Interpretation Rules for Incomplete Contracts: A Laboratory Experiment*

Sean Inoue[†]

June 14, 2019

Abstract

This paper provides an experimental test for incomplete contracting theory where interpretation is crucial. The paper tests simplified versions of models detailed by Heller & Spiegler (2008) and Shavell (2006). In the experiment, one player takes the role of a Writer of a contract, while the other player takes the role of a Decider who decides on a rule of interpretation to be used. Interpretation is used in the cases where the contract does not specify an action for a state of the world. The experiment uses a 2x2 experimental design, where the order of play is changed in one dimension and there is an increasing conflict of interest between the two players in the other dimension. As the conflict of interest grows, contracts should become more obligationally complete. The experimental results support that prediction. With few exceptions, play is found to be in accordance with subgame perfect equilibrium.

JEL: C90, D23, D86, K12
Keywords: *incomplete contracting, interpretation in contracts, experimental economics*

*Thanks to Charles Noussair and ESL at the University of Arizona for funding this experiment and for providing a venue in which to run this experiment, to Martin Dufwenberg for feedback involving my instructions and for allowing me to run a pilot in his class, to Mark Walker for his detailed feedback, and to Andreas Blume for his repeated feedback and motivation to work on this project.

[†]University of Arizona. Email: sinoue@email.arizona.edu

1 Introduction

Interpretation plays a central role in how contracts are written. In construction, for example, it is too costly to write a fully detailed contract. These contracts rely on interpretation when the present state is not mentioned in the contract. This interpretation is carried out either by the person whose responsibility it is to take an action or by a court. Although the kind and degree of incompleteness in a contract may be identifiable (Mansor, Rashid [2014]), it is impossible to map data on incomplete contracts directly to theory. The goal of this paper is to examine whether the way contracts are interpreted and written matches existing theory models in a laboratory experiment.

This experiment will examine a setting similar to Shavell (2006) and Heller & Spiegler (2008). These two papers focus on how people write incomplete contracts and how interpreting parties optimally interpret contracts. There are two agents: a Writer and a Decider. The Writer writes a contract that may be incomplete, because writing a contract that specifies some action for every possible state is too costly due to the plethora of possible states. The Decider provides an interpretation in the case that the state falls into a gap. Following the Writer's contract and the Decider's interpretation, the state is drawn. An action for this state is either written into the contract or is provided by the interpretation of the Decider.

In Shavell and in Heller & Spiegler (H&S), the Writer writes a contract that specifies a number of terms and is either restricted by a complexity bound (H&S) or is subject to a cost per term (Shavell). Each term in the contract dictates a specific action to be taken when one of the states listed in the term is realized. There are different orders of play in the two papers. In Shavell, the Decider publishes the interpretive rule first, while in H&S, the Writer writes a contract first. Both cases are relevant. In some instances, an interpretive rule may be published by a court, in which case the rules of interpretation are common knowledge when someone goes to write a contract with a gap in it. In other instances, when a contract written has no standard interpretation, an interpretive clause may not be defined, in which case the party interpreting the contract must decide on an interpretation after the contract is written. In addition, increasing conflict between the party interpreting a contract and the party writing the contract is predicted to affect how complete a contract will be. Both papers illustrate a key problem: When the Writer writes a contract, including an extra state gives the Writer more control over the action taken and gives the Decider less control over the action taken. This problem of how to optimally assign control is the main focus of this paper.

This experiment will test four primary predictions. A main result of H&S is that contracts become more obligationally complete as the conflict of interest grows. Another main result is that each contract should include either the lowest state or the highest state. In addition to testing these predictions of equilibria, this experiment will test to see if the actions in contracts are optimal given the states included in the contract. This experiment will also test for whether Decider actions are subgame perfect. The experiment includes four treatments in total, using a 2x2 design. Treatments vary over the levels of conflict of interest and over the order of play. The conflict treatment has a low level of conflict and a high level of conflict. The variation in the order of play is so that both Shavell and H&S are represented.

Over the last ten rounds of play, the contracts are more complete in the high-conflict of interest treatment, providing support for the predictions. Participants also tend to place contracts such that one of the extremes of the state space is covered in the contract. In addition, there is evidence that subjects write actions into their contract that are optimal given the states included. Writers who move first write contracts that more consistent with equilibrium predictions than a writer who plays randomly. Deciders mirror this prediction when the conflict is high, but when the conflict is low Deciders overwhelmingly pick the action in the middle of the state space.

This observation, along with the post-experimental questionnaire, indicate that fairness and reciprocity could play a role in the distribution of Default Actions and Writer Actions. In employment contracts, fairness and reciprocity are sometimes implied by the terms of the contract. This could occur both on the employer's side, when employers are expected to give employees bonuses for good work, and on the employee's side, when employees are expected to work overtime.

Exploring this more, a level-k model is analyzed in this setting. When a level-k model in which level 0s randomize uniformly across all actions that are not dominated is tested, a level-k test fails, as too many subjects are classified as level-0. This could be due to the complexity of the game or to a bad starting point for level-0.

Dye (1985) details a model in which contracts are costly to write. Because of the cost, firms can decide to write a contract that does not cover all states of the world. This idea that contractual incompleteness can be caused by a writing cost is echoed in Shavell (2006) and Heller & Spiegler (2008), who focus on intentionally incomplete contracts in settings with interpretation rules. There are many important papers discussing incomplete contracts following Dye, including Simon (1991), who discusses incompleteness in labor contracts.

There is also a modest experimental literature on incomplete contracting. However, the literature tends to focus on the hold-up problem and on how people invest, whereas this experiment focuses on how contracts are written in settings where people provide interpretive rules for contracts. Hackett (1993) is a primary example of this. Fehr, Powell, & Wilkening (2014) aim to test the performance of Maskin & Tirole's subgame-perfect implementation mechanism in a setting with observable but non-verifiable effort, and they find that this mechanism fails to achieve good outcomes in an experimental setting.

Within the hold-up literature, many papers analyze various behavioral influences. Dufwenberg, Smith, & Van Essen (2013) analyze negative reciprocity in conjunction with the hold-up problem. Ellingsen & Johannesson (2004) look at promises and threats in this context.

In addition to this literature on the hold-up problem, there are many recent papers that analyze the best way for contracts to be written in different settings. Fehr et al. (2007) has a paper that looks at fairness in contracting, analyzing bonus contracts. Brands, Charness, & Ellman (2012) analyze how communication plays a role in rigid and flexible contracting.

Cai & Wang (2006) analyze communication games in an experimental setting using a level-k model. Unlike this paper, they find that level-k supports subjects' play in a communication game.

Section two of this paper will outline the model and the predictions. Section three will discuss the experimental design and the results. Section four will conclude the paper.

2 Theoretical Predictions

2.1 Description

There are two players: A Writer and a Decider. The Writer has payoff $\pi_W = 20- \mid 2s - a \mid$, and the Decider has payoff $\pi_D = 20- \mid 2s - a + b \mid$, where $s \in S = \{1, 2, 3, 4, 5, 6\}$ is a state of the world drawn from a uniform distribution over the state space S, $a \in \mathbb{N}$ is the action taken, and $b > 0$ is the bias term.

To determine the action that is taken, the Writer writes a contract (or a rule) that is a triplet $c = (s_{low}, s_{high}, a_W)$, where $s_{low} \in S$, $s_{high} \in S$, $a_W \in \mathbb{N}$, and $s_{low} \leq s_{high}$. The 'event' is denoted $e = [s_{low},\ s_{high}]$. a_W, will sometimes be referred to as the 'rule action,' since this is the action named in the rule that the Writer writes. In the case that the realized state $s \in e$, $a = a_W$. In simple language, a contract has one term in it that states: If the state falls between s_{low} and s_{high}, take action a_W.[1]

A few terms that will be used throughout the paper:

[1]This experiment will allow only one term in the contract, and all of the results and predictions are based on this. However, Shavell and Heller & Spiegler allow for a much richer space of contracts.

A contract in this setting will be considered **obligationally complete** if $s_{low} = 1$ and $s_{high} = 6$, and will be considered incomplete otherwise.

A contract is said to have a **gap** if the contract is incomplete. The **gap** is the set of states not covered by the contract $\{s \notin e\} = \{s < s_{low} \cup \ s > s_{high}\}$.

A **lacuna** in a contract is either $\{s < s_{low}\}$ or $\{s > s_{high}\}$. The union of all **lacunae** is the gap. In this setting, a contract contains at most 2 **lacunae**.

For example, suppose that the the Writer writes the contract $c = (s_{low}, s_{high}, a_W) = (2, 4, 7)$. The gap in the contract is the set $\{1, 5, 6\}$. There are two lacunae, which are the sets $\{1\}$ and $\{5, 6\}$.

The Decider publishes a default action a_D. In the case that $s < s_{low}$ or $s > s_{high}$, $a = a_D$. That is to say, when the state falls into a gap, the default action takes place. When the state is between the low state and the high state, the action specified in the contract will be taken instead.

Continuing the example above, suppose that the Decider decides $a_D = 12$. If $s \in \{2, 3, 4\}$, then $a = a_W = 7$. If $s \in \{1, 5, 6\}$, then $a = a_D = 12$.

The experiment utilizes two versions of this model to reflect the difference between Heller & Spiegler (2006) and Shavell (2008):

(Wfirst) The Writer moves first, publishing a contract $c = (s_{low}, s_{high}, a_W)$. The Decider observes c and then chooses a_D. The state is drawn after both players have made their decisions. (H&S)

(Dfirst) The Decider moves first, publishing a default action a_D. The Writer observes a_D and then writes a contract $c = (s_{low}, s_{high}, a_W)$. The state is drawn after both players have made their decisions. (Shavell)

The remainder of the paper will use Wfirst and Dfirst when referring to the two different versions of the game.

Here, a small discussion is warranted about the differences between the models in this paper, in H&S, and in Shavell. The model presented in Shavell (2006) is a generalized version of Dfirst: payoffs for each player are not specified, and instead of just interpreting a gap, the Decider is allowed to interpret terms in a contract as well, which is not allowed here. Wfirst is a discretized version of H&S, with a couple of other minor differences. Here, the payoffs are linear in the difference between the action and the state, whereas in H&S, the payoffs are quadratic loss functions. In addition, H&S allow for a more complex set of rules, where either multiple intervals can map to an action or where the allowed contract can have multiple terms. In addition, this experiment multiplies the state by two in the payoff function to get rid of the need for non-integer actions.

The next section will analyze the models of Wfirst and Dfirst using a subgame-perfect equilibrium. It will cite the result from the appropriate paper and, if necessary, will prove the analog of the result.

2.2 Observations

Observation 1 *The optimal action that the Writer chooses is such that* $2\mathbf{E}\left[s \mid s \in e\right] - a_W = 0$. *If* $2\mathbf{E}\left[s \mid s \in e\right]$ *is odd,* $a_W = 2E\left[s \mid s \in e\right] \pm 1$ *is also optimal. then In Wfirst, The optimal default action the Decider chooses is such that* $2E[(s \mid s \notin e)] - a_D = -b$. *If* $2E[(s \mid s \notin e)]$ *is odd,* $a_D = 2E[(s \mid s \notin e)] + b \pm 1$ *is also optimal.*

This observation trivially arises from how the payoffs are setup, and this experiment uses it to test whether Writers are expected payoff maximizers. Multiple actions being optimal is a result of a lack of risk aversion in the model.

Observation 2 *All SPE are solved for by doing backward induction: For any decision that the first mover could do, the optimal action for the second mover is found. Given any particular combination of actions that yields a Nash equilibrium within each proper subgame, the first mover picks an action that maximizes their payoff. All possible SPE paths are detailed in Table 1.*

This observation is obtained through writing the game tree in detail for each treatment. There are a few noteworthy things about the set of subgame perfect equilibria. In Wfirst, as b increases from 2 to 10 the number of states included in a contract in an SPE weakly increases. The degree of incompleteness in Wfirst is specified completely in H&S under quadratic payoffs, and in their setting it is generally true that increasing b increases the size of the optimal contract. Note that in Dfirst there is not a (weak) increase in the size of contracts due to the presence of equilibria in which the Decider dedicates to a low action that favors the Writer, causing the Writer to best respond by writing a contract that covers few states.

It is also noteworthy that in each almost every SPE a gap is also a lacuna. With a small degree of risk aversion, as is the case in H&S, this is easy to show. In this setting, in Wfirst b=10 a contract with $s_{low} = 2$ and $s_{high} = 5$ can be supported as a SPE due to the fact that all default actions yield the same payoff.

$(s_{low},\ s_{high},\ a_W), a_D$	Dfirst	Wfirst
$b=2$	$(5,6,12), 6$	$(1,2,2), 10$
	$(5,6,12), 7$	$(1,2,3), 10$
	$(1,3,4), 11$	$(1,2,4), 10$
	$(1,3,4), 12$	$(5,6,10), 6$
	$(4,6,10), 5$	$(5,6,11), 6$
	$(4,6,10), 6$	$(5,6,12), 6$
	$(1,4,6), 12$	$(1,3,4), 12$
	$(1,4,6), 13$	$(4,6,10), 6$
	$(1,4,6), 14$	$(1,4,4), 12$
		$(1,4,5), 12$
		$(1,4,6), 12$
		$(3,6,8), 4$
		$(3,6,9), 4$
		$(3,6,10), 4$
$b=10$	$(5,6,10), 7$	$(2,5,6), 12$
	$(5,6,11), 7$	$(2,5,7), 12$
	$(5,6,12), 7$	$(2,5,8), 12$
	$(1,3,4), 12$	$(1,6,6)$, Anything
	$(1,4,5), 14$	$(1,6,7)$, Anything
	$(1,4,6), 12$	$(1,6,8)$, Anything
	$(1,4,6), 13$	
	$(1,4,6), 14$	
	$(1,5,6), 17$	
	$(1,5,6), 18$	
	$(1,6,8), 18$	
	$(1,6,8), 19$	
	$(1,6,8), 20$	
	$(1,6,8), 21$	
	$(1,6,8), 22$	

Table 1: SPE Paths

2.3 Predictions

1 Each gap, if it exists, will be a lacuna. That is, $s_{low} = 1$ or $s_{high} = 6$.

A majority of equilibria have the feature that a gap is a lacuna. This has an intuition that both Writers and Deciders receive bad payoffs when the default action covers two different lacunae, and both receive good payoffs when the default action covers a single lacuna. The exception to this intuition in Observation 2 came from a corner case where the written contract is such that every default action yields the same payoff.

2 The action the Writer writes in the contract maximizes the Writer's payoff given the contract structure.

The contract that the Writer writes is such that $a_w = s_{low} + s_{high}$. If $s_{low} + s_{high}$ is odd, then it is also optimal that $a_w = s_{low} + s_{high} \pm 1$.

3 As b increases from 2 to 10, the number of states covered in a contract increases under both Wfirst and Dfirst.

For Wfirst, this prediction comes straight from Observation 2. For Dfirst, this prediction is supported by the intuition that as the conflict of interest grows between the Writer and the Decider, the Writer wants to control more of the state space. Conditional on the Decider picking a specific default action, the Writer wants to include the same number of states. However, if the Decider picks higher default actions, which they may with a higher bias, the Writer will respond by writing contracts that give the Decider less control.

4 The first mover in each treatment takes an action consistent with the one of the non-dominated SPE classified in table 1.

Prediction 4 relies on the first mover being able to do some degree of backward induction. Prior literature shows that subjects fail at higher degrees of backwards induction (Binmore et al. [2002]). However, as there are many equilibria, this prediction tests whether first movers select into specific equilibria if they are playing in equilibrium.

3 Experimental Design

For this experiment, Subjects played the game in Z-Tree (Fischbacher [2007]) at the Experimental Science Laboratory at the University of Arizona. There were 8 total sessions (2 per treatment), with 8-12 subjects in each session, all of whom were University of Arizona students. Each subject played 2 practice rounds of the game with themselves, and then were quizzed about the results of the practice rounds. Following this, each person was randomly assigned to be either a Writer or a Decider and then played 15 rounds of the game, being anonymously and randomly matched with a partner in each round. Subjects were paid for 2 of the 15 rounds of play, chosen randomly by them, and in addition were paid a $6 show-up fee. The only change between the theory covered above and the experiment is that the experiment restricted the action space to be integers that are undominated for both players. So in the $b = 2$ treatments, the actions possible were the integers between 2 and 14, while in the $b = 10$ treatments, the actions possible were the integers between 2 and 22.

Before the experiment began, the instructions were read aloud and the subjects were allowed to look over the instructions for five minutes. At the end of each practice round, there was a quiz on what action was taken based on the state drawn and what payoffs each player would have received. The subjects were paid in Experimental Currency Units, and received 30 cents for each ECU they earned.

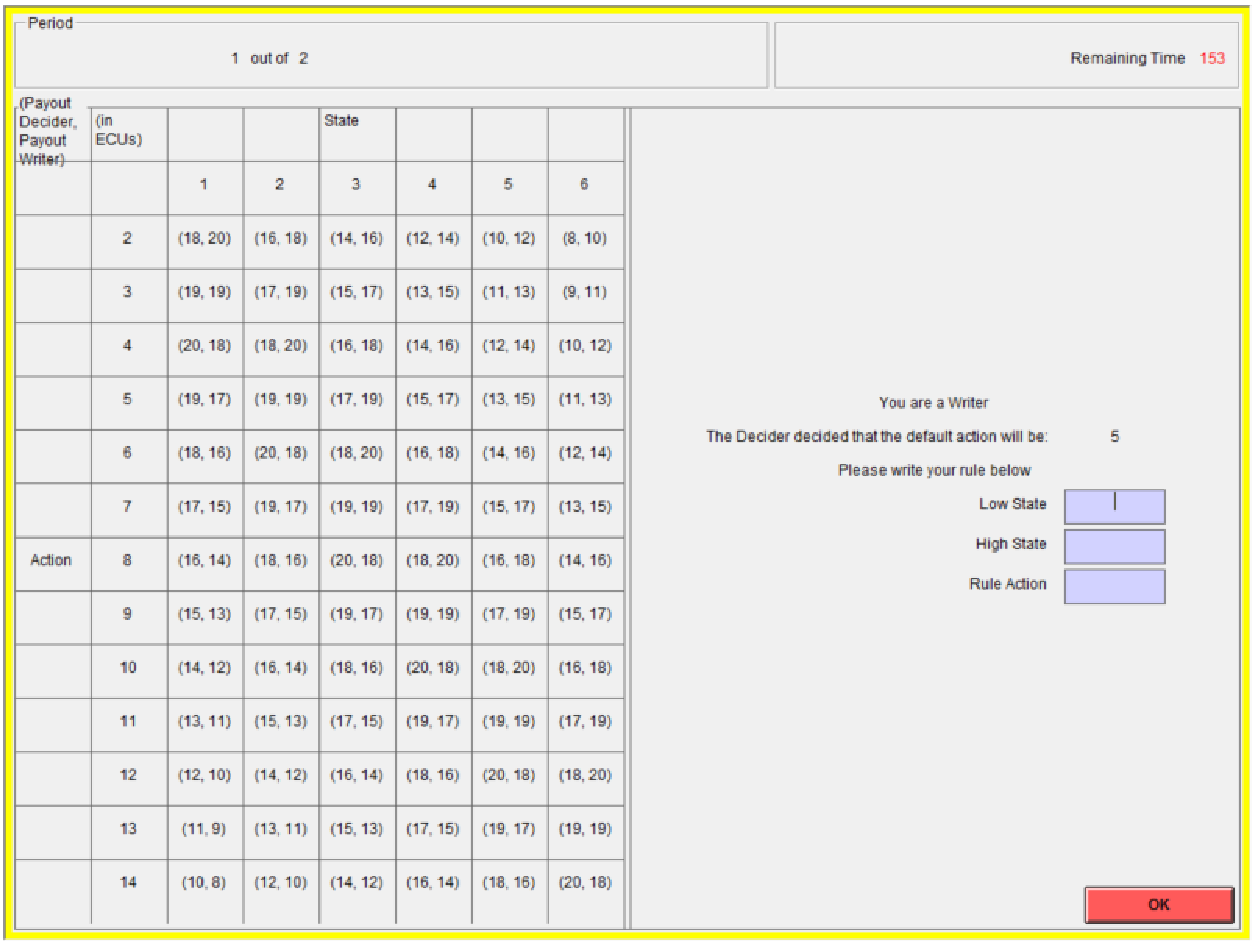
4 Experimental Results

4.1 How Contracts are Written

In general, subjects wrote contracts that were on one edge of the state space according to Table 2. There appears to have been a move toward writing a contract on the edge of the state space as well (Figure 4), as in the last five rounds, 86% of people were writing contracts that contained state 1 or state 6. Across treatments there is little variation in this. Although this is still significantly different from 100% of people covering states 1 or 6 in their contract, a person who played randomly would have a .524 probability of contracting on an edge state (Since there are 21 possible s_{low}, s_{high} combinations and 11 of those contain 1 or 6). Using a t-test, all play is significantly different from pure randomization except for in the Dfirst b=10 treatment for periods 6-10. This demonstrates that people tend to write contracts that include the bounds of the state space.

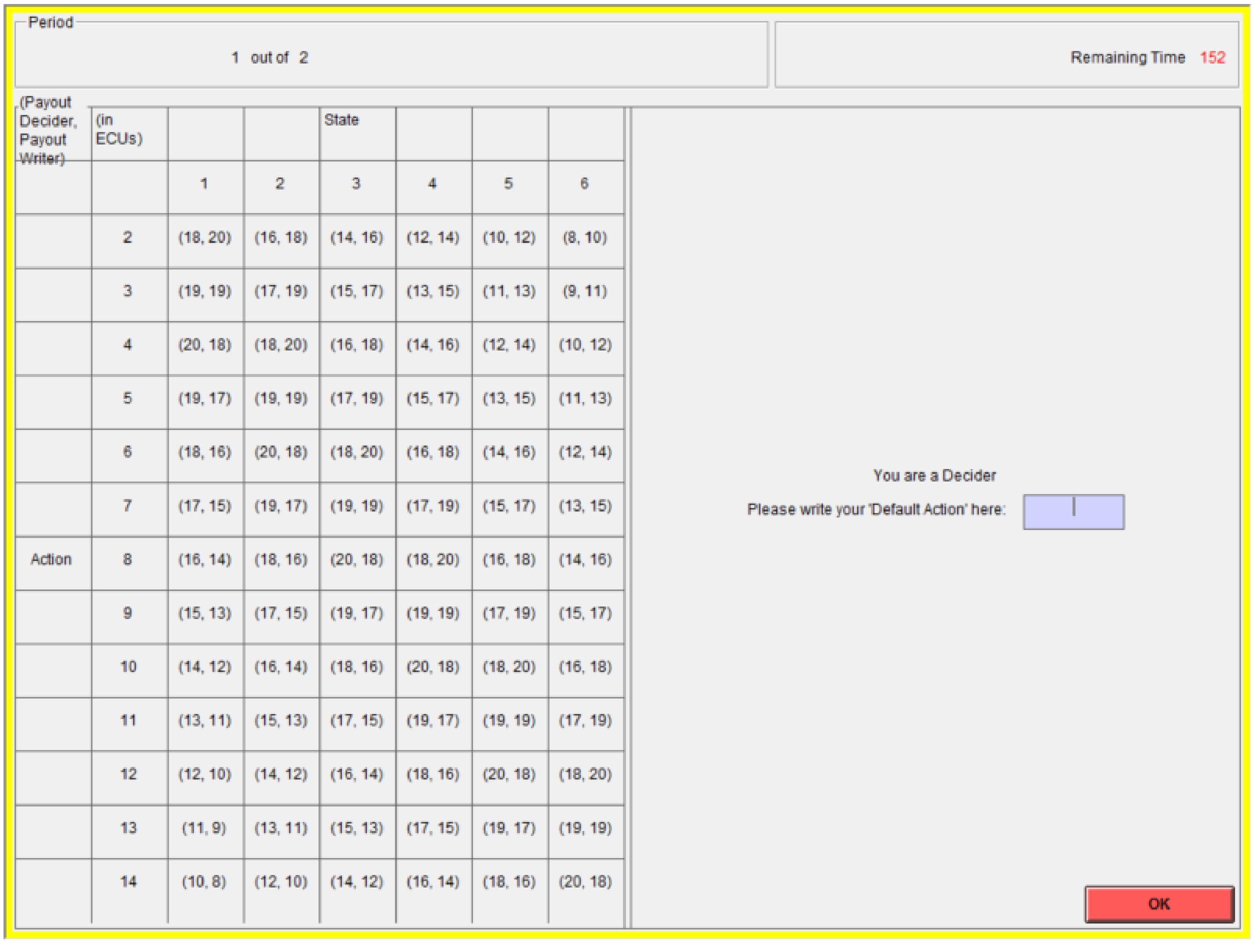
Result 1: Prediction 1 is supported. A significant majority of the contracts written include $s_{low} = 1$ or $s_{high} = 6$. The number of contracts that include $s_{low} = 1$ or $s_{high} = 6$ is significantly more than random play.

For whether people write their actions optimally (Figure 5 and Table 3)), approximately



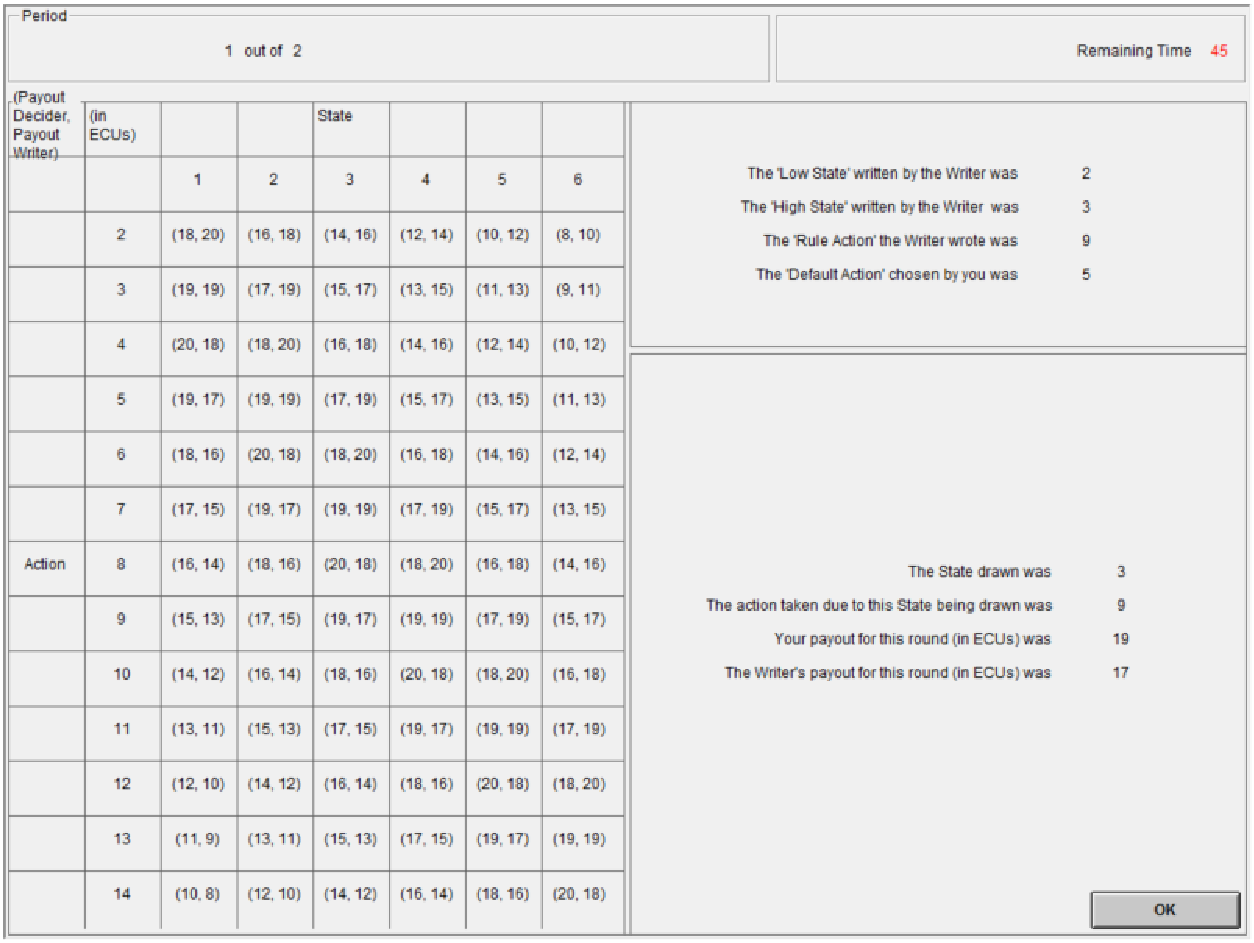
(Payout Decider, Payout Writer)	(in ECUs)	State					
		1	2	3	4	5	6
	2	(18, 20)	(16, 18)	(14, 16)	(12, 14)	(10, 12)	(8, 10)
	3	(19, 19)	(17, 19)	(15, 17)	(13, 15)	(11, 13)	(9, 11)
	4	(20, 18)	(18, 20)	(16, 18)	(14, 16)	(12, 14)	(10, 12)
	5	(19, 17)	(19, 19)	(17, 19)	(15, 17)	(13, 15)	(11, 13)
	6	(18, 16)	(20, 18)	(18, 20)	(16, 18)	(14, 16)	(12, 14)
	7	(17, 15)	(19, 17)	(19, 19)	(17, 19)	(15, 17)	(13, 15)
Action	8	(16, 14)	(18, 16)	(20, 18)	(18, 20)	(16, 18)	(14, 16)
	9	(15, 13)	(17, 15)	(19, 17)	(19, 19)	(17, 19)	(15, 17)
	10	(14, 12)	(16, 14)	(18, 16)	(20, 18)	(18, 20)	(16, 18)
	11	(13, 11)	(15, 13)	(17, 15)	(19, 17)	(19, 19)	(17, 19)
	12	(12, 10)	(14, 12)	(16, 14)	(18, 16)	(20, 18)	(18, 20)
	13	(11, 9)	(13, 11)	(15, 13)	(17, 15)	(19, 17)	(19, 19)
	14	(10, 8)	(12, 10)	(14, 12)	(16, 14)	(18, 16)	(20, 18)

Figure 1: The Writer Screen in Dfirst, b=2



(Payout Decider, Payout Writer)	(in ECUs)			State			
		1	2	3	4	5	6
	2	(18, 20)	(16, 18)	(14, 16)	(12, 14)	(10, 12)	(8, 10)
	3	(19, 19)	(17, 19)	(15, 17)	(13, 15)	(11, 13)	(9, 11)
	4	(20, 18)	(18, 20)	(16, 18)	(14, 16)	(12, 14)	(10, 12)
	5	(19, 17)	(19, 19)	(17, 19)	(15, 17)	(13, 15)	(11, 13)
	6	(18, 16)	(20, 18)	(18, 20)	(16, 18)	(14, 16)	(12, 14)
	7	(17, 15)	(19, 17)	(19, 19)	(17, 19)	(15, 17)	(13, 15)
Action	8	(16, 14)	(18, 16)	(20, 18)	(18, 20)	(16, 18)	(14, 16)
	9	(15, 13)	(17, 15)	(19, 17)	(19, 19)	(17, 19)	(15, 17)
	10	(14, 12)	(16, 14)	(18, 16)	(20, 18)	(18, 20)	(16, 18)
	11	(13, 11)	(15, 13)	(17, 15)	(19, 17)	(19, 19)	(17, 19)
	12	(12, 10)	(14, 12)	(16, 14)	(18, 16)	(20, 18)	(18, 20)
	13	(11, 9)	(13, 11)	(15, 13)	(17, 15)	(19, 17)	(19, 19)
	14	(10, 8)	(12, 10)	(14, 12)	(16, 14)	(18, 16)	(20, 18)

Figure 2: The Decider Screen in Dfirst, b=2



Period

1 out of 2

Remaining Time 45

(Payout Decider, Payout Writer)	(in ECUs)	State					
		1	2	3	4	5	6
	2	(18, 20)	(16, 18)	(14, 16)	(12, 14)	(10, 12)	(8, 10)
	3	(19, 19)	(17, 19)	(15, 17)	(13, 15)	(11, 13)	(9, 11)
	4	(20, 18)	(18, 20)	(16, 18)	(14, 16)	(12, 14)	(10, 12)
	5	(19, 17)	(19, 19)	(17, 19)	(15, 17)	(13, 15)	(11, 13)
	6	(18, 16)	(20, 18)	(18, 20)	(16, 18)	(14, 16)	(12, 14)
	7	(17, 15)	(19, 17)	(19, 19)	(17, 19)	(15, 17)	(13, 15)
Action	8	(16, 14)	(18, 16)	(20, 18)	(18, 20)	(16, 18)	(14, 16)
	9	(15, 13)	(17, 15)	(19, 17)	(19, 19)	(17, 19)	(15, 17)
	10	(14, 12)	(16, 14)	(18, 16)	(20, 18)	(18, 20)	(16, 18)
	11	(13, 11)	(15, 13)	(17, 15)	(19, 17)	(19, 19)	(17, 19)
	12	(12, 10)	(14, 12)	(16, 14)	(18, 16)	(20, 18)	(18, 20)
	13	(11, 9)	(13, 11)	(15, 13)	(17, 15)	(19, 17)	(19, 19)
	14	(10, 8)	(12, 10)	(14, 12)	(16, 14)	(18, 16)	(20, 18)

The 'Low State' written by the Writer was 2
The 'High State' written by the Writer was 3
The 'Rule Action' the Writer wrote was 9
The 'Default Action' chosen by you was 5

The State drawn was 3
The action taken due to this State being drawn was 9
Your payout for this round (in ECUs) was 19
The Writer's payout for this round (in ECUs) was 17

OK

Figure 3: The Payout Screen for the Decider in Dfirst, b=2

50% of rounds had contracts that included an optimal action given the states. Interestingly, there were a higher number of contracts written with an optimal action in the b=2 Wfirst treatment, which may have to do with the fact that in Wfirst, Writers could write their contracts optimally without having to worry about what the Deciders have done. In Wfirst b=10, in a fair number of periods, people wrote the contract $s_{low} = 1$, $s_{high} = 6$, $a_W = 10$, which is suboptimal, but in a way that helps the Decider in all states. In addition, there is a slight upward trend, as can be seen in figure 5. Statistically all treatments are different from pure randomization, with most specifications being different at the 1% level.

Result 2: Prediction 2 is supported. Subjects write rule actions that do statistically better than a subject who purely randomizes.

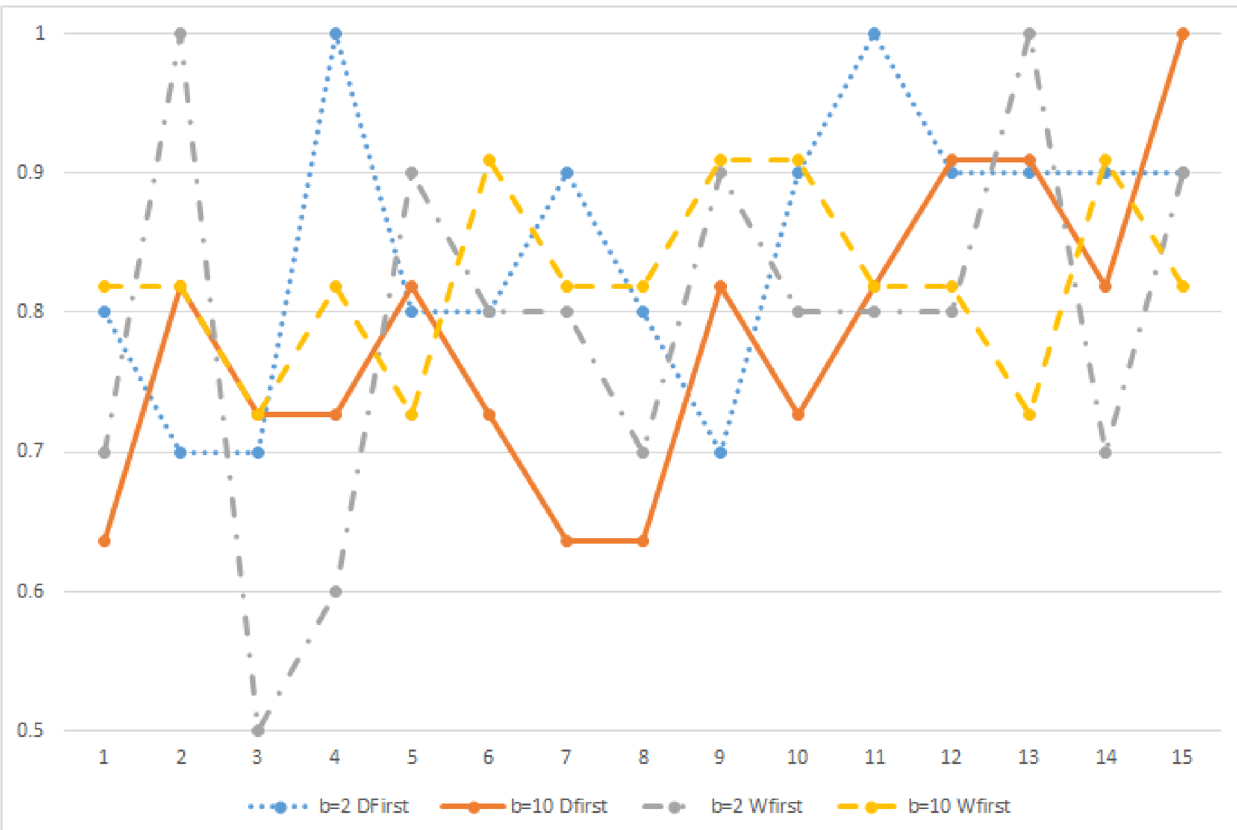


Figure 4: Rules That Cover States 1 or 6 by Period

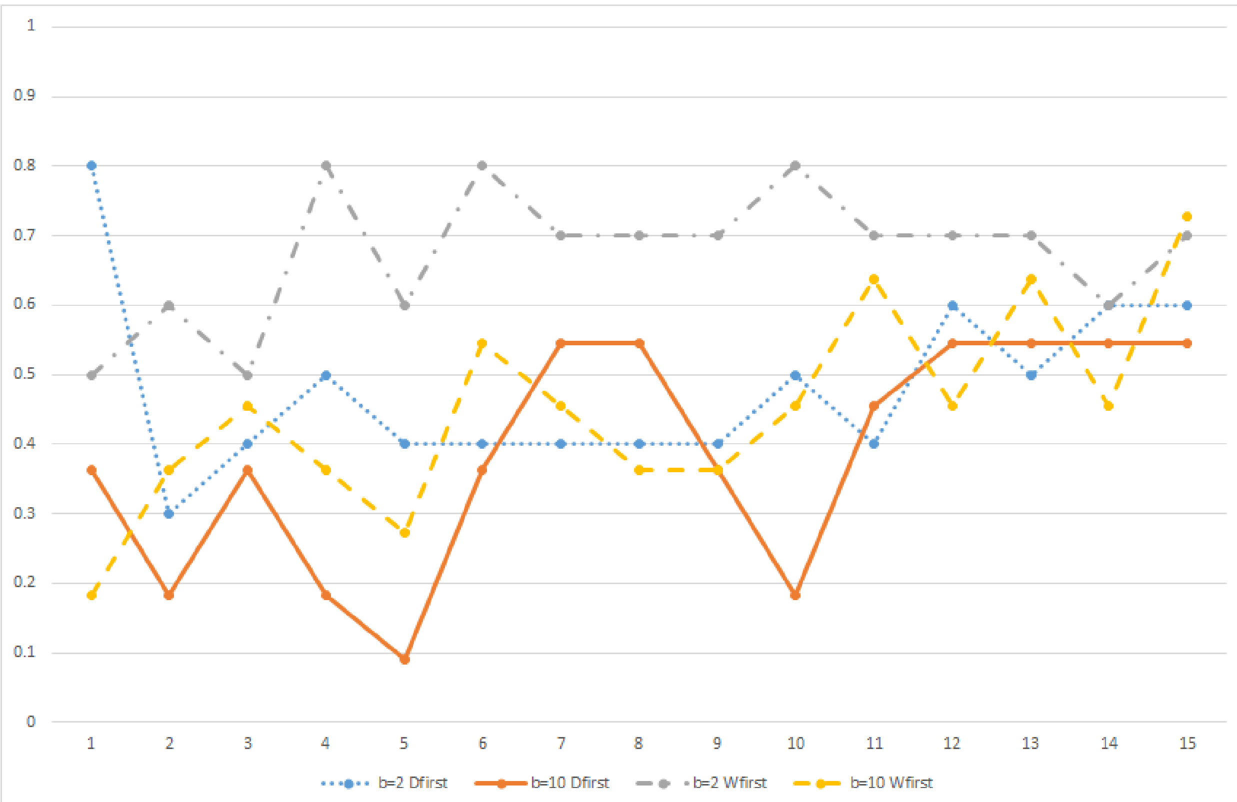


Figure 5: Fraction of Subjects Who Wrote Optimal Rule Actions Given Their Rules by Period

Table 2: Does the Contract Cover an Edge?

Treatments	Contract Covers Edge	Periods 1-5	Periods 6-10	Periods 11-15
Randomization Probability		0.524	0.524	0.524
b=2, Dfirst	0.847*** $(n = 150)$	0.8*** $(n = 50)$	0.82*** $(n = 50)$	0.92*** $(n = 50)$
b=10, Dfirst	0.782*** $(n = 165)$	0.745*** $(n = 55)$	0.709*** $(n = 55)$	0.891*** $(n = 55)$
b=2, Wfirst	0.793*** $(n = 150)$	0.74*** $(n = 50)$	0.8*** $(n = 50)$	0.84*** $(n = 50)$
b=10, Wfirst	0.824*** $(n = 165)$	0.782*** $(n = 55)$	0.873*** $(n = 55)$	0.818*** $(n = 55)$

*, **, and *** indicate statistical significance at the 10%, 5%, and 1% confidence intervals respectively using a one-sample t-test to test whether the indicated treatment has a significantly higher mean than is expected by random play.

Table 3: Do Subjects Write Rule Actions Optimally?

Treatments	Action Optimal Given Rule	Periods 1-5	Periods 6-10	Periods 11-15	Predicted Under Randomization
b=2, Dfirst	$0.48^{***}(n = 150)$	$0.48^{***}(n = 50)$	$0.42^{***}(n = 50)$	$0.54^{***}(n = 50)$	0.25
b=10, Dfirst	$0.448^{***}(n = 165)$	$0.236^{*}(n = 55)$	$0.4^{***}(n = 55)$	$0.527^{***}(n = 55)$	0.15
b=2, Wfirst	$0.673^{***}(n = 150)$	$0.6^{***}(n = 50)$	$0.74^{***}(n = 50)$	$0.68^{***}(n = 50)$	0.25
b=10, Wfirst	$0.448^{***}(n = 165)$	$0.327^{***}(n = 55)$	$0.436^{***}(n = 55)$	$0.582^{***}(n = 55)$	0.15

*, **, and *** indicate statistical significance at the 10%, 5%, and 1% confidence intervals respectively using a one-sample t-test to test whether the indicated treatment has a significantly higher mean than is expected by random play.

Table 4: Number of States in Each Contract

Treatments	Average Number of States	Periods 1-5	Periods 6-10	Periods 11-15
b=2, Dfirst	3.92 (n=150)	4.04 (n=50)	3.64 (n=50)	4.08 (n=50)
b=10, Dfirst	4.461 (n=165)	4.109 (n=55)	4.382 (n=55)	4.891 (n=55)
t-Stat for difference between b=2 and b=10	-3.063***	-0.218	-2.464***	-2.795***
b=2, Wfirst	4.053 (n=150)	4.18 (n=50)	4.04 (n=50)	3.94 (n=50)
b=10, Wfirst	4.879 (n=165)	4.564 (n=55)	5.091 (n=55)	4.982 (n=55)
t-Stat for difference between b=2 and b=10	-4.485***	-1.176	-3.283***	-3.357***
All	0.811	0.767	0.8	0.867

*, **, and *** indicate statistical significance at the 10%, 5%, and 1% confidence intervals respectively using a two-sample t-test for differences in means.

In periods 1-5, the number of states covered by contracts (Figure 6 and Figure 7) are statistically indistinguishable. However, in periods 6-10 and 11-15, there is a large separation in both Dfirst and Wfirst. This is strong evidence of a learning effect. In both treatments, periods 6-10 and 11-15 have t-statistics that indicate significant difference at the 5% level, which can be observed in table 4.[2]

Result 3: Prediction 3 is supported. The number of states covered in each contract in the $b = 2$ treatment is significantly different from the $b = 10$ treatment.

[2]In addition, a Z-test was run with clustered standard errors to combat any doubt about independence across periods for the same individual. Unfortunately, the significance disappears (becomes significant at the 15% level). However, when the treatments are pooled, the two periods are still significantly different. This is likely due to the smallness of the sample size.

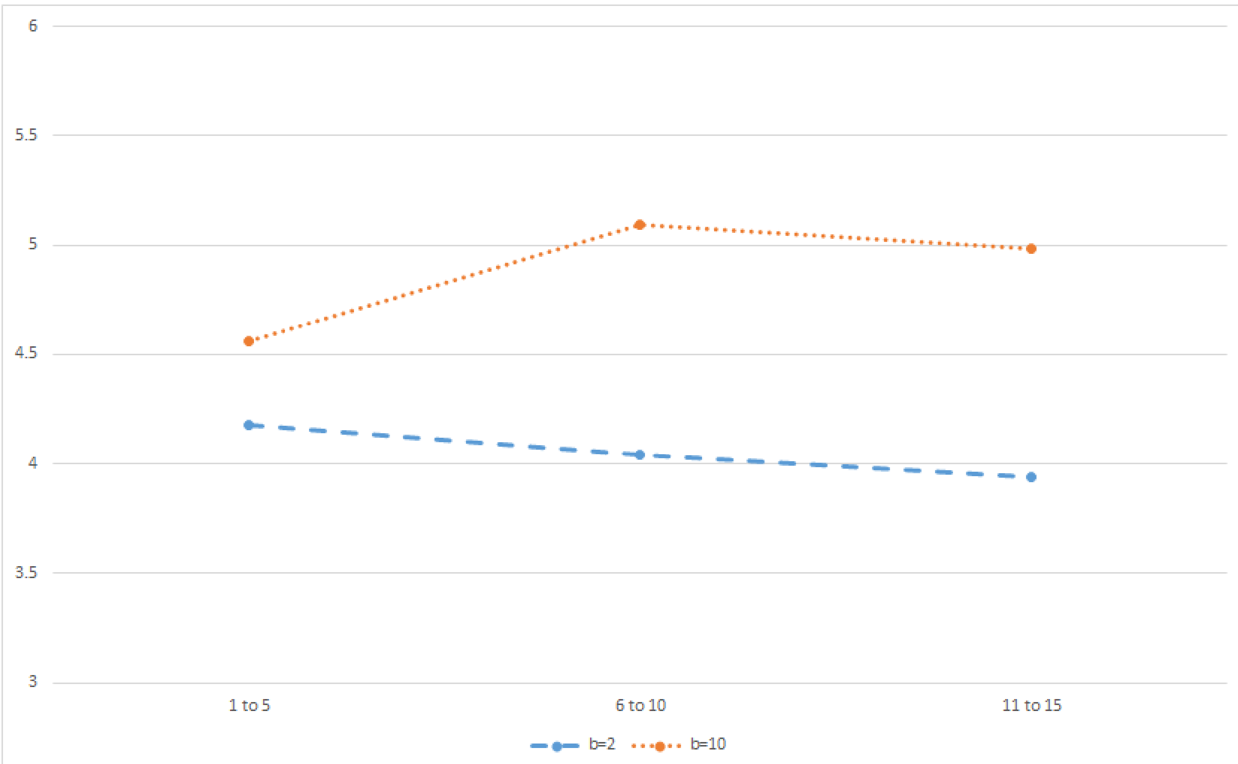


Figure 6: Number of States Covered by a Contract in Wfirst for b=2 and b=10, Including the Predicted SPE

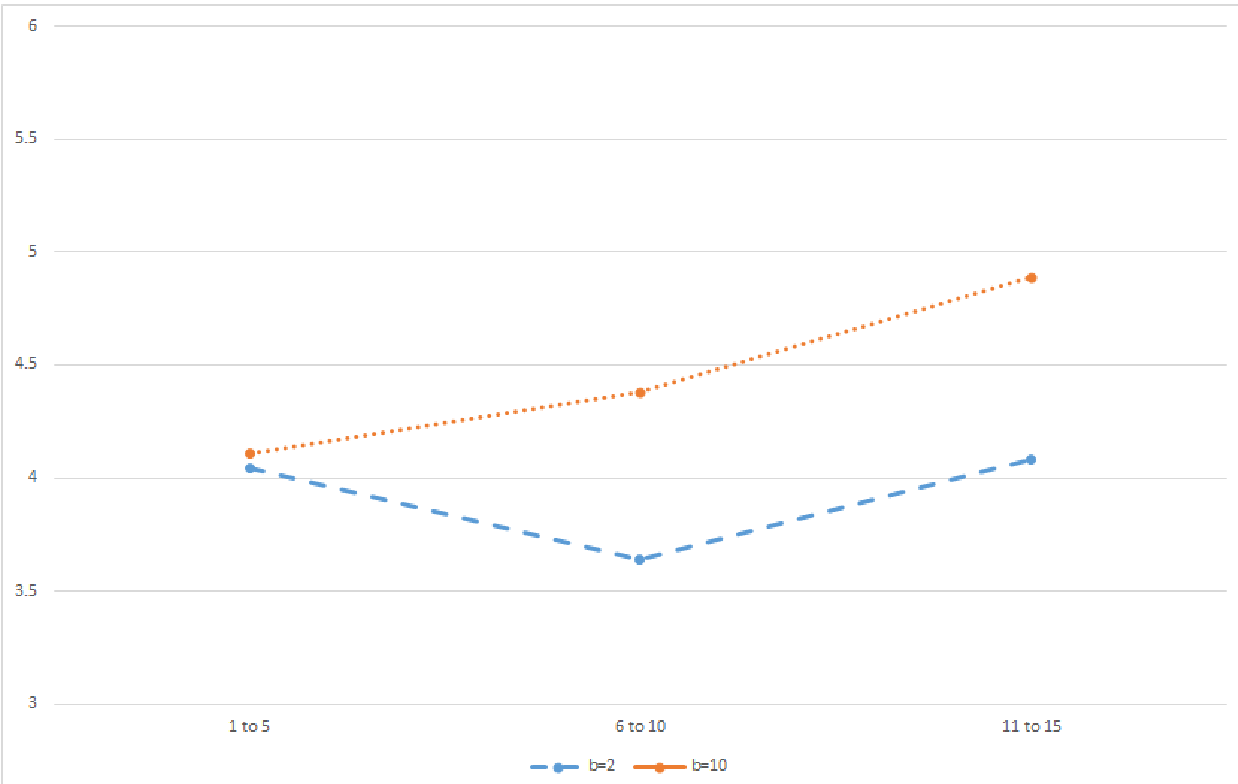
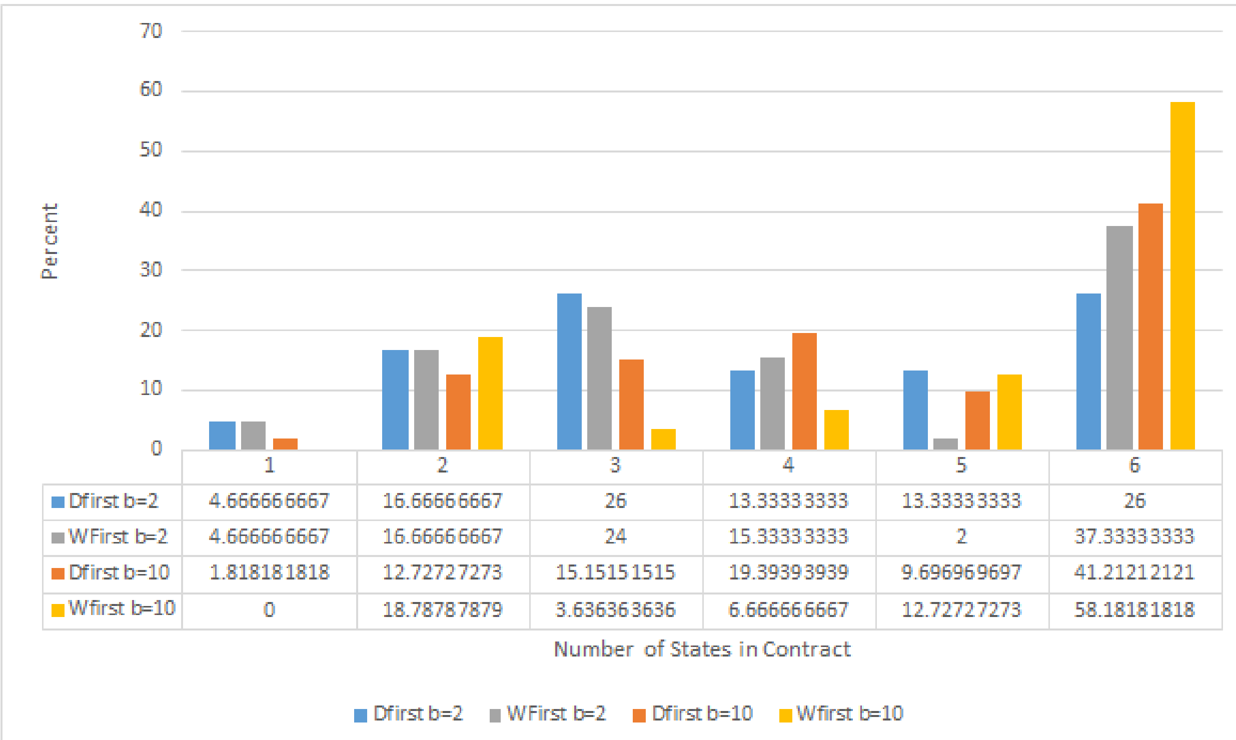


Figure 7: Number of States Covered by a Contract in Dfirst for b=2 and b=10, Including the Predicted SPE



	1	2	3	4	5	6
Dfirst b=2	4.666666667	16.66666667	26	13.33333333	13.33333333	26
WFirst b=2	4.666666667	16.66666667	24	15.33333333	2	37.33333333
Dfirst b=10	1.818181818	12.72727273	15.15151515	19.39393939	9.696969697	41.21212121
Wfirst b=10	0	18.78787879	3.636363636	6.666666667	12.72727273	58.18181818

Figure 8: Distribution of the Number of States in a Contract over All Treatments

Table 5: In Dfirst, Are Decider Choices Possible SPE?

Treatments	Action Optimal Given Rule	Periods 1-5	Periods 6-10	Periods 11-15	Predicted Under Randomization
b=2, Dfirst	$0.269 (n = 150)$	$0.190 (n = 50)$	$0.26 (n = 50)$	$0.44 (n = 50)$	0.583
b=10, Dfirst	$0.636^{***} (n = 165)$	$0.655^{**} (n = 55)$	$0.618^{**} (n = 55)$	$0.636^{**} (n = 55)$	0.5

*, **, and *** indicate statistical significance at the 10%, 5%, and 1% confidence intervals respectively using a one-sample t-test to test whether the indicated treatment has a significantly higher mean than is expected by random play.

Table 6: In Wfirst, are s_{low} and s_{high} consistent with SPE?

Treatments	Action Optimal Given Rule	Periods 1-5	Periods 6-10	Periods 11-15	Predicted Under Randomization
b=2, Wfirst	$0.373^{**}(n = 150)$	$0.28(n = 50)$	$0.38^{*}(n = 50)$	$0.46^{***}(n = 50)$	0.286
b=10, Wfirst	$0.612^{***}(n = 165)$	$0.545^{***}(n = 55)$	$0.655^{***}(n = 55)$	$0.636^{***}(n = 55)$	0.095

*, **, and *** indicate statistical significance at the 10%, 5%, and 1% confidence intervals respectively using a one-sample t-test to test whether the indicated treatment has a significantly higher mean than is expected by random play.

Table 5, displays whether Deciders who move first have play consistent with SPE, while Table 6 shows the same test for Writers who move first. Writers in the Wfirst $b = 10$ treatment and Deciders in the Dfirst $b = 10$ treatment chose actions that align with SPE predictions far more often than would be predicted by random play. Writers in the Wfirst $b = 2$ treatment write contracts that do better than a random subject. However Deciders in the Dfirst $b = 2$ treatment do far worse than a random player, although the trend is in the correct direction over time. As can be seen in figure 9, 38% of Deciders pick a default action of 8. This, coupled with the fact that 27% of Deciders pick a default action of 12 in the Dfirst $b = 10$ treatment (Figure 10), indicate that some bias is influencing the way that Deciders choose default actions.

Result 4: Prediction 4 is supported for Writers and Deciders in the Dfirst, $b = 10$ treatment. Writers write contracts that align with predictions far more frequently than random play would suggest. Prediction 4 is not supported for Deciders in the Dfirst, $b = 2$ treatment. Although more Deciders select an equilibrium default action over time, Deciders select fewer default actions that are consistent with predictions than a completely random player.

One aspect of these results worth discussing is that across all of the findings, there is some evidence for learning taking place, which is seen for the Writers in Figures 6, and 7 and for the Deciders in figures 11 and 12. This game has a high degree of complexity, so it is encouraging that there seems to be a learning effect. The following section will examine whether play is in accordance with a level-k model.

4.2 Behavioral Preferences

In many comments on the post-experimental questionnaire, subjects mentioned fairness when discussing their view of the game. Subjects wrote about working together to create outcomes that were good for both parties and sometimes discussed the Writer having too much control. In addition, when looking at the Decider's best response in period two (Fig. **??**), Deciders play actions that tend to be well below the best action to play. Also, as discussed in the previous section, Deciders and Writers in the first period write suboptimal default actions and contracts. This indicates that analysis of fairness and level-k models may be fruitful.

Level-k

The starting point this paper will use for level-0 is that the Decider and Writer both randomize over all actions for level 0. Level 1 for the Decider is a best response to the Writer's

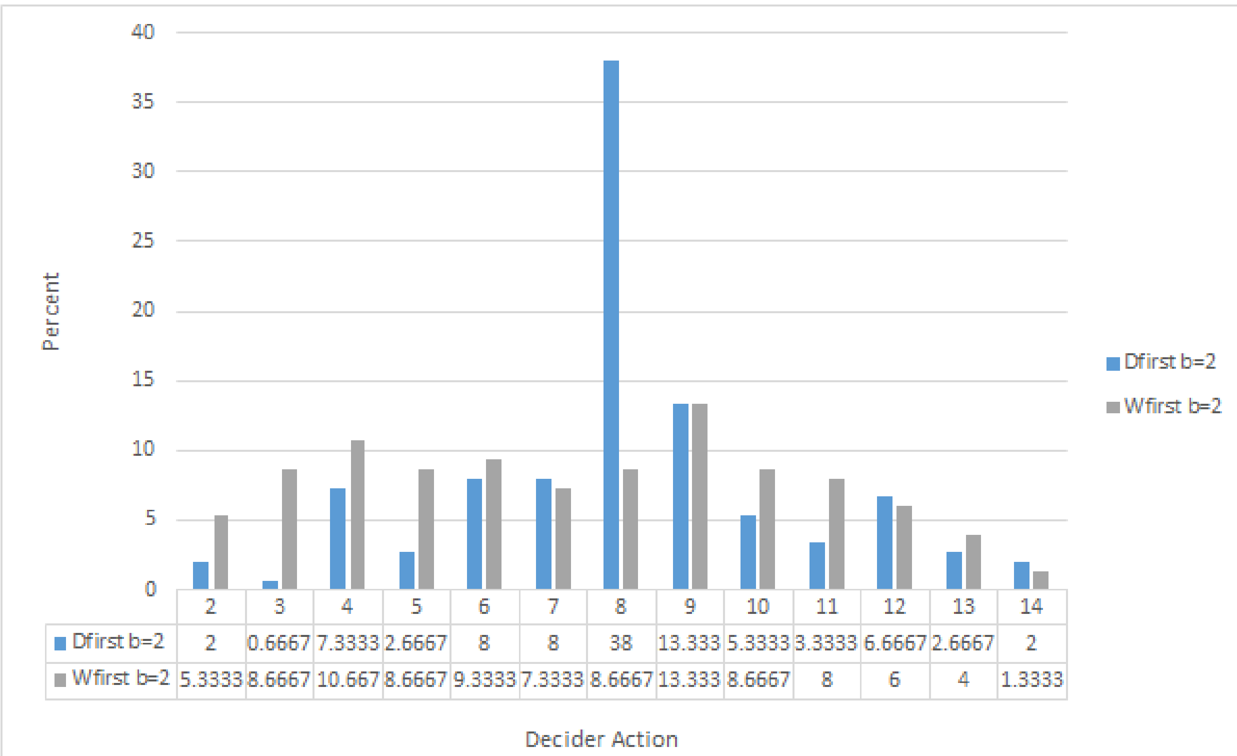
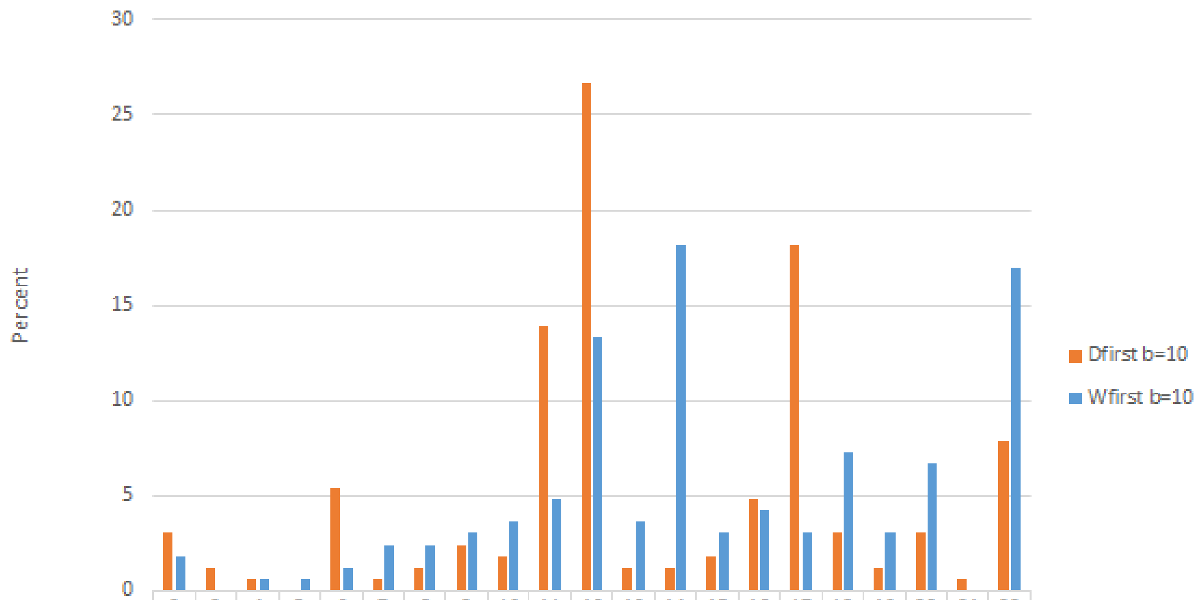


Figure 9: Distribution of Decider Actions in the b=2 Treatment



	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Dfirst b=10	3	1.2	0.6	0	5.5	0.6	1.2	2.4	1.8	14	27	1.2	1.2	1.8	4.8	18	3	1.2	3	0.6	7.9
Wfirst b=10	1.8	0	0.6	0.6	1.2	2.4	2.4	3	3.6	4.8	13	3.6	18	3	4.2	3	7.3	3	6.7	0	17

Figure 10: Distribution of Decider Actions in the b=10 Treatment

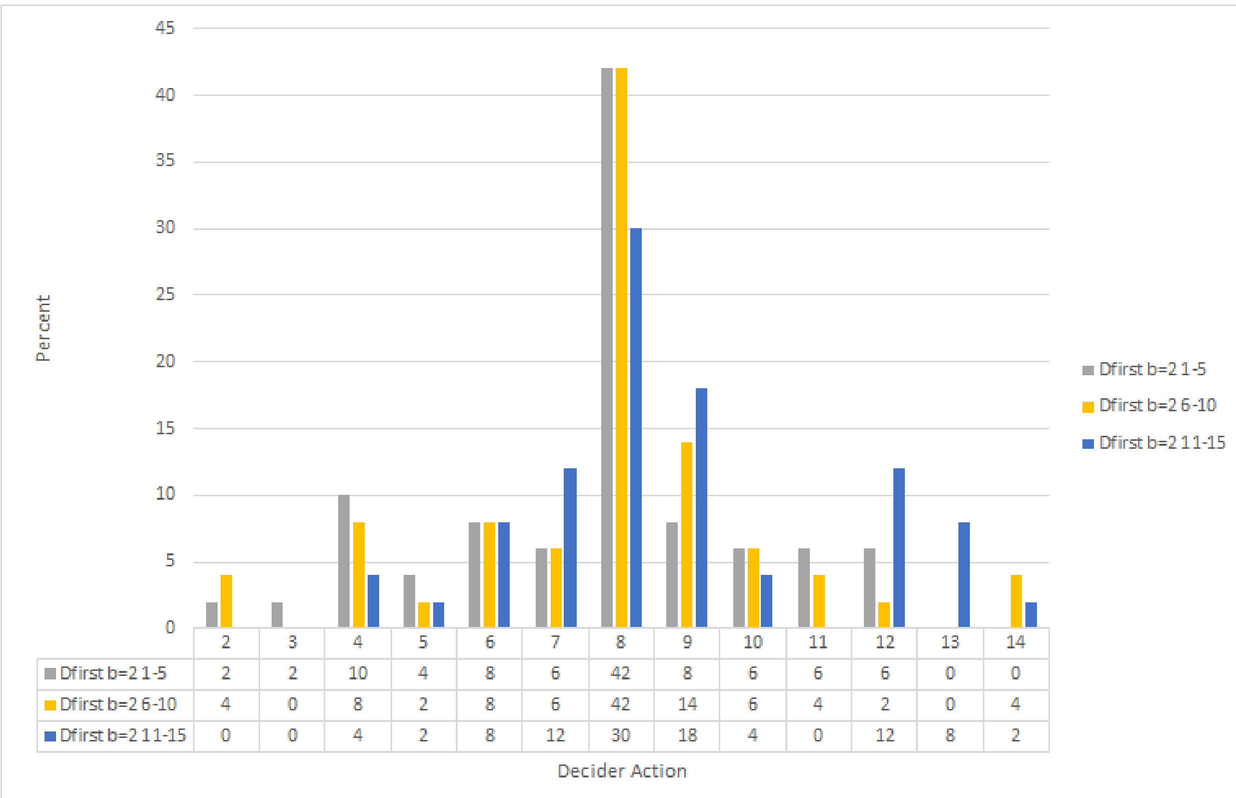
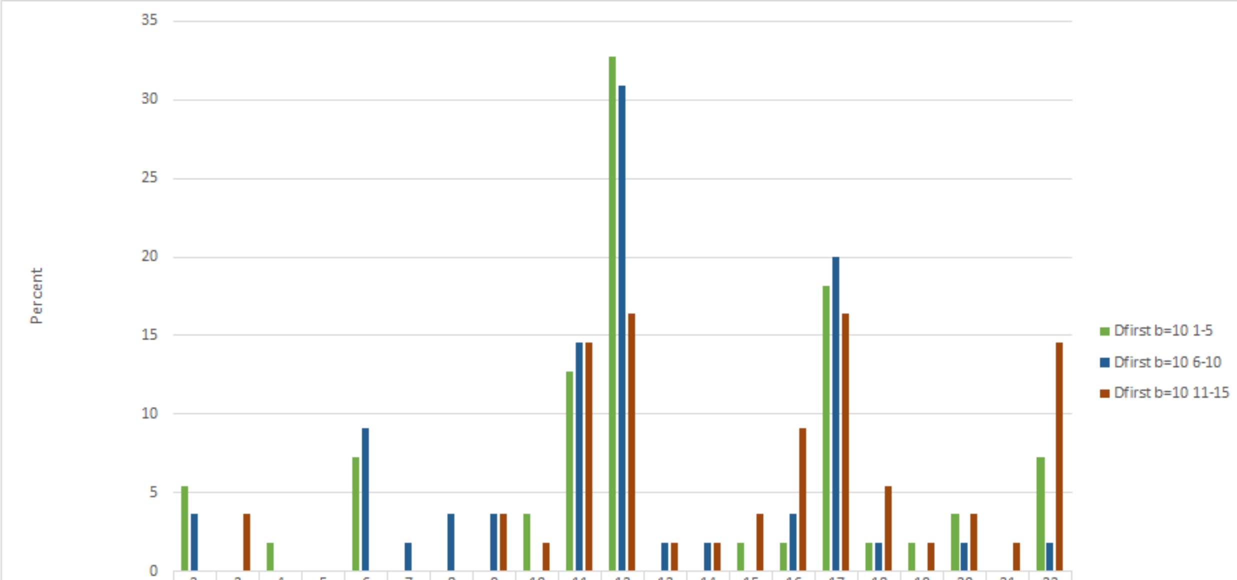


Figure 11: Distribution of Decider Actions by Period in b=2 Treatment



	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Dfirst b=10 1-5	5.455	0	1.818	0	7.273	0	0	0	3.636	12.73	32.73	0	0	1.818	1.818	18.18	1.818	1.818	3.636	0	7.273
Dfirst b=10 6-10	3.636	0	0	0	9.091	1.818	3.636	3.636	0	14.55	30.91	1.818	1.818	0	3.636	20	1.818	0	1.818	0	1.818
Dfirst b=10 11-15	0	3.636	0	0	0	0	0	3.636	1.818	14.55	16.36	1.818	1.818	3.636	9.091	16.36	5.455	1.818	3.636	1.818	14.55

Figure 12: Distribution of Decider Actions by Period in b=10 Treatment

level 0, level 1 for the Writer is a best response to the Decider's level 0, and so on. This starting point seems the most sensible, as players have many things to consider and may just pick an action that is good against a random selection of the other players. The $\frac{3}{5}$ rule is effective for fitting people into levels: If, in the last 10 periods of the game, at least $\frac{3}{5}$ of a subject's play corresponded to a certain level, then they will be classified as that level. In addition, for Writers, I will allow them an error of 1 away from the Writer Action that fits into a level. This means that, for example, a Writer in the $b = 2$ treatment who writes the contract $(1, 2, 5)$ would be classified as a level 2 rather than a level 0 in that period. In tables 6 and 7, after the first and second level, the third level looks similar to the first level. For this reason, I will only focus on levels 0 through 2.

Table 7: Table 6: Levels in b=2

	Writer Contract (s_{low}, s_{high}, a_W)	Decider Action
Level 0	Randomize uniformly between all possible contracts	Randomize uniformly between $\{4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}$
Level 1	$\{((1, 3), 4)\}$	$\{8, 9, 10\}$
Level 2	$\{((1, 2), 2), ((1, 2), 3), ((1, 2), 4), ((1, 3), 4)\}$	$\{10\}$
Level 3	$\{((1, 3), 4)\}$	$\{8, 9, 10\}$

Table 8: Levels in b=10

	Writer	Decider
Level 0	Randomize uniformly between all possible contracts	Randomize uniformly between $\{12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22\}$
Level 1	$\{((1, 5), 6)\}$	$\{16, 17, 18\}$
Level 2	$\{((1, 5), 6), ((1, 6), 5), ((1, 6), 6), ((1, 6), 7)\}$	$\{22\}$
Level 3	$\{((1, 6), 5), ((1, 6), 6), ((1, 6), 7)\}$	$\{22\}$, Randomize uniformly between $\{12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22\}$

Table 9: Results of Level-k Analysis When Player Moves First

	Writers in level in $b = 2$	Deciders in level in $b = 2$	Writers in level in $b = 10$	Deciders in level in $b = 10$
Level 0	8	5	8	6
Level 1	1	5	1	2
Level 2	1	0	2	0
Total Number of Subjects	10	10	11	11

Table 10: Results of Level-k Analysis When Player Moves Second

	Writers in level in $b = 2$	Deciders in level in $b = 2$	Writers in level in $b = 10$	Deciders in level in $b = 10$
Level 0	7	8	7	9
Level 1	1	1	0	1
Level 2	2	1	4	0
Total Number of Subjects	10	10	11	11

For the results of the level-k analysis, which can be seen in Tables 9 and 10, the results immediately point away from level-k being explanatory. Firstly, in the $b = 2$ and $b = 10$ treatments, if level-k held it is expected that first movers in both treatments would appear to be similar in the data. Secondly, far too much of the data is explained by level 0, which is where most of the subjects are classified since they don't tend to play according to the levels.[3] This indicates that level-k does not explain behavior in this setting.

Result 5: This model of level-k does not explain behavior in this experiment, as many of the subjects are classified as level-0 which coincides with random play.

5 Conclusion

This paper proposes an experiment that studies how incomplete contracts are written and how well theory predicts what interpretive rules and contracts are played in a laboratory setting. Subjects tend to write their contracts correctly in many ways. Writer actions, contract length, and the location of the contract are all played in accordance with theory. Additionally there is some evidence that first movers select actions that are consistent with subgame perfect equilibrium.

One factor that goes against predictions is that default actions for Deciders who move first are not focused on equilibrium. Additionally, in the post-experimental questionnaire, subjects indicated that fairness may be at play. This would make the out-of-equilibrium play make sense, as play by Deciders when they were first movers was focused around actions right in the middle of what was possible. One way to test this would be to rerun the experiment with different possible actions, as this could be an unconscious bias towards default actions in the middle of the state space.

In total, there is room for further exploration involving how people learn to write these contracts and default rules, since there is evidence that subjects learn how to take actions that have characteristics predicted in equilibrium. The findings presented in this paper indicate that there is more exploration necessary to grasp the full picture that links the intentionally incomplete contracting literature to how contracts may be written by human subjects.

[3]Some of the subjects are classified as nothing in the $b = 10$ treatment where Deciders played actions below 12 a majority of the time.

References

[1] Aghion, P., & Bolton, P. (1987). Contracts as a Barrier to Entry. The American economic review, 388-401.

[2] Battigalli, P. and Maggi, G. (2002). Rigidity, discretion and the costs of writing contracts, American Economic Review, vol. 92, pp. 798–817.

[3] Binmore, K., McCarthy, J., Ponti, G., Samuelson, L., & Shaked, A. (2002). A backward induction experiment. Journal of Economic theory, 104(1), 48-88.

[4] Brandts, Gary Charness, and Matthew Ellman. "How communication affects contract design: An experimental study of formal and informal contracting." (2012).

[5] Cai, Hongbin, and Joseph Tao-Yi Wang. "Overcommunication in strategic information transmission games." Games and Economic Behavior 56.1 (2006): 7-36.

[6] Dufwenberg, Martin, Alec Smith, and Matt Van Essen. "Hold-Up: With a Vengeance." Economic Inquiry 51.1 (2013): 896-908.

[7] Dye, R. (1985). Costly contract contingencies, International Economic Review, vol. 26, pp. 233–50.

[8] Ellingsen, Tore, and Magnus Johannesson. "Promises, threats and fairness." The Economic Journal 114.495 (2004): 397-420.

[9] Fehr, Ernst, Alexander Klein and Klaus Schmidt (2007), " Fairness and Contract Design," Econometrica 75, 121–154.

[10] Fehr, Ernst, Michael Powell, and Tom Wilkening. "Handing out guns at a knife fight: behavioral limitations of subgame-perfect implementation." (2014).

[11] Hackett, Steven C. "Incomplete contracting: a laboratory experimental analysis." Economic Inquiry 31.2 (1993): 274.

[12] Hart, O., & Moore, J. (1988). Incomplete contracts and renegotiation. Econometrica: Journal of the Econometric Society, 755-785. Chicago

[13] Heller, D. and Spiegler, R. (2008). Contradiction as a Form of Contractual Incompleteness, The Economic Journal, vol. 118, pp. 875-888.

[14] Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. American economic review, 92(5), 1644-1655.

[15] Mansor, Nur Syaimasyaza Bt, and Khairuddin Abd Rashid. "IDENTIFYING THE PRESENCE OF INCOMPLETE CONTRACT IN PFI CONTRACTS."

[16] Marques, Rui Cunha, and Sanford Berg. "Risks, contracts, and private-sector participation in infrastructure." Journal of Construction Engineering and Management 137.11 (2011): 925-932.

[17] Maskin, Eric, and Jean Tirole (1999). "Unforeseen contingencies and incomplete contracts." The Review of Economic Studies 66.1 : 83-114.

[18] Simon, Herbert A. "Organizations and markets." The Journal of Economic Perspectives 5.2 (1991): 25-44.

[19] Shavell, S. (2006). On the writing and the interpretation of contracts, Journal of Law, Economics and Organization, vol. 22, pp. 289–314.

[20] Urs Fischbacher (2007): z-Tree: Zurich Toolbox for Ready-made Economic Experiments, Experimental Economics 10(2), 171-178.

Appendix A: Instructions

The following are the Instructions for the Wfirst, b=2 treatment:

Instructions C

Welcome! In this experiment, your earnings will depend on your choices, the choices of others, and chance. Please refrain from talking to others until the experiment has concluded. In addition, please silence and put away any electronic devices.

The participants of this experiment will be randomly split between Writers and Deciders, such that half will be Writers and the other half will be Deciders. You will play only as a Writer or as a Decider for the duration of the experiment. In each round, each Writer will be randomly paired with a Decider. You will not see the identity of the person you are paired with, but you will see each player's decisions at the end of each round.

Your payment in each round, in Experimental Currency Units (ECUs), depends on a randomly drawn state and on choices both players will make that dictate an action for each of those states. This action is decided in part by the Writer, who moves first, and in part by the Decider, who moves second after observing the Writer's choices. The details of this process will be described below.

States: There are 6 random states that can occur, numbered 1, 2, 3, 4, 5, and 6. The states that occur in this experiment will be computer generated and **all states will be equally likely** in each round. There will be a state drawn after the Writer and the Decider have moved at the end of each round.

Writer's Rule: In this experiment, the Writer will be writing a rule. This rule will indicate a 'low state,' a 'high state,' and a 'rule action.' The 'low state' can be any state (1 through 6). The 'high state' can be any state (1 through 6) that is higher than or equal to the 'low state.' The 'rule action' can be any action from 2 to 14. This rule will help to determine the action that is taken. **If a state is drawn that is between 'low state' and 'high state', or equal to either of these states, the rule will dictate that the 'rule action' is taken.**

Decider's Decision: After the Writer writes his/her rule, the Decider will observe this and then decide on a 'default action,' which can be any action between 2 and 14. **In the case that the state that is drawn at the end of the period is either lower than the 'low state' or higher than the 'high state' specified in the Writer's rule, the 'default action' will be taken.**

In summary, each round of the experiment will be as follows:

First: The Writer will write a rule that indicates three things: 'low state', 'high state', and 'rule action.' The 'high state' must be a state with number higher than or equal to the number the Writer writes down for 'low state.' The 'rule action' can be any action from 2 to 14.

Second: Then, after the Writer writes a rule, the Decider will decide on a 'default action.' This action will be the 'default action' taken in the case that the state drawn is lower than 'low state' or higher than 'high state'. The 'default action' can be any action from 2 to 14.

Third: After the Decider decides on a 'default action,' the state will be drawn. The 'rule action' will be taken if the state is between 'low state' and 'high state' or at 'low state' or 'high state.' Otherwise, if the state drawn is less than 'low state' or higher than 'high state,' the 'default action' will be taken.

At the end of each round, you will be shown the decisions of both you and your partner for the period, the action taken, and your earnings for the round. Your payout, depending on the action and state, is detailed on the next page.

You will play 2 practice rounds of the task by yourself as both the Writer and the Decider where you will be quizzed on the things that happen in those trials at the end of each trial period. After that, the task will be repeated 15 times, with random matching in each round, and where your role will stay fixed as either the Writer or the Decider. **Your total earnings from this experiment will be your earnings from 2 of the 15 periods, drawn randomly by you at the end of the experiment, plus your show up fee of $6.** The payments in each period will be recorded in Experimental Currency Units (ECUs). **Each ECU is worth 30 cents (.3 dollars).**

The Writer's payout (in ECUs) followed by The Decider's payout (in ECUs) is presented below for each action and state:

Payouts		*State*					
		1	2	3	4	5	6
	2	(20, 18)	(18, 16)	(16, 14)	(14, 12)	(12, 10)	(10, 8)
	3	(19, 19)	(19, 17)	(17, 15)	(15, 13)	(13, 11)	(11, 9)
	4	(18, 20)	(20, 18)	(18, 16)	(16, 14)	(14, 12)	(12, 10)
	5	(17, 19)	(19, 19)	(19, 17)	(17, 15)	(15, 13)	(13, 11)
	6	(16, 18)	(18, 20)	(20, 18)	(18, 16)	(16, 14)	(14, 12)
	7	(15, 17)	(17, 19)	(19, 19)	(19, 17)	(17, 15)	(15, 13)
Action	8	(14, 16)	(16, 18)	(18, 20)	(20, 18)	(18, 16)	(16, 14)
	9	(13, 15)	(15, 17)	(17, 19)	(19, 19)	(19, 17)	(17, 15)
	10	(12, 14)	(14, 16)	(16, 18)	(18, 20)	(20, 18)	(18, 16)
	11	(11, 13)	(13, 15)	(15, 17)	(17, 19)	(19, 19)	(19, 17)
	12	(10, 12)	(12, 14)	(14, 16)	(16, 18)	(18, 20)	(20, 18)
	13	(9, 11)	(11, 13)	(13, 15)	(15, 17)	(17, 19)	(19, 19)
	14	(8, 10)	(10, 12)	(12, 14)	(14, 16)	(16, 18)	(18, 20)