discuss in more details below. We proceed by first characterizing the strategy choices we observed in the laboratory. We examine strategy choices empirically by considering average private values of subjects making certain choices, the proportion of subjects choosing each strategy, and an estimated "cut-point" that gives for each subject the private value that optimally predicts their decision at each node. We then leverage the differences in behavior induced by changes in the cost of player 1 not fighting to explore the intersection of beliefs and strategies more thoroughly.

Tables 4a (low cost treatment) and 4b (high cost treatment) break down the mean private values by all possible moves of player 1 (1-Not E, 1-Enter, 2 Backs Down, 1-Enter, Not Fight, 1-Enter and Fight) and player 2 (2-NotApplicNoEntry, 2-No Resistance, 2-Resistance). For example, if player 1 Enters and 2 backs down, then the outcome for player 1 is coded as "1-Enter, 2 Backs Down". If 1 enters, 2 resists, and 1 backs down, then the outcome for player 1 is coded as "1-Enter, Not Fight" and player 2 is coded as "2-Resistance". All the outcomes for player 1 except "1-Not E" imply that player 1 entered. In the not Enter case, 2's decision is listed as "2-NotApllicNoEntry". We also compare the earlier and later repetitions in the tables for possible learning effects.

$B_1 = -0.1$	All Repetitions			Repetitions 1-4			Repetitions 5-9		
Actions	N	Mean	Std	N	Mean	Std	N	Mean	Std
1-Not E	163	-1.82	1.28	70	-1.84	1.35	93	-1.81	1.24
2-NotApplicNoEntry	163	-0.09	1.7	70	0.03	1.62	93	-0.17	1.76
1-Enter, 2 Backs Down	142	0.18	1.79	58	0.09	1.88	84	0.25	1.73
2-No Resistance	142	-1.99	1.25	58	-1.87	1.2	84	-2.08	1.29
2-Resistance	451	0.66	1.65	208	0.62	1.77	243	0.7	1.54
1-Enter, Not Fight	73	-1.7	1.05	31	-1.54	0.94	42	-1.81	1.12
1-Enter and Fight	378	0.77	1.47	177	0.69	1.46	201	0.83	1.48
Total N			1512			672			840

Table 4a: Mean Private Value by Actions (Single Play Low Cost)

$B_1 = -0.3$	All Repetitions			Repetitions 1-4			Repetitions 5-9		
Actions	N	Mean	Std	N	Mean	Std	N	Mean	Std
1-Not E	205	-1.94	1.3	99	-2.04	1.57	106	-1.86	0.99
2-NotApplicNoEntry	205	0.14	1.6	99	0.09	1.63	106	0.19	1.58
1-Enter, 2 Backs Down	177	0.88	1.71	68	0.61	1.78	109	1.05	1.65
2-No Resistance	177	-1.85	0.94	68	-1.8	0.79	109	-1.88	1.03
2-Resistance	374	0.87	1.34	169	0.7	1.35	205	1.01	1.33
1-Enter, Not Fight	28	-1.75	0.9	12	-1.88	1.02	16	-1.64	0.82
1-Enter and Fight	346	0.96	1.26	157	1.02	1.18	189	0.91	1.32
Total N			1512			672			840

Table 4b: Mean Private Value by Actions (Single Play High Cost)

Hypothesis 1 predicts that there should be more Entry in the low cost treatment than in the high cost treatment. Indeed, table 4b shows that the high cost treatment saw more decisions to not enter. Player 1 subjects in the low cost treatment entered 78% of the time, while the entrance rate was 73% in the high cost treatment. This difference in proportions is statistically significant with a corresponding z-score of -2.15 ($\Pr(Z < z) = 0.015$). While there was more entry in the low cost treatment, the prediction of hypothesis 1 that all subjects would enter is not observed. Nevertheless, we observe the proper ordering of entry rates given the theory.

Hypothesis 2 predicts that the average private value of player 1 subjects that decide to enter should be lower in the low cost treatment. This is because the cutpoint for entry in the high cost treatment is -0.72, meaning that subjects with private values below this value should not enter. We also observe statistically significant support for hypothesis 2. The average private value of subjects that entered in the high cost treatment was 0.8, whereas the average private value in the low cost treatment was 0.32. This difference is statistically significant with a t-value of 4.6, and a Mann-Whitney (MW) test statistic of z=3.89.

Table 5 displays our estimates of the average optimal cutpoints using the procedure described above. The first column lists the treatment type. The second column takes the optimal classification cutpoint for each subject and then takes the average of all subjects within the particular treatment. Values in the parentheses are the 95% confidence intervals corresponding to these estimates (we ignore the error distribution created when averaging an individual subject's optimal cutpoint when there were multiple optimal classification values). The third column calculates the median of this distribution. The fourth column considers an alternative strategy, where we ignore subject by subject differences, and calculate the optimal classification cutpoint for our subject pool as whole within each treatment.

Focusing on the single play case we find that the optimal classification cut point for entry is lower in the low cost case, though there is a small amount of overlap in the confidence intervals. An alternative non-parametric test, two sample Mann-Whitney, generates a test statistics of z=1.4, indicating that the probability that the high cost cutpoint is greater than the low cost cutpoint is .62. The median and group cut points are similarly ordered, but these values do not have associated confidence intervals. Interestingly, the group cut point for the high cost case is exactly that predicted in the theory.

TreatmentType	Mean	Median	GroupCutPoint	PBE
Enter Repeat High	-2.32(-2.84,-1.8)	-1.91	-4.14	
Enter Repeat Low	-1.86(-2.43,-1.3)	-1.37	-1.8	
Enter Single High	-1.11(-1.49,-.72)	-0.78	-0.8	$-.72$
Enter Single Low	-1.99(-2.65,-1.33)	-0.97	-1.1	$\forall\mu_1$
Fight Repeat High	-1.87(-2.19,-1.54)	-1.68	-0.85	
Fight Repeat Low	-1.41(-1.76,-1.05)	-0.95	-0.6	
Fight Single High	-1.71(-2.05,-1.37)	-1.2	-0.85	-0.60
Fight Single Low	-1.95(-2.5,-1.39)	-2.65	-0.6	$\forall\mu_1$
Resist Repeat High	-1.4(-1.69,-1.1)	-1.15	-1.2	
Resist Repeat Low	-1.84(-2.26,-1.41)	-1.45	-1	
Resist Single High	-.85(-.97,-.74)	-0.85	-0.8	-0.80
Resist Single Low	-1.49(-2.02,-.96)	-0.9	-0.8	-1.23

Table 5: Optimal Classification

Given these results it seems reasonable to expect that beliefs of player 2 should differ across the two treatments. This is predicted by our theoretical model, which assumes that player 1 subjects following specific equilibrium strategies. While these subjects do not exactly follow the cutpoint strategies involved, there are clear differences across treatments. But do beliefs pickup on this difference, and how do the strategies we observe explain any of the deviations discussed above?

Tables 3a and 3b report the mean expectations elicited from the subjects after each particular belief sequence. For example, the beliefs elicited in the "1-Enter, 2-Backs Down", "1-Enter, Not Fights", and "1-Enter, Fight" rows are all from Player 2 after Player 1 has Entered but we break them down into these categories based on possible subsequent moves, partly for ease of comparison to the mean private value tables. We find support for Hypothesis 5 which states that the expectation of Player 1's private value after she has chosen Enter should be higher in the high cost case than in the low cost case. Player 2's belief does get updated to a significantly higher level after observing Entry in the high cost case with a t-value of 3.99 (Mann-Whitney: z=5.29). We also clearly see why player 2 beliefs after observing entry in the low cost case change (an off the equilibrium track case because in theory all player 1's should enter). 22% of entrants subjects screen out with low private values in the low cost treatment.

We now turn our attention to the second player's decision to Resist or not Resist and player 1's subsequent decision to Fight or not Fight.[5] Again, we first ask whether the cost treatments had an effect on strategy choice and then see whether this translates to differences in beliefs. Tables 4a and 4b above report the mean private value of subjects who decided to resist or not resist. According to hypothesis 4, player 2 subjects who decide to resist in the low cost treatment should have a lower mean private value than those who do so in the high cost treatment. We do observe that the proportion of subjects resisting in the single play low cost case is significantly higher, with a z-score of -2.65 in keeping with the lower equilibrium cutpoint. The mean private value of those who do resist is slightly lower in the low cost treatment (t=1.53, $Pr(T > t) = 0.06$; (MW: z=1.7)).

Table 5 reports the optimal classification cutpoints for resist decisions. Our estimates here also show that the cutpoint is lower in the high cost case. Again, the 95% confidence intervals barely overlap. Mann-Whitney tests put the probability that high cost case has a higher cutpoint at .62 (z=1.4) (ranksum CutPointResist_OpClass if ExpType==1 , by(IncentiveType) porder). The median is also properly ordered, though interestingly the group estimates are the same. Thus we find some support for the

[5] There is a potential problem with analyzing the resist and fight decisions. The above mode of analysis may produce results that suffer from post-treatment bias. By the time the fight node has been reached we might not have comparable samples between the 'control' condition (low cost) and 'treatment' condition (high cost). Because the treatment presumably influences entry decisions (as documented above), subpopulations of subjects with low private values decide to not enter. We are thus unable to know what they would have done had they reached the fight node. Thus the sample that has made it to the fight node is a specially selected sample, and that sampling process depends on the very treatment effect we wish to observe. This does not appear to be a problem for our investigation for two reasons. First, the model we are testing abstracts away from any unit (individual) level heterogeneity and assumes random assignment into our treatments conditional on any actual heterogeneity. While reasonable, we feel this is still unsatisfying. Typical candidates like the role of gender never appeared to have an effect. We did not measure subject risk preferences, though the fact that gender had no effect suggests this might not be a problem (Eckel and Grossman 2008). Second, the distributions of privates values for subjects that reach the fight and resist nodes appear to have sizeable overlaps in their support, and thus empirically we are able to make these comparisons.

theory's prediction that the cutpoint for resistance will be lower in the low cost case, though the statistical evidence is weaker.

Do these differences in cutpoints for resistance translate to difference in beliefs, as predicted by the model? Similarly we look at the expectation of Player 2's private value after he has chosen both Resist and not Resist in the two cost treatments. We find no statistically significant difference (Resist: t=.97, Not Resist t=-.37) in either case. Thus we do not find support for Hypothesis 6 which posits that the expectations should be higher in the high cost case after Resistance and no Resistance.

Finally we turn to Player 1's choice to Fight or not Fight given that they entered and faced resistance. Because this decision is made at the final node it does not generate additional updating in the single play case we consider here. We nevertheless consider whether the hypotheses derived from the model are supported in order to completely cover the comparative statics on strategies discussed in the theoretical section. According to Hypothesis 1, Player 1 should always fight in the high cost case. Although we do not observe this in the data, we do observe a very low rate of not Fighting. Furthermore, the proportion of subjects who chose Fight in the high cost case is significantly higher (z=3.99) than in the low cost case in keeping with the hypothesis 1. We can also compare the average private value of subjects who chose to Fight or not Fight across the low and high cost treatments. The average private value of subjects who fought in the high cost case is higher than in the low cost case, though this difference is only significant at the 0.1 level using differences in means (t=1.3, $Pr(T > t) = 0.09$) and insignificant using Mann-Whitney (z=.36). This result only weakly contradicts Hypothesis 3 which states that the average private value should be lower in the high cost case where every player is supposed to choose Fight in equilibrium. However, the average private value of subjects who decided to not Fight was not significantly different across the two cost treatments which suggests that nothing conclusive can be said about the relative mean private values of subjects choosing fight across cost treatments. The optimal classification cutpoints also show little difference between the two cost treatments for the decision to fight. The confidence intervals overlap considerably, thus this procedure suggests little difference in cutpoints at this node. Mann-Whitney tests suggest a similar conclusion. Interestingly, the estimated group cut point for the low cost case is exactly that predicted by the theory.

5.2 Effect of Repeat Play on Strategy Choice and Belief Updating

We now turn to our repeat play treatments. While we are not able to derive formal equilibrium solutions[6], repetition allows us to investigate how beliefs change within a particular dyad over multiple interactions. Given the prevalence of repeated interactions in the applications we discussed in Section 2.2, we want explore whether such repetition would improve the accuracy of a player about her opponent's type. As with our analysis in the previous section we examine first whether repetition changed behavior, but do so within a payoff treatment. This then sets us up to ask whether beliefs reflect any differences across the single and repeat play setting.

Are strategy choices in the first period influenced by whether the game will be played once or three times? Within the repeat play design, are strategy choices influenced by the cost treatment? Tables 6a and 6b are the repeat play analogs of Tables 4a and 4b. These tables only report decisions in the first period of play, as behavior in subsequent periods can be influenced by the exact course of play in the first period. Likewise, we also report the analogs of tables 3a and 3b for beliefs in our repeat play treatment in Tables 7a and 7b.

[6]Existing techinques for analyzing repeat games of incomplete information often exploit the existence of one-sided incomplete information to fix the strategy profile of one player. Our case obviously complicates formal analysis significantly.

$B_1 = -0.1$	All Repetitions			Repetitions 1			Repetitions 2&3		
Action	N	Mean	Std	N	Mean	Std	N	Mean	Std
1-Not E	69	-1.81	1.13	16	-2.12	1.19	53	-1.71	1.11
2-NotApplicNoEntry	69	-0.08	1.83	16	0.82	2.19	53	-0.35	1.64
1-Enter, 2 Backs Down	53	0.75	1.76	12	1.51	1.35	41	0.53	1.82
2-No Resistance	53	-2.11	1.18	12	-2.09	1.26	41	-2.12	1.17
2-Resistance	214	0.45	1.45	84	0.42	1.31	130	0.47	1.55
1-Enter, Not Fight	39	-1.53	1.16	20	-1.19	0.92	19	-1.89	1.28
1-Enter and Fight	175	0.94	1.3	64	0.72	1.28	111	1.07	1.3
Total N			672			224			448

Table 6a: Mean Private Value by Actions (Repeated Play First Period Low Cost)

$B_1 = -0.3$	All Repetitions			Repetitions 1			Repetitions 2&3		
Action	N	Mean	Std	N	Mean	Std	N	Mean	Std
1-Not E	47	-1.7	0.78	15	-1.58	0.58	32	-1.76	0.86
2-NotApplicNoEntry	47	-0.13	1.45	15	-0.31	1.56	32	-0.04	1.42
1-Enter, 2 Backs Down	79	0.09	1.63	25	0.06	1.64	54	0.1	1.64
2-No Resistance	79	-2.18	1.14	25	-2.01	1.04	54	-2.26	1.19
2-Resistance	210	0.61	1.63	72	0.4	1.72	138	0.71	1.57
1-Enter, Not Fight	30	-2.14	0.91	8	-2.27	0.83	22	-2.1	0.95
1-Enter and Fight	180	0.73	1.22	64	0.74	1.18	116	0.72	1.25
Total N			672			224			448

Table 6b: Mean Private Value by Actions (Repeated Play First Period High Cost)

$B_1 = -0.1$	All Repetitions			Repetitions 1			Repetitions 2&3		
Action	N	Mean	Std	N	Mean	Std	N	Mean	Std
1-Not E	69	-1.00	1.07	16	-0.54	1.36	53	-1.14	1.11
2-NotApplicNoEntry	69	-0.23	0.79	16	0.13	1.04	53	0.25	1.64
1-Enter, 2 Backs Down	53	0.72	0.66	12	0.73	0.57	41	0.71	1.82
2-No Resistance	53	-0.98	0.91	12	-0.36	1.08	41	-1.16	1.17
2-Resistance	214	0.71	0.70	84	0.62	0.81	130	0.78	1.55
1-Enter, Not Fight	39	0.46	0.75	20	0.27	0.62	19	0.66	1.28
1-Enter and Fight	175	0.56	0.70	64	0.63	0.75	111	0.52	1.3
Total N			672			224			448

Table 7a: Mean Belief by Actions (Repeated Play First Period Low Cost)

$B_1 = -0.3$	All Repetitions			Repetitions 1			Repetitions 2&3		
Action	N	Mean	Std	N	Mean	Std	N	Mean	Std
1-Not E	47	-1.29	0.78	15	-1.01	0.76	32	-1.42	0.76
2-NotApplicNoEntry	47	-0.10	0.35	15	0.07	0.16	32	0.11	0.42
1-Enter, 2 Backs Down	79	0.43	0.50	25	0.47	0.63	54	0.41	0.43
2-No Resistance	79	-1.31	0.62	25	-1.10	0.77	54	-1.41	0.52
2-Resistance	210	0.50	0.51	72	0.43	0.54	138	0.54	0.49
1-Enter, Not Fight	30	0.41	0.54	8	0.14	0.35	22	0.51	0.57
1-Enter and Fight	180	0.42	0.47	64	0.45	0.56	116	0.41	0.42
Total N			672			224			448

Table 7b: Mean Belief by Actions (Repeated Play First Period High Cost)

We begin first with comparing behavior within a cost treatment and across the single and repeat play designs. We include only the first round of the repeat play treatment. Table 5 shows that repeat play influences entry decisions in different ways depending on the cost treatment. For the high cost case, the estimated cutpoint for entry is significantly lower in the repeat play case (t=3.6, MW: z=3.07). For the low cost case, the estimated cutpoint for entry is generally higher in the repeat play case but not significantly so. Under the high cost repeat play treatment, player 1's choose to enter 86% of the time versus 72% in the single play case, a significant difference (z=-4.8). Under the low cost repeat play treatment player 1s chose to enter 79% of the time versus 78% in the single play case, an insignificant difference. These findings are perhaps not surprising given the incentives for reputation building under incomplete information so that Player 1 can appear to be someone with a high private value by entering in the high cost case. Since the equilibrium behavior is already always enter in the low cost case even in single play, the equilibrium entry percentage predicted for the repeated play case could not possibly be higher. Thus, only those in the high cost treatment can expect to influence future play with a "bluff" by entering, as these subjects may expect that the potentially costliness of this decision will significantly influence player 2's beliefs.

Do beliefs reflect these differences within cost treatments and across single versus repeated play? In the high cost case we see that beliefs are significantly lower in the repeat play case after observing entry (t=5.75, MW: z=8.6). Repetition appears to decrease the extent that player 2 is willing to update after observing entry in the high cost case, possibly reflecting their knowledge that more player 1's (with low private values) are attempting to bluff by entering in the first round. We do not see a significant difference in the low cost case, where there was little difference in strategies. Thus indeed beliefs reflect differences (or lack thereof) in the entry choices we observe between the single play and repeat play designs.

What about decisions to resist? We see the same pattern as with entry. In the high cost condition there is signficantly higher levels of resistance (80% vs. 75%, z=-1.63), lower average private values (t=2.01, MW: z=2.22), and lower estimated cutpoints (t=3.4, MW: z=2.6) in the repeat play design. This possibly reflects the lower beliefs by player 2 about entrants in the repeat play case, as the higher levels of resistance indicate a greater suspicion that there were more low type entrants in the repeat play case of the high cost treatment. In the low cost condition we see a higher proportion of subjects choosing resist (z=-1.62), lower average private values for subjects that resist (t=2.07, MW: z=1.61), and slightly lower cutpoints in the repeat play treatment (t=1, z=1.7). Due to the two-sided incomplete information, Player 2 possibly has the incentive to build a reputation as a strong type as well by choosing to resist when there is repetition. We do observe more Resistance in the repeated case for both costs in keeping with this reputation building motive. The equilibrium solution of the repeated play case is needed to better understand these results.

Do player 1 beliefs reflect these differences? Consistent with the strategy differences we see that beliefs are lower on average after observing a resistance decision in the repeat play case (t=1.78, MW: z=4.54). We observe however that in the low cost case beliefs are higher in the repeat play case after observing resistance, though this difference is less significant under the Mann-Whitney test (t=-2.3, z=-1.3). Thus there is at least some evidence that the pattern in differences between beliefs and differences in strategies diverge slightly in the low cost case from the single to repeat play treatments for resistance decisions.

Finally, we compare the effect of cost treatments within the repeat play design. Under the high cost repeat play treatment, player 1's choose to enter 86% of the time. Under the low cost repeat play treatment player 1s chose to enter 79% of the time. This difference is significant (z=2.25). The average private values of those choosing to enter in the first period is also lower in the high cost case (t=-2.1, MW: z=-2.4). We, however, do not find a significant difference in the estimated cutpoint across the two treatments in the repeat play case. These results are the opposite of what we find in the single play case–a high cost of backing down to player one leads to a lower entry cutpoint. Because we have not solved the repeat play model, we do not have theoretical predictions about this. Nevertheless, this difference is interesting and we look forward to pursuing them further. What about for resistance decisions? We see a higher proportion of subjects resisting in the low cost case (z=-2.1), slightly lower but not significant differences of the average private values in the low cost case (t=1.04, MW: z=.6), and estimated cutpoints slightly lower in the low cost case (t=1.67, MW: z=1.6).

Do beliefs reflect any of these differences in strategy choices within the repeat play design? After observing an entry decision, player 2 subjects in the high cost case had significantly lower beliefs on average than in the low cost case (t=-2.3, MW: z=-3.4). Thus beliefs indeed track differences in strategy choices for entry within the repeat play treatment. After observing a resist decision, player 1 subjects reported a higher average belief in the low cost case (t=-3.6, MW: z=-3.92). This is the opposite of what would be predicted given the differences in strategy choices, where the set of subjects resisting had slightly lower private values in the low cost case. We look forward to thinking abut this divergence more in the future.[7]

[7] We of course also explored whether repetition of the game has any effect on fight decisions within a fixed cost treatment. We find that in the high cost case, a lower proportion of subjects choose to Fight in the repeat play case (z=2.63), but there is no difference for the low cost case. Subjects in the high cost case that decided to fight however had a higher average private value in the single play treatment (t=2.03, MW: z=2.1). Estimated cutpoints do not differ for the high cost case, but difference in means and Mann-Whitney tests are close to signficant for the low cost case where repetition tended to increase the estimated cutpoints (t=-1.6, MW: z=-1.5). Thus the repeat play design had a mixed effect on Fight decisions. The effect of repetition tended to decrease both the proportion of subjects fighting, and the average private value of those that fought. This contrasts slightly with our findings on the effect of repetition on entry, where entry was more likely in the high costs case under repeat play, though no different in the single play case. One simple explanation for this is that repeat play induced more subjects to enter (atleast in the high cost case), but upon facing resistance many of these subjects decided not to

5.3 Comparison to Alternative Beliefs

In this section we compare how earnings from the belief elicitation task would have differed if a subject had reported some other belief. This allows us to look at how close subject beliefs were to actual private values of their opponent *relative to other beliefs that they might have reported*. In particular, we investigate six comparisons. *Actual-Equilibrium* takes the earnings the subject actually received and subtracts off the earnings they would have received were they to follow the equilibrium beliefs. Because we do not solve for an equilibrium in the repeat play case, we simply use equilibrium parameters derived from the single play case. For cases considered "off the equilibrium track", for example, for player 2's belief when player 1 does not enter in the low cost case, we use as the equilibrium prediction the mean of the prior, 0. *Actual-Guess* takes the actual earnings and subtracts off what the player would have earned if they had simply taken a random draw from the prior distribution used in the experiment. To do this we programmed a random number generator with a fixed seed to generate a random draw from our distribution. *Actual-MeanPrior* takes the player's actual earnings and subtracts off what they would have earned if they simply guessed the mean of the prior 0. *Actual-Perfect* takes the player's actual earnings and subtracts off their earnings if they had perfectly predicted the other player's private value. *Equilibrium–MeanPrior* takes the player's earnings if they had used the equilibrium predictions and subtracts off the earnings had they always guess the mean of 0. Finally, *Motivated-NonMotivated* takes all the cases where player 1 did not enter, and then calculates for these player 1's their actual earnings (which were often "motivated" in the sense discussed above) and subtracted off what they would have earned had they reported the mean of the prior, 0.

The tables below report the means of these differences for each of our payoff treatments separately, and includes the 95% confidence intervals. Table 8 reports results

fight. We do not focus on this case as it does not relate to our focus on beliefs in this paper.

from all rounds together, table 9 reports results when only the first round of the repeat play case is used. We see that subjects could have moderately improved their belief payoff earnings were they to use equilibrium values, but did substantially better than randomly guessing. Subjects also tended to do better than if they simply guessed the mean of the prior, though this difference was not significantly different from 0 for the singly play high cost case. The negative values on *Actual-Perfect* indicate that subject had considerable room to improve. Comparison of *Actual-Perfect* across the single and repeat conditions illustrates how subjects earned more in the belief elicitation on average in the repeat play conditions. However, if we only consider this difference in the first round of the repeat play sessions, this difference is smaller. We find that this directly due to subject updating their beliefs upon observing a particularly informative course of action within the three rounds of play. For example, if player 1 enters, faces resistance by player 2, and then backs down, player 2 has observed additional information from when their beliefs were first measured. This additional knowledge can then be used to report more accurate beliefs (we return to this below). The positive values of *Equilibrium-MeanPrior* indicate that the selecting the equilibrium expectation was better than playing the mean of 0, though interestingly this difference was much higher in both high cost cases. Finally, we see that subject could have modestly improved their earnings were they to report a belief of 0 after choosing not to Enter. However, this effect appears strongest in both the low cost cases.

Category	**SingleLow**	**SingleHigh**	**RepeatLow**	**RepeatHigh**
Actual-Equilibrium	-.02(-.04,0)	-.12(-.17,-.08)	-.01(-.02,.01)	-.01(-.02,0)
Actual-Guess	.36(.31,.41)	.36(.29,.43)	.38(.34,.42)	.42(.38,.46)
Actual-MeanPrior	.03(.01,.05)	.01(-.04,.06)	.04(.02,.06)	.09(.08,.1)
Actual-Perfect	-.32(-.36,-.29)	-.32(-.37,-.27)	-.28(-.3,-.25)	-.22(-.24,-.21)
Equilibrium-MeanPrior	.05(.04,.06)	.13(.11,.15)	.05(.04,.06)	.1(.09,.12)
Motivated-NonMotivated	-.05(-.08,-.03)	-.01(-.04,.01)	-.04(-.08,0)	.01(-.01,.04)

Table 8: All Rounds

Category	SingleLow	SingleHigh	RepeatLow	RepeatHigh
Actual-Equilibrium	-.02(-.04,0)	-.12(-.17,-.08)	-.02(-.05,0)	-.01(-.02,.01)
Actual-Guess	.36(.31,.41)	.36(.29,.43)	.39(.31,.47)	.37(.31,.44)
Actual-MeanPrior	.03(.01,.05)	.01(-.04,.06)	.02(-.01,.04)	.06(.04,.08)
Actual-Perfect	-.32(-.36,-.29)	-.32(-.37,-.27)	-.3(-.35,-.25)	-.25(-.29,-.22)
Equilibrium-MeanPrior	.05(.04,.06)	.13(.11,.15)	.04(.02,.06)	.07(.04,.09)
Motivated-NonMotivated	-.05(-.08,-.03)	-.01(-.04,.01)	-.02(-.09,.06)	-.01(-.04,.02)

Table 9: Only Round 1 In Repeat Play

5.4 Updating within a Dyad

Another way to look at updating in the repeat play case is to exploit the panel structure of the dyadic interaction. How do beliefs respond to strategic choices made after initial beliefs are elicited? In particular, we coded subjects that observed a course of action over the three rounds of interaction that would lead them to update their beliefs beyond their first round belief. There are two main ways that this can happen. First, player 1 subject can initially enter but then subsequently not fight or not enter in a later period. Second, player 2 can resist in the first period but then subsequently choose to not resist in later period (usually after observing the first player choose to fight). We then identified all cases where players increased their belief earnings from round 1 to round 3. Table 10 reports a simple cross-tabulations of subjects who improved/did not improve their belief payoff and whether or not they observed an "updateable" strategy choice beyond initial choices in the first round. Estimating a probit tells an identical story; observing an "updateable" strategy choice beyond initial round 1 choices signficantly relates to the event of a player increasing their belief payoffs ($\beta = 1.32(.19)$). Thus player beliefs actively respond to events that allow them to adjust the beliefs after reporting their initial beliefs in the first round. A separate project currently is working on how to estimate the process using a structural model (Mukherjee et al. 2007).

	Did Not Improve Belief Payoff	Improved	Total
No Updateable Observations	372	212	584
Row %	(63.70%)	(36.30%)	
Updateable Observations	11	56	67
Row %	(16.42%)	(83.58%)	
Total	383	268	651
Row %	(58.83%)	(41.17%)	

Table 10. Updating beliefs within a dyad

5.5 Best Responses

Another important check on the assumptions of the PBE is whether the actions chosen by the players are optimal give their beliefs. That is, are the players best responding given their expectations about the other player's private value? We can only answer this question for Player 2 and not for Player 1 because the expectation of Player 1 about Player 2's private value is elicited at Player 1's last decision node after Player 2 has decided to Resist or not Resist. The optimality of Player 1's decision to Fight or not Fight is not dependent on her expectation of Player 2's private value at that last decision node. To examine whether the Player 2s are best responding given their expectations and the other player's actions, we compare Player 2's expected payoff under Resistance vs. no Resistance with the assumption that Player 1 behaves according to the PBE. If Player 2 chose the action that yields the higher expected payoff, we code it as a 1 and code it as a 0 otherwise to construct our hit-rate measure. We find that under this measure, Player 2s best respond 94% of the time in the single play high cost treatment and 92% of the time in the other three treatments (single play low cost, repeat play low and high cost). With our experimental design in which actions are taken and beliefs are elicited together, we are unable to distinguish between the possibilities of actions best responding to beliefs and stated beliefs that best respond to/rationalize the chosen actions. We can only conclude that this result is consistent with the interpretation that subjects in the Player 2 role best responded to their expectations about the other player's private value under our equilibrium assumptions an overwhelming majority of the time. This does suggest that, contrary to previous findings (Costa-Gomes and Weiszacker 2006), a player's elicited beliefs are strongly predictive of her action choice under the principle of best response.

6 Conclusion

We have provided one of the first (only?) experimental tests of a sequential game of two-sided incomplete information. By using the Perfect Bayesian Equilibrium concept we derive a number of predictions and comparative statics that we then test using incentivized laboratory experiments. We also directly observe beliefs and updating using an innovative elicitation procedure. This procedure is proper and incentivized, in that it does not give anyone an incentive to report anything other than their true beliefs. Thus we explore an experimentally understudied class of game theoretic models, and use innovative experimental procedures to measure beliefs. Because beliefs play a critical role in our theories we feel this is an important contribution.

We investigated a simple 2x2 design: high versus low cost of not fighting and single versus repeat play. This not only gave us a richer set of data to explore the intersection of strategies and beliefs, but it also captures important substantive concepts in both the economics and political science literatures. A highlight of using this crossed design is exploring the differential effect of repetition on entry rates across the cost treatments. We find that giving player 1 a high cost of backing down decreases their incentive to entry in the single play case, but that repetition actually increases the incentive to entry. Furthermore, beliefs pick-up on these changes. Insofar as discussions about credibility turn on concepts like costs (e.g., audience costs in the international relations literature), we feel this is an important finding that deserves more theoretical attention in the future.

There are two important behavioral findings. The first is what we have termed 'motivated beliefs'. In this case, when player 1 decides to not enter, presumably because they have a bad private value, their reported belief about player 2's type is significantly greater than zero. This is despite the fact that they should not have updated their prior. Our second behavioral finding also involves beliefs. Here, we find that updating is 'sticky' around the prior. Subjects regularly updated their beliefs less in magnitude than they should have according to the model. Future research will push this later finding further by calculating the extent to which subjects should update their beliefs given actual behavior in the experiment, instead of setting the benchmark as the theoretical predictions of the model.

Future research will work on estimating a Quantal Response Equilibrium model to try and get a better fit to the data. We also plan to exploit our findings about motivated and conservative beliefs, potentially looking at how these patterns are found in other strategic settings. We are also in the process of setting up a similar set of experiments but with a different, simpler, distribution governing the generation of player types. We are interested in how properties of the type generating function relate to how subjects update. More generally, we are interested in how different strategic contexts (i.e., the structure of the game) make beliefs more or less relevant to strategy choices. Finally, we hope to provide clearer theoretical differences between the single and repeat play versions of our game. Unfortunately, solving such models is at the forefront of game theoretic research and the appropriate technologies may not yet be developed.

7 Appendix

We now characterize an equilibrium solution for the entire range of payoffs to player 1 for not Fighting (B_1). Furthermore, we show that the equilibria we found for the low cost treatment and the high cost treatment are indeed unique.

There are three possible action profiles that Player 1 could have: Not Enter, Enter and Not Fight, Enter and Fight. We now look for the general class of equilibria where the private value of Player 1 is partitioned into three regions with each region mapping to a possible action profile. Note that Player 1's private value does not affect the comparison of payoffs in the Not Enter vs. Enter and Not Fight cases because the private value only enters into the payoff when Fight is chosen. Therefore Player 1 has the same preference over the two action profiles regardless of private value. If Player 1 strictly prefers Entering and Not Fighting to Not Entering, then Player 1 will always Enter and then decide to Fight or Not Fight subject to incentive compatibility at the last node. Note that this is the type of equilibrium we find for the Low Cost, $B_1 = -0.1$. If Player 1 strictly prefers Not Entering to Entering and Not Fighting, then Player 1 will either Enter or Not Enter with some cutoff rule and then once she Enters, she will always Fight. Note that this is the type of equilibrium we find for the High Cost, $B_1 = -0.3$ case. The third possibility is that Player 1 is indifferent between Not Entering and Entering and Not Fighting. In this case, for the private values under which Player 1 would not Enter and Fight, we can partition them any way we want for the Not Enter region and the Enter and Not Fight region due to the indifference.

We will solve for the action probabilities that can sustain this indifference condition and show that in the Low Cost and High Cost treatments, the indifference condition cannot be sustained and the equilibria we solved for are in fact the unique one. In fact, we show that above a certain threshold B_1^H only the equilibrium in which Player 1 either chooses to Enter and Fight or Enter and Not Fight is sustained. We also show that below a certain threshold B_1^L only the equilibrium in which Player 1 either chooses to Not Enter or Enter and Fight is sustained. For the intermediate region between B_1^H and B_1^L, we do find an equilibrium in which the distribution of private values has three partitions and each one maps to an action profile.

7.1 Indifference Condition

Player 1 is indifferent between Not Entering and Entering but Not Fighting if

$$1 - \Pr(R|E)]D_1 + \Pr(R|E)B_1 = 0$$

Which simplifies to

$\Pr(R|E) = \frac{D_1}{D_1 - B_1}$

From the definition of $\Pr(R|E)$ in Section 3.1, we know that

$$\begin{aligned}\Pr(R|E) &= \Pr(\mu_2 > \frac{\Pr(F|E)(D_2 - W_2) + B_2 - D_2}{\Pr(F|E)}) = \\ &= \Pr(\mu_2 > \frac{-0.8}{\Pr(F|E)}) = 1 - \frac{1}{1 + e^{\frac{0.8}{\Pr(F|E)}}}\end{aligned}$$

From this, we can solve for the probability of Fight given Entry that would sustain the indifference condition

$$\Pr(F|E) = \frac{0.8}{ln(\frac{Pr(R|E)}{1 - Pr(R|E)})} \tag{4}$$

7.2 Low Cost: $B_1 = -0.1$

We first consider the Low Cost treatment, $B_1 = -0.1$ and show that the indifference condition cannot be sustained. Plugging in the payoff values, the probability of Player 2 Resisting given Entry by Player 1 is

$\Pr(R|E) = \frac{D_1}{D_1 - B_1} = \frac{0.5}{0.6} = 0.833\,33$

According to (4), $\Pr(F|E) = 0.49708 = \frac{\Pr(E \cap F)}{\Pr(E)}$

We know that for all Player 1s who Enter, those with $\mu_1 > B_1 - W_1$ will Fight so

$$\Pr(E \cap F) = \Pr(W_1 + \mu_1 > B_1) = \int_{-0.6}^{\infty} \frac{e^x}{(1 + e^x)^2} dx = 0.65$$

In order for $\Pr(F|E)$ to be 0.49708, $\Pr(E)$ must take an impossible value of greater than 1.

For plausible values of $\Pr(F|E)$, that is $\Pr(F|E) \geq 0.65$), $\Pr(R|E) < 0.833\,33$ according to (4). Therefore $[1 - \Pr(R|E)]D_1 + \Pr(R|E)B_1 > 0$ indicating that choosing to Enter and Not Fight is always preferred to choosing to Not Enter. By incentive compatibility, Enter and Fight would not be chosen over Enter and Not unless it was preferred. Thus, in the Low Cost treatment, Enter is always preferred to Not Enter. The equilibrium we characterized in Section 3.1.1 where $\Pr(E) = 1$ is in fact the only one for $B_1 = -0.1$ because Player 1 would not want to Not Enter instead regardless of her private value under this payoff structure.

To find the threshold B_1^H above which only the equilibrium with Player 1 choosing either Enter and Fight or Enter and Not Fight is sustained, we first note that $\Pr(F|E) = \frac{\Pr(E \cap F)}{\Pr(E)} = 0.65$ if $\Pr(E) = 1$. From (4) we know that $\Pr(R|E) = 0.774$ and under the indifference condition, $B_1 = -0.145$. If $B_1 > -0.145$, then $[1 - \Pr(R|E)]D_1 + \Pr(R|E)B_1 > 0$ meaning that Enter and Not Fight is always preferred to Not Enter. Therefore $B_1^H = -0.145$.

7.3 High Cost: $B_1 = -0.3$

We now consider the High Cost treatment, $B_1 = -0.3$ and show that indifference condition again cannot be sustained. Plugging in the payoff values, the probability of Resistance given Entry is

$\Pr(R|E) = \frac{D_1}{D_1 - B_1} = \frac{0.5}{0.8} = 0.625$

From (4), we know that $\Pr(F|E) = 1.5661$ which is not possible.

For feasible values of $\Pr(F|E)$, that is ($\Pr(F|E) \leq 1$), $\Pr(R|E) > 0.625$ according to (4). It follows that $[1 - \Pr(R|E)]D_1 + \Pr(R|E)B_1 < 0$ so that Not Enter is always preferred to Enter and Not Fight. If Not Enter is preferred to Enter and Not Fight, then

Player 1 would always choose to Fight if she chooses to Enter or else she would choose to Not Enter in the the first place. Therefore The equilibrium we characterized in Section 3.1.2 where $\Pr(F|E) = 1$ is in fact the only one for $B_1 = -0.3$.

To find the threshold B_1^L below which only the equilibrium with Player 1 choosing either Enter and Fight or Not Enter is sustained, we first note that for $\Pr(F|E) = 1$, $\Pr(R|E) = 0.690$ according to (4). For this $\Pr(R|E)$, $B_1 = -0.225$ under the indifference condition. if $B_1 < -0.225$, then $[1 - \Pr(R|E)]D_1 + \Pr(R|E)B_1 < 0$ meaning that Not Enter is always preferred to Enter and Not Fight. Therefore $B_1^L = -0.225$.

7.4 Intermediate Region: $B_1^L < B_1 < B_1^H$

We now show that in this intermediate region of B_1 values, there is an equilibrium with three partitions of Player 1 private values, where those with the lowest private values will choose to Not Enter, those in the middle region of private values choose to Enter and Not Fight, and those with the highest private values will Enter and Fight. The indifference condition must hold in this equilibrium so we know that

$\Pr(R|E) = \frac{D_1}{D_1 - B_1}$

(4) gives us the probability of Fight after Entry given $\Pr(R|E)$

$\Pr(F|E) = \frac{0.8}{ln(\frac{\Pr(R|E)}{1-\Pr(R|E)})}$

We know that for all Player 1s who Enter, those with $\mu_1 > B_1 - W_1$ will Fight so

$$\Pr(E \cap F) = \Pr(\mu_1 > B_1 - W_1) = 1 - \frac{1}{1 + e^{W_1 - B_1}}$$

To find the cutoff point for Entry, C, we first note that

$\Pr(E) = \Pr(\mu_1 > C) = \frac{\Pr(E \cap F)}{\Pr(F|E)}$

After plugging in $\Pr(E \cap F)$ and $\Pr(F|E)$, we have

$1 - \frac{1}{1+e^{-C}} = \frac{\frac{0.8}{ln(\frac{\Pr(R|E)}{1-\Pr(R|E)})}}{1 - \frac{1}{1+e^{W_1 - B_1}}}$

Letting $\alpha = \frac{\frac{0.8}{ln(\frac{\Pr(R|E)}{1-\Pr(R|E)})}}{1 - \frac{1}{1+e^{W_1 - B_1}}}$, we find

$C = -ln(\frac{\alpha}{1-\alpha})$

Thus, in this equilibrium, Player 1 with $\mu_1 \geq B_1 - W_1$ Enters and Fights, Player 1 with $C \leq \mu_1 < B_1 - W_1$ Enters and Not Fights, and Player 1 with $\mu_1 < C$ chooses to Not Enter.

To illustrate this equilibrium solution, we use the example $B_1 = -0.2$ which lies in this intermediate region. First, we know that under the indifference condition

$\Pr(R|E) = \frac{D_1}{D_1 - B_1} = 0.714$

Next, we solve for probability of Fight given Entry using (4)

$\Pr(F|E) = \frac{0.8}{ln(\frac{\Pr(R|E)}{1-\Pr(R|E)})} = 0.873$

Using the cutoff value for Enter and Fight that follows from the incentive compatibility condition at the terminal node, $\mu_1 > B_1 - W_1 = -0.7$, we solve for the probability of Enter and Fight

$\Pr(E \cap F) = 1 - \frac{1}{1+e^{0.7}} = 0.668$

Finally, we find the cutoff value for Entry

$C = -ln(\frac{\frac{0.873}{0.668}}{1-\frac{0.873}{0.668}}) = -1.182$

Thus, in this example, Player 1 with $\mu_1 \geq -0.7$ Enters and Fights, Player 1 with $-1.182 \leq \mu_1 < -0.7$ Enters and Not Fights, and Player 1 with $\mu_1 < -1.182$ chooses to Not Enter.

We have shown that above a certain threshold B_1^H only the equilibrium in which Player 1 either chooses to Enter and Fight or Enter and Not Fight is sustained. We also show that below a certain threshold B_1^L only the equilibrium in which Player 1 either chooses to Not Enter or Enter and Fight is sustained. For the intermediate region between B_1^H and B_1^L, we do find an equilibrium in which the distribution of private values has three partitions and each one maps to an action profile, Not Enter, Enter and Not Fight, or Enter and Fight. For the payoff values we have chosen for our experiment, $B_1^H = -0.145$ and $B_1^L = -0.225$. Thus for our low-cost of Player 1 backing down treatment, $B_1 = -0.1$, the unique PBE is one in which Player 1 either chooses to Enter and Fight or to Enter and not Fight so she Enters for certain regardless of private value. For our high-cost treatment, $B_1 = -0.3$, the unique PBE is one in which Player 1 either chooses to Not Enter or to Enter and Fight so that she will Fight for certain given Entry. We also show that had we picked a cost of Player 1 backing down in the intermediate region, then there exists a PBE such that Player 1 would choose one of all three action profiles, Not Enter, Enter and Not Fight, or Enter and Fight, depending on her private value.

References

[1] Akerlof, G. A. and W. T. Dickens (1982): "The Economic Consequences of Cognitive Dissonance" *American Economic Review*, 72, 307-319

[2] Benabou, R. and J. Tirole (2006): "Belief in a Just World and Redistributive Politics" *Quarterly Journal of Economics*, 121, 699-746.

[3] Bolton, G. E. and A. Ockenfels (2007): "Information Externalities, Matching and Reputation Building - A Comment on Theory and an Experiment" *Working Paper.*

[4] Brandts, J. and C. A. Holt (1993): "Adjustment Patterns and Equilibrium Selection in Experimental Signaling Games" *International Journal of Game Theory,* 22, 279-302.

[5] Camerer, C. and K. Weigelt (1988): "Experimental Tests of a Sequential Equilibrium Model" *Econometrica,* 56, 1-36.

[6] Carrillo, J. D. and T. R. Palfrey (2006): "The Compromise Game: Two-sided Adverse Selection in the Laboratory" *Working Paper.*

[7] Charness, G. and M. Dufwenberg (2006): "Promises and Partnership" *Econometrica*, 74, 1579-1601.

[8] Costa-Gomes, M. and G. Weiszacker (2006): "Stated Beliefs and Play in Normal Form Games" *Working Paper.*

[9] Croson, R. T. A. (2000): “Thinking like a Game Theorist: Factors Affecting the Frequency of Equilibrium Play” *Journal of Economic Behavior and Organization*, 41, 299-314.

[10] Dominitz, J. and A. Hung (2004): "Homogeneous Actions and Hetergeneous Beliefs: Experimental Evidence on the Formation of Information Cascades" *Working Paper.*

[11] Dufwenberg, M. and U. Gneezy (2000): "Measuring Beliefs in an Experimental Lost Wallet Game" *Games & Economic Behavior*, 30, 163-82.

[12] Eckel, C. C. and P. J. Grossman (2008): "Forecasting Risk Attitudes: An Experimental Study Using Actual and Forecast Gamble Choices" *Journal of Economic Behavior and Organization*, Forthcoming.

[13] Edwards, W. (1968): "Conservatism in Human Information Processing" *Formal Representation of Human Judgment*, Kleinmuntz, B. (ed.), New York.

[14] Esarey, J. E., S. Mukherjee, and W. H. Moore (2006): "A Fixed-Effects Bayesian Quantal Response Estimator for Incomplete Information Games" *Working Paper.*

[15] Holt, C. A. and S. Laury (2002): "Risk Aversion and Incentives: New Data Without Order Effects" *American Economic Review,* 95, 902-912.

[16] Huck, S. and G. Weizsäcker (2002): "Do Players Correctly Estimate What Others Do? Evidence of Conservatism in Beliefs" *Journal of Economic Behavior and Organization*, 47, 71-85.

[17] Jung, Y. J., J. H. Kagel, and D. Levin (1994): "On the Existence of Predatory Pricing: an Experimental Study of Reputation and Entry Deterrence in the Chain-Store Game" *The Rand Journal of Economics,* 25, 72-93.

[18] Kadane, J. and R. Winkler. (1988), "Separating Probability Elicitation from Utilities" *Journal of the American Statistical Association*, 83, 357-63.

[19] Lewis, J. B. and K. A. Schultz (2003): "Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information" *Political Analysis*, 11, 345-367.

[20] Manski, C. (2004): “Measuring Expectations” *Econometrica*, 72, 1329-1376.

[21] Mather, M., E. Shafir, and M. K. Johnson (2000): "Misremembrance of Options Past: Source Monitoring and Choice" *Psychological Science*, 11, 132-138.

[22] McKelvey, R. D., and T. Page (1990): "Public and Private Information: An Experimental Study of Information Pooling" *Econometrica*, 58, 1321-1339.

[23] Mukherjee, S., D. H. Tingley, and S. W. Wang (2007): "Comparing Statistical Estimators for the Crisis Bargaining Model" *Working Paper.*

[24] Nyarko, Y. and A. Schotter (2002): " An Experimental Study of Belief Learning Using Elicited Beliefs" *Econometrica*, 70, 971-1005.

[25] Offerman, T., J. Sonnemans, and A. Schram (1996): "Value Orientations, Expectations, and Voluntary Contributions in Public Goods" *Economic Journal*, 106, 817-45.

[26] Palfrey, T. R. and S. W. Wang (2007): "On Eliciting Beliefs in Strategic Games" *Working Paper.*

[27] Risen, J. L., and T. Gilovich (2007): "Another Look at Why People Are Reluctant to Exchange Lottery Tickets" *Journal of Personality and Social Psychology*, 93, 12-22.

[28] Savage, L. J. (1971): "Elicitation of Personal Probabilities and Expectations" *Journal of the American Statistical Association*, 66, 783-801.

[29] Tingley, D. H. and B. Walter (2007): "Reputation Building in International Relations: An Experimental Approach" *Working Paper.*

[30] Tingley, D. H. and B. Walter (2008): "Can Cheap Talk Help Build Reputations? An Experimental Analysis" *Working Paper.*

[31] Wand, Jonathan (2006): "Comparing Models of Strategic Choice: The Role of Uncertainty and Signaling." Political Analysis 14(101-120).