Belief Updating in Sequential Games of Two-Sided Incomplete Information: An Experimental Study of a Crisis Bargaining Model[1]

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

September 30, 2010

[2]Government Department, Harvard University, Cambridge, MA 02138.

[3]Division of Humanities and Social Sciences, Caltech, Pasadena, CA 91125.

Abstract

We investigate theoretically and experimentally the "crisis bargaining model," a dynamic game of two-sided incomplete information with player types drawn from a commonly known distribution. Little work has been done to analyze whether and how players update their beliefs in such games. Within the experiment we elicited beliefs from players about their opponent's type using a proper scoring rule. We implement two treatments that vary the cost of backing down to the first mover after initial entry, generating sharp comparative static predictions in both beliefs and strategies. We find that players do update their beliefs in the predicted directions after observing some of the action choices. However, we highlight evidence of conservative updating relative to rational expectations.

1   Introduction

We study experimentally a strategic model of conflict, the "crisis bargaining model," widely used in the international relations literature (Fearon, 1994; Lewis and Schultz, 2003; Schultz, 2001; Esarey *et al.*, 2008) but also related to work on sequential games of two-sided incomplete information in economics and other literatures (Kreps and Wilson, 1982; Milgrom and Roberts, 1982; Wilson, 1992; Carrillo and Palfrey, 2009; Fudenberg and Tirole, 1991, chap. 9) that study the role of signalling. Although theoretical predictions about this game using solution concepts such as the perfect Bayesian equilibrium (PBE) are well known, less is known about actual strategic behavior and beliefs in such dynamic games of two-sided incomplete information. Our focus is on the beliefs that both agents hold about the other agent's private information and how these beliefs are updated. Such beliefs are at the foundation of strategic interactions with private information but it is difficult to observe beliefs directly with field data (Manski, 2004) and here we take an experimental approach.

In our experimental setting, we elicit beliefs about the opponent's private information at each decision node during game play using a strictly proper quadratic scoring rule. We vary the payoff to one of the players at a potential terminal node as our treatment variable. This payoff represents a cost to one player for initiating bargaining but then backing down. Belief updating should be related to the magnitude of the cost. The high versus low cost variation to the first player generates sharp differences in PBE behavior and beliefs. Our framework allows us to 1) test comparative statics of belief updating at each decision node depending on whether the cost for backing down is high or low and 2) characterize specific departures from equilibrium beliefs. Although the direction of belief updating usually conforms to theory, we find evidence of conservatism in beliefs rather than rational expectations.

We survey the existing literature on experimental sequential games, beliefs and expectations, and applications in economics and political science in Section 2. In Section 3, we describe the game we study and solve for the perfect Bayesian Equilibrium. We derive the equilibrium cutpoints and expectations for both treatments and propose several

hypotheses based on the comparative statics. We present the experimental design and procedures in Section 4. Section 5 presents our results. We test theoretical predictions about belief updating and highlight conservative updating relative to rational expectations. We conclude with an overview of our findings and suggestions for future work in Section 6.

2   Experimental Studies of Updating

There has been a wealth of theoretical work on sequential games with one-sided and two-sided incomplete information. While a corresponding experimental literature exists on sequential games with one-sided incomplete information (Camerer and Weigelt, 1988; Brandts and Holt, 1992; Banks *et al.*, 1994; Jung *et al.*, 1994; Cooper *et al.*, 1997; Cai and Wang, 2006; Tingley and Walter, 2011; Drouvelis *et al.*, 2009), far fewer experimental studies focus on two or multi-sided incomplete information sequential games (McKelvey and Page, 2000; Carrillo and Palfrey, 2009; Dominitz and Hung, 2009). Furthermore, few studies directly investigate the role of beliefs through belief elicitation even though belief formation and updating figure prominently in sequential equilibrium concepts.

Most elicited beliefs in experimental studies can be categorized into three types: probabilistic beliefs about opponents' actions in matrix games (Nyarko and Schotter, 2002; Costa-Gomes and Weiszäcker, 2008), probabilistic beliefs about the state of the world (McKelvey and Page, 1990; Dominitz and Hung, 2009), 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). Beliefs about the other player's private information, his type in this case, are not elicited as often (Dominitz and Hung, 2009). Furthermore, beliefs are usually elicited simultaneously from the players even in sequential games (Dufwenberg and Gneezy, 2000; Charness and Dufwenberg, 2006) so that the response of beliefs to observed actions cannot be ascertained.

We fill these gaps in the literature 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. Our elicitation mechanism is an easily understood incentive compatible

Savage (1971) analog of the quadratic scoring rule for probabilistic forecasts: the deduction from a fixed payoff is the squared distance between the expectation and the realized value of the variable. By eliciting beliefs about the opponent's private value rather than chosen action, we can test equilibrium predictions for beliefs that previous studies have not been able to address.

3   Model and Hypotheses

Our game follows from the crisis bargaining literature (Lewis and Schultz, 2003; Esarey *et al.*, 2008). There are two players, 1 and 2. The sequence of moves of this game, with generic payoff notation (left figure) and with the specific numerical payoffs for each treatment provided to subjects, is depicted in Figure 1.

Figure 1
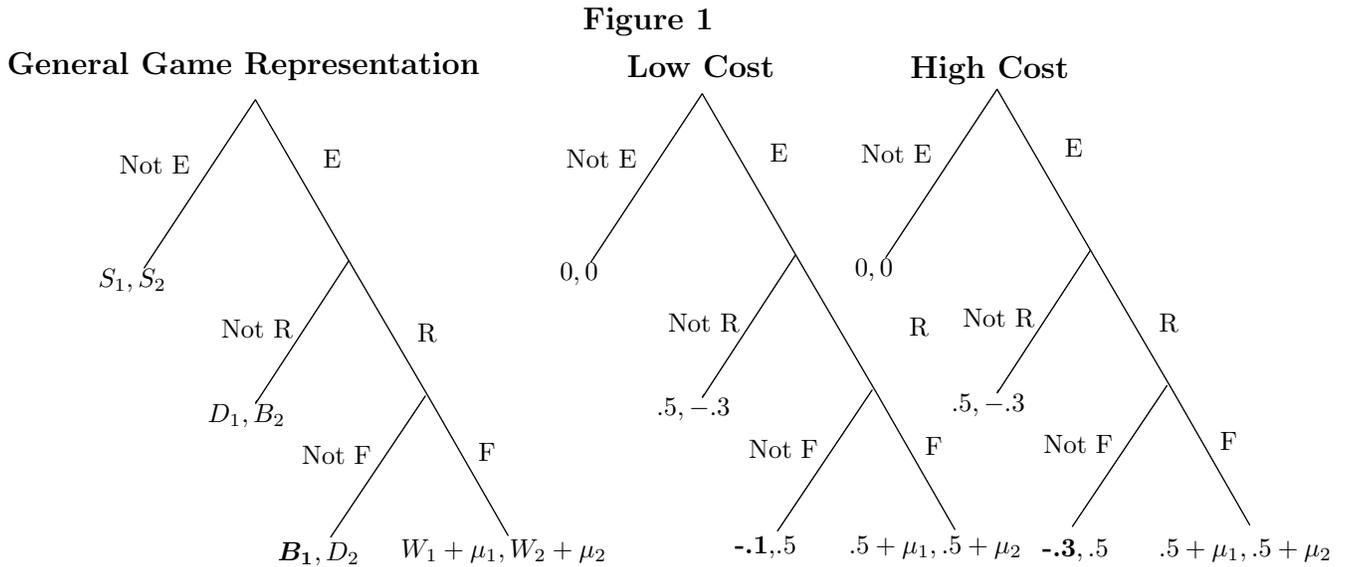
First, Nature draws each player's type, μ_i, independently from the same continuous logistic distribution with PDF $g(x) = \frac{e^x}{(1+e^x)^2}$. The distribution from which the types are drawn is common knowledge while the actual types of the respective players are private information. Player 1, the potential entrant, then chooses whether or not to Enter (E or not E) and challenge player 2. If 1 chooses not to Enter, the status quo prevails and the players receive $S_1, S_2 = 0$. If 1 does Enter, then 2 decides whether or not to Resist 1's challenge (R or Not R). If 2 does not Resist, the game ends with 2 conceding to 1 and the players receive $D_1 = .5$ and $B_2 = -.3$. If player 2 Resists, then 1 must decide whether or

not to Fight (F or Not F). If 1 chooses not to Fight, then the game ends and the players receive $B_1 < 0$ and $D_2 = .5$. We vary the fixed value of B_1 to a high cost $(-.3)$ and low cost $(-.1)$ level to generate crisp comparative static predictions that we detail below. If 1 decides to Fight then the game ends and the players receive a fixed payoff, $W_1, W_2 = .5$, plus their respective private values μ_1, μ_2. The equilibrium strategies are characterized by equilibrium cutpoints which players 1 and 2 use to decide between the binary choices 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.

3.1 Perfect Bayesian Equilibrium and Hypotheses

We solve for the perfect Bayesian Equilibria using the standard backward induction method and assuming risk neutrality of the players. A summary of results is presented in Table 1 and a complete derivation is presented in the supplementary materials. Here we sketch out the intuition behind the equilibrium strategies which differ as a function of B_1, the cost to Player 1 of choosing to Not Fight and our treatment variable in the experiment.

If the cost of choosing Not Fight is low enough, as in our low cost treatment, then there is an unique equilibrium in which Player 1s choose Enter regardless of their private value. In this case, the potential gains from having Player 2 choosing Not Resist are sufficiently high compared to the small cost of backing down that all Player 1s pool on Enter. Suppose all Player 1s choose to Enter. The probability of a Player 1 choosing Fight given Enter is substantially less than 1 because all players who chose Enter but have a private value below the cutpoint for Fight ($\mu_1 = -0.80$) will choose Not Fight. Player 2, knowing that some portion of Player 1s will ultimately choose to Not Fight, have a low cutpoint ($\mu_2 = -1.23$) and hence are more likely to choose Resist. In this low cost condition, the probability of choosing Resist given Enter is high since the probability of facing Fight is low. Nevertheless, when the cost of choosing Not Fight is sufficiently low (e.g. $B_1 = -0.1$), the expected gain outweighs the expected loss and Enter is always preferred to Not Enter at the first node.

On the other hand, if the cost of choosing Not Fight is high enough, as in our high

cost treatment ($B_1 = -0.3$), then there is an unique equilibrium in which Player 1s below a cutpoint ($\mu_1 = -0.72$) choose Not Enter and Player 1s who choose Enter in equilibrium also always choose Fight at the last node. In contrast with the low cost treatment there is separation at the Entry stage in the high cost condition. The Player 1s with sufficiently low private values will not choose Enter because if the last node is reached, the loss there is a sufficiently high deterrent. Suppose that the probability of Fight given Enter is 1. Then the probability of Resist given Enter is low because of the maximal probability of facing Fight. The cutpoint for Player 2s to choose Resist is in this case higher than in the low cost condition ($\mu_2 = -0.8$) because fewer will want to face a Fight choice. Still, the potential loss from Player 2 choosing Resist and then having to choose Not Fight is high enough that Not Enter is always preferred to Enter then Not Fight when the cost of choosing Not Fight is high enough (e.g. $B_1 = -0.3$). Thus, if Player 1 chooses Enter, Fight must be chosen at the final node in the unique equilibrium.[1]

For both the Low Cost and High Cost treatments, we report in Table 1 the expectations one player should have (E_i) about the other player's private value after every possible action along with the cutpoints (μ_1^F, μ_2^R, and μ_1^E) at each decision node for both cost treatments. The cutpoints give us the truncated distribution from which the players form their equilibrium beliefs about the expectation of the other player's value. Calculations for all cutpoints and expectations in both treatments can be found in the supplementary materials.

We focus on two testable comparative statics for the low versus high cost treatment from the equilibrium beliefs. We focus on the beliefs instead of action profiles, though in equilibrium they are related. The supplementary materials contain relevant tests of

---

[1]In the supplementary materials we show that there exists a threshold B_1^H such that for all $B_1 > B_1^H$ Player 1 always chooses to Enter in the unique PBE. The B_1 we chose for our low cost treatment, -0.1, is higher than B_1^H. We also show there exists a threshold B_1^L such that for all $B_1 < B_1^L$, Player 1 always choose Fight given Entry in the unique PBE. The B_1 we chose for our high cost treatment, -0.3, is lower than B_1^L. For intermediate values of B_1, Player 1 could be made indifferent between Not Entering versus Entering and Not Fighting given a certain probability of Resist. In this case, for the private values under which Player 1 would not choose 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. The supplementary materials also derive an Agent-Quantal Response Equilibrium (AQRE) model. The direction of the comparative statics for the AQRE are similar to that of the PBE though we note several important differences below.

	Backdown payoff	
	Low Cost: $B_1 = -0.1$	High Cost: $B_1 = -0.3$
Cutpoints		
μ_1^E	$\forall \mu_1$	-0.72
μ_2^R	-1.23	-0.80
μ_1^F	-0.60	-0.80
Beliefs		
E_2: Enter	0	0.94
E_2: Not Enter	N/A	-1.93
E_1: Resist	0.69	.9
E_1: Not Resist	-2.36	-2
E_1: N/A	0	0

Table 1: Equilibrium Predictions

whether action choices conform to the PBE predictions.

Hypothesis B1: After Player 1 has chosen E, Player 2's belief about Player 1's private value should be higher when $B_1 = -0.3$ than when $B_1 = -0.1$.

Hypothesis B2: 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 $B_1 = -0.3$ than when $B_1 = -0.1$

4   Experimental Design and Procedures

We conducted 12 sessions with a total of 96 subjects from spring 2007 to spring 2008. The subjects were registered undergraduate students at a university who were recruited by e-mail solicitation. Sessions were conducted in a computer lab and all interaction was computerized and anonymous. No subject participated in more than one session. The cost to Player 1 of choosing to not Fight was fixed within an experimental session.

After entering the laboratory, subjects were seated at workstations with dividers and were given several instructional worksheets to read over. A set of instructions was also presented verbally with projected slides. Subjects were informed of the structure of the game (strategies and payoffs), re-matching procedure, and a thorough description of the properties of the distribution from which private values were drawn. During this period subjects were guided through an example series of interactions using their computers, and then were asked to answer a series of comprehension questions. They could not proceed unless all answers were correct.

In the experiment, two subjects were paired together and randomly assigned both 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 before being paired with someone else. In each pair Player 2 would report their belief about Player 1's type following Player 1's choice of "E" or "not E" and Player 1 would report their belief following either Player 2's choice or if "not E" was chosen. To incentivize this belief elicitation we utilized the incentive compatible quadratic scoring rule (Savage, 1971) 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 $\beta = 0.1$.[2] While the realized payoffs from beliefs are lower than the realized payoffs from game play on average, we believe that the payoff for accuracy and the substantial punishment for inaccuracy were sufficient incentives for forming and updating to accurate beliefs. We present further evidence in the results section that subjects were incentivized to state sensible beliefs.[3]

Subjects were re-paired after each round and received randomly assigned private values and positions. When all subjects had been paired with all other subjects once, we repeated the experiment with randomly matched subjects again. We repeated the experiment 9 times. We randomly selected one repetition of the experiment with which to pay subjects, and reminded subjects of this procedure so as to maintain their focus.

5 Results

Table 2 presents several summary statistics. The gender ratio is close to 50/50 in the low cost treatment and a slightly larger majority of subjects are female in the high cost treatment. Because we found relatively little gender difference, we are not concerned

---

[2]Incentivizing beliefs can generate various problems. Hedging (Blanco *et al.*, 2011) is less of a concern here because we are eliciting beliefs about the opponent's private value rather than action. Incentivization might create additional incentives for making particular strategy choices, such as an additional (unmodelled) incentive to enter in order to find out information about their opponent. This is a valid concern and one that we had to trade off against our desire to elicit beliefs and strategies from the same individual. In practice, the total difference in belief payoffs between entrants deciding not to enter and those entering was small enough to not systematically change the comparative static predictions. Alternative designs such as using observers who only make predictions about private values would ensure greater focus on belief formation and updating (Offerman *et al.*, 1996; Palfrey and Wang, 2009).

[3]At the bottom of the screen all subjects were provided with a history table of the current period's decisions, point earnings from strategy choices, and reported belief about the other subject's private value. Point earnings from beliefs not reported until completing all interactions with that subject.

	Private Value		Belief		Game Payoff		Belief Payoff		%Female	N
Treatment	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.		
High Cost	0.02	1.82	0.11	1.49	0.66	1.14	0.15	1.12	65%	3024
Low Cost	-0.01	1.84	0.15	1.21	0.68	1.23	0.16	0.72	56%	2968

Table 2: Summary statistics by treatment condition

about this imbalance.[4] In addition to a show-up fee of $10, subjects in the high and low cost treatment earned on average $7 with a maximum of $17.[5]

5.1 Belief Accuracy and Best Response

We first present evidence that subjects are stating sensible beliefs before going on to analyze the elicited beliefs in more detail. Subjects' beliefs are significantly more accurate than beliefs randomly drawn from the logit distribution. To measure this difference we took each observation and calculated the difference between what was earned based on the subject's stated belief and what they would have earned if they had stated a randomly drawn belief taken from a logit distribution. In the low cost treatment this difference was on average $.34(.26, .41)$ (standard errors clustered at subject level and 95% confidence intervals in parentheses) and $.3(.15, .43)$ for the high cost treatment. Given that a subject could earn at most .5 points for a belief that guessed exactly the other player's private value, this is a substantial improvement.[6]

5.2 Testing the Belief Hypotheses

We find support for Hypothesis B1 which states that following an Enter choice Player 2's expectation of Player 1's private value should be higher when the cost of backing down is higher. Player 2's belief does get updated to a higher level after observing Entry in the high cost case (.77) compared with the low cost case (.58). A difference in means

---

[4]Results available from author.

[5]A computer malfunction caused us to lose the first round of data from our final experimental session in the low cost treatment, and hence the sample sizes are slightly different.

[6]Player 2's expectation of Player 1's type also allows a rough test of whether action choices are optimal given stated beliefs. Under the assumption that Player 1 behaves according to the PBE, Player 2's belief about Player 1's private value also implies a belief about the probability of Fight. Using this probability, we calculate that 91% and 95% of Player 2 decisions were best responses in the low and high cost case, respectively. Actual Player 1 decisions reveal similar high levels of best response rates (97% in the high cost treatment and 95% in the low cost treatment). These high rates of best response indicate that beliefs are mostly consistent with rather than orthogonal to action choices, though we acknowledge that the best response function is constant for wide ranges of beliefs/private values.

test with standard errors clustered by individual, which we use below as well, shows these beliefs are significantly different from each other ($p < .05$, one-tailed tests used for directional hypotheses). Player 2s appear to understand that Player 1s who Enter in the high cost treatment have, on average, higher private values than Player 1s who enter in the low cost treatment because the cost of choosing to Not Fight down the line is higher.[7] Substantively, this means that decision-makers who are challenged in a conflict adjust their beliefs conditional on the incentives of their opponent further down the game tree. If, for example, backing down is relatively cheap for player 1 at the final node, then the amount of updating will be more moderate. Higher costs, such as those hypothesized in the audience cost literature in international relations, will lead to more updating (e.g., Tomz, 2007).

We do find substantial separation in Player 1 beliefs after observing a Resist choice versus Not Resist in both the low and high cost treatments. As with the Entry choice, subjects were on average updating in the predicted direction. However, our results do not support Hypothesis B2 at conventional significance levels. Hypothesis B2 predicted that when faced with resistance, player 1s would update to larger values in the high cost treatment compared to the low cost treatment. Conversely, when faced with no resistance, player 1s should update to less extreme negative values in the high cost case.[8] Following the decision of Player 2 to Resist, Player 1s' reported beliefs of Player 2's private value were on average 0.55 in the low cost treatment and 0.64 in the high cost treatment, a difference in the correct direction but insignificant ($p = .23$). For beliefs following a Not Resist choice, the averages were -1.18 for the low cost treatment and -1.36 for the high cost treatment, which is in the opposite direction predicted by hypothesis 2 but still insignificant ($p = .12$). The lack of significant difference in Player 1 beliefs across treatments suggests that the indirect impact of the Not Fight cost on Player 2's action

---

[7]The PBE prediction for the low cost case is that there should be no updating following Entry. The supplementary materials show how in the AQRE players will update some in a positive direction.

[8]Intuitively, in the low cost case player 2s should believe that some player 1s will ultimately back down if faced with resistance. This induces player 2s with lower private values in the low cost treatment to resist. Conversely, when entry is chosen in the high cost case player 2s expect that player 1 is unlikely to back down (see audience cost discussion above). Hence beliefs following resistance and no resistance should be higher in the high cost case.

choice is not as obvious as its direct impact on Player 1's decision to Enter or Not Enter. This suggests that Player 1s might not properly take into account how the cost of choosing Not Fight impacts both their Entry decisions but also Player 2s subsequent Resistance decisions. In the game that we investigate, Player 1's beliefs about Player 2's private value does not directly impact her decision about Fight or Not Fight as her payoff in these nodes depends only on her private value. While our design allows us to observe belief updating by both players, empirically this updating is most consistent with theory when the beliefs are consequential for choice. Crisis bargaining models in the literature typically assume that updating will happen when it can happen. However, there are related games in which the payoff does depend on the other player's private value (e.g., a bonus to the player with the higher private value only). Whether Player 1's lack of differentiated beliefs across cost treatments about Player 2's private value is due to this lack of strategic necessity is an open question.

Overall, we find some support for comparative static predictions over beliefs derived from our PBE model. Our subjects updated beliefs following entry choices conditional on the incentives they and their opponents faced. This provides some behavioral support for the beliefs conjectured to exist in PBE models. Next we turn to instances where our subjects departed from our equilibrium predictions.

5.3   Testing Rational Expectations

Do the subjects have rational expectations about their opponents? By rational expectations, we mean that the players' expectations about their opponents' private values are correct on average. A first pass is to simply compare average beliefs and average private values following a particular action choice. We present these values and their standard errors in Table 3. Here we see, for example, that beliefs about private values following a no Entry choice were less negative than the average of actual private values of subjects choosing not Enter. A similar pattern is seen for Resist and Not Resist, though not for Entry as we discuss below.

Next we examine conservative updating in more detail by comparing our subjects' beliefs after observing each strategy choice to the *actual* private values of subjects who

10

Choice	High Cost		Low Cost	
	Value	Belief	Value	Belief
Entry	.75(.05)	.77(.09)	.45(.06)	.58(.06)
No Entry	−1.87(.07)	−1.44(.10)	−1.80(.06)	−1.16(.14)
Resist	.86(.06)	.64(.10)	.72(.06)	.55(.05)
No Resist	−1.88(.05)	−1.36(.09)	−1.93(.07)	−1.19(.12)

Table 3: Average private value of opponent vs. average belief (Standard errors clustered at subject level in parentheses).

made the corresponding strategy choice.[9] In most cases beliefs are significantly closer to the 0 prior than the actual private values for both cost treatments. We first compare the stated beliefs after observing no Entry to the private values of their opponents who chose not Enter. We find that the stated beliefs are significantly less negative than the actual private values and thus closer to the zero prior in both treatments (low cost: $p < .001$; high cost $p < .001$). Similarly, after observing no Resistance, Player 1s update to stated beliefs that are less negative (closer to the zero prior) than the private values of the players who chose to not Resist in both the low cost ($p < .001$) and high cost ($p < .001$) treatments. After observing Resistance, Player 1s update to stated beliefs that are less positive (closer to the zero prior) than the private values of Player 2s who chose Resist in both the low cost ($p < .05$) and high cost ($p < .1$) cases. In all three of these instances we observe conservatism in updating.[10] In contrast, after observing Entry in both cost treatments,

---

[9]That is, we take the difference between the belief and private value of the opponent and see whether this difference is significantly different from 0. We cluster standard errors at the subject-level and report p-values corresponding to a two-tail test.

[10]The supplementary materials derive expected beliefs under the Agent Quantal Response Equilibrium (AQRE) model using the maximum likelihood estimate of the precision parameter λ. While these beliefs are closer to the observed beliefs, they still do not account for this finding. Furthermore, AQRE is still a rational expectations model and so it cannot in principle explain the conservatism we observe. It is possible that subject misunderstanding of the distribution of private values generates this behavior. We highlight some reasons why these results are unlikely. First, in the no Resist case, the conservative updating result is largely driven by Player 1s stating a belief of 0 for the expected private value of Player 2. Following a Not Resist choice, 33% and 34% of beliefs were 0 in the high and low cost case and in both cases this was the highest percentage of beliefs. Since it is unlikely that subjects did not understand the symmetry of the distribution or that it was centered at 0 as this was stressed in the instructions, no misunderstanding about the variance of the distribution could have led to these 0 beliefs. Second, subject misunderstanding of the variance of the private value distribution–here that the variance was smaller than it was in reality–could have led to the conservatism in updating that we observe. The average values of Player 1s who chose no Enter are −1.87 and −1.80 in the high and low cost treatments respectively which roughly corresponds to cutpoint at −0.6. Suppose that Player 2s believe the cutoff value to be −0.6 but misunderstood the variance of the logistical distribution. Numerical calculations for this example and others suggest that the perceived variance would need to be substantially smaller in order to generate the observed average beliefs.

Player 2s actually update to stated beliefs that are higher than (further away from the zero prior) the average private value of the Player 1s who chose to Enter. However, the difference is small in magnitude and only marginally significant in the low cost case (low cost: $p = .07$; high cost: $p = .79$).[11]

Conservatism in beliefs (Edwards, 1968) has been documented for probabilistic stated beliefs about action choices made in the experimental population (Huck and Weizsäcker, 2002). Here we provide the first systematic account of conservatism in beliefs about opponent type after observing the opponent's action in a sequential game of two-sided incomplete information (similarly, see Dominitz and Hung, 2009). A number of recent models such as cursed equilibrium (Eyster and Rabin, 2005) and behavioral equilibrium (Esponda, 2008) are built upon the observation that agents do not properly account for the connection between other agents' private information and action choice in bilateral trade with adverse selection, auctions, voting games, etc.. The conservatism in belief updating we find offers evidence of this disconnect. More specifically, our subjects underestimate what some actions reveal about private values more than other actions. After their opponent chooses to Not Enter or Not Resist, they fail to expect the extent to which such a choice indicates a low private value/strength. Such "inertia" in updating can have considerable implications for optimal strategies in a broad class of games with private information. For instance, a number of experiments on signaling games have documented behavior that is consistent with sluggish updating about the opponent's type within or across trials (Cooper *et al.*, 1997; Potters and van Winden, 1996; Kübler *et al.*, 2008). Kübler et al.'s study of a Spence education game find that the wage spread is not as large as predicted by the equilibrium, suggesting that the wage-setters might not be updating as much as they should about the workers' types based on their signals. They suggest that "for future research it could be useful to elicit the beliefs of players in order to pin down the equilibrium more precisely" (pg. 234). Indeed, our results about conservative updating in a related setting offer suggestive evidence that the wage setting could have

---

[11]We suggest one possible account of the discrepancy between the lack of conservative updating by Player 2s following Entry and the strong conservative updating results after no Entry. There are a number of Player 1s with negative private values who choose to Enter, thus pulling the average private value down following entry while also leaving subjects with more negative values not entering.

been driven by conservative updating about the workers' types.[12] More work can be done in the future to elicit beliefs in various games of incomplete information to examine if belief updating patterns within or across trials can account for particular deviations from equilibrium behavior.[13]

6 Conclusion

Our study is one of the first to examine beliefs and strategies in a sequential game with two-sided incomplete information, specifically the "crisis bargaining" game, in the laboratory. This class of games has relevant applications in industrial organization, international relations, and legal studies, among other fields. We directly elicit the expectation each player holds about her opponent's type at various decision nodes rather than her expectation about her opponent's strategy choice, as in most previous belief elicitation studies. We find some evidence that players update their beliefs in directions consistent with the perfect Bayesian equilibrium. However, players update conservatively compared to the average private values of subjects for most strategy choices. That is, the beliefs remain "sticky" around the prior and violate rational expectations.

References

Banks, J. C., Camerer, C. F., and Porter, D. (1994). An experimental analysis of nash refinements in signaling games. *Games and Economic Behavior*, **6**(1), 1–31.

Blanco, M., Engelmann, D., Koch, A., and Norman, H.-T. (2011). Belief elicitation in experiments: Is there a hedging problem? *Experimental Economics*.

Brandts, J. and Holt, C. A. (1992). An experimental test of equilibrium dominance in signaling games. *American Economic Review*, **82**(5), 1350–1365.

Cai, H. and Wang, J. T. (2006). Overcommunication in strategic information transmission games. *Games and Economic Behavior*, **56**(1), 7–36.

Camerer, C. F. and Weigelt, K. (1988). Experimental tests of a sequential equilibrium model. *Econometrica*, **56**(1), 1–36.

---

[12]Whether this underestimation can be exploited or serve as a deterrent is the topic of future work. It is also possible that having subjects work together could reduce the amount of conservatism (Cooper and Kagel, 2005).

[13]The supplementary materials show that players' action choices are mostly in line with the PBE predictions but there are deviations such as some no-Entry in the low cost treatment and not Fight choices in the high cost treatment.

Carrillo, J. D. and Palfrey, T. R. (2009). The compromise game: Two-sided adverse selection in the laboratory. *American Economic Journal: Microeconomics*, **1**(1), 151–181.

Charness, G. and Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, **74**(6), 1579–1601.

Cooper, D., Garvin, S., and Kagel, J. (1997). Signalling and adaptive learning in an entry limit pricing game. *Rand Journal of Economics*, **28**(4), 662–683.

Cooper, D. J. and Kagel, J. H. (2005). Are two heads better than one? team versus individual play in signaling games. *American Economic Review*, **95**(3), 477–509.

Costa-Gomes, M. A. and Weiszäcker, G. (2008). Stated beliefs and play in normal form games. *Review of Economic Studies*, **75**(3), 729–762.

Croson, R. (2000). Thinking like a game theorist: Factors affecting the frequency of equilibrium play. *Journal of Economic Behavior and Organization*, **41**(3), 299–314.

Dominitz, J. and Hung, A. A. (2009). Empirical models of discrete choice and belief updating in observational learning experiments. *Journal of Economic Behavior and Organization*, **69**(2), 94 – 109.

Drouvelis, M., Müller, W., and Possajennikov, A. (2009). Signaling without common prior: An experiment. *Working Paper*.

Dufwenberg, M. and Gneezy, U. (2000). Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior*, **30**(2), 163–182.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz, editor, *Formal Representation of Human Judgment*, pages 17–52. John Wiley and Sons, New York.

Esarey, J., Mukherjee, B., and Moore, W. H. (2008). Strategic interaction and interstate crises: A bayesian quantal response estimator for incomplete information games. *Political Analysis*, **16**(3), 250–273.

Esponda, I. (2008). Behavioral equilibrium in economies with adverse selection. *American Economic Review*, **98**(4), 1269–1291.

Eyster, E. and Rabin, M. (2005). Cursed equilibrium. *Econometrica*, **73**(5), 1623–1672.

Fearon, J. D. (1994). Domestic political audiences and the escalation of international disputes. *American Political Science Review*, **88**(3), 577–592.

Fudenberg, D. and Tirole, J. (1991). *Game Theory*. MIT Press, Cambridge, Massachusetts.

Huck, S. and Weizsäcker, G. (2002). Do players correctly estimate what others do? evidence of conservatism in beliefs. *Journal of Economic Behavior and Organization*, **47**(1), 71–85.

Jung, Y. J., Kagel, J. H., and Levin, D. (1994). On the existence of predatory pricing: an experimental study of reputation and entry deterrence in the chain-store game. *Rand Journal of Economics*, **25**(1), 72–93.

Kreps, D. and Wilson, R. (1982). Reputation and imperfect information. *Journal of Economic Theory*, **27**(2), 253–279.

Kübler, D., Müller, W., and Normann, H.-T. (2008). Job-market signaling and screening: An experimental comparison. *Games and Economic Behavior*, **64**, 219–236.

Lewis, J. B. and Schultz, K. A. (2003). Revealing preferences: Empirical estimation of a crisis bargaining game with incomplete information. *Political Analysis*, **11**(4), 345–367.

Manski, C. F. (2004). Measuring expectations. *Econometrica*, **72**(5), 1329–1376.

McKelvey, R. D. and Page, T. (1990). Public and private information: An experimental study of information pooling. *Econometrica*, **58**(6), 1321–1339.

McKelvey, R. D. and Page, T. (2000). An experimental study of the effect of private information in the coase theorem. *Experimental Economics*, **3**(3), 187–213.

Milgrom, P. and Roberts, J. (1982). Predation, reputation, and entry deterrence. *Journal of Economic Theory*, **27**(2), 280–312.

Nyarko, Y. and Schotter, A. (2002). An experimental study of belief learning using elicited beliefs. *Econometrica*, **70**(3), 971–1005.

Offerman, T., Sonnemans, J., and Schram, A. (1996). Value orientations, expectations, and voluntary contributions in public goods. *The Economic Journal*, **106**(437), 817–845.

Palfrey, T. R. and Wang, S. W. (2009). On eliciting beliefs in strategic games. *Journal of Economic Behavior and Organization*, **71**(2), 98–109.

Potters, J. and van Winden, F. (1996). Comparative statics of a signaling games: An experimental study. *International Journal of Game Theory*, **25**, 329–353.

Savage, L. J. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, **66**(336), 783–801.

Schultz, K. A. (2001). Looking for audience costs. *Journal of Conflict Resolution*, **45**(1), 32–60.

Tingley, D. and Walter, B. (2011). The effect of repeated play on reputation building:an experimental approach. *International Organization*.

Tomz, M. (2007). Domestic audience costs in international relations: An experimental approach. *International Organization*, **61**(4), 821–840.

Wilson, R. (1992). Strategic models of entry deterrence. In R. Aumann and S. Hart, editors, *Handbook of Game Theory*, pages 281–304. Elsevier, New York.