Predictive Repeated Game Theory: Measures and Experiments*

November 11, 2014

Laurent Mathevet† Julian Romero‡

Abstract:

One of the fundamental results in repeated games is the Folk theorem, which predicts a plethora of equilibrium outcomes. Many have argued that this extreme number of equilibria is a virtue, as it can explain a variety of different behaviors. However, this result leaves us with almost no predictive power. This paper provides measures for evaluating the predictive power of a theory given experimental data. After running experiments with human subjects in the experimental laboratory, we use these measures to compare a variety of different theories including Mathevet (2014)'s axiomatic approach, and Ioannou and Romero (2014)'s learning model.

Keywords: Repeated Games, Prisoners' Dilemma, Battle of the Sexes, Stag Hunt, Chicken, Experimental Economics, Equilibrium Selection

† Economics Department, New York University • Email:lmath@nyu.edu
‡ Economics Department, Krannert School of Management, Purdue University • Email:jnromero@purdue.edu

1 Introduction

There exist many methodological approaches to economics.[1] On one end of the spectrum, apriorism argues that theoretical conclusions are credible if they follow from sound assumptions and correct deductive logic (Rothbard (1957)). According to this school of thought, there is genuine knowledge outside sensory experience; there is no need for empirical "testing;" and the only proper evaluation of a theory is internal.[2] On the other end of the spectrum, extreme empiricism asserts that knowledge comes only from sensory experience, and thus, requires that both the conclusions *and* the assumptions of a theory be empirically validated (Homan (1931) and Commons (1934)). According to this school of thought, the only proper evaluation of a theory is external. In his famous methodological essay, Friedman (1953) adopts an intermediate position: empirical verification of theoretical conclusions is proper, but that of assumptions is irrelevant. Indeed, a model can only be useful if it abstracts from reality, which often requires unrealistic assumptions; but ultimately, theories must make valid predictions about phenomena not yet observed.

This paper focuses on predictive power along the lines of Friedman's precepts. We first axiomatize a binary relation that orders theories according to their predictive power. We then run repeated-game experiments in the laboratory. Finally, based on our measure and experiments (and other measures), we rank various theories of repeated interactions according to predictive power. In this sense, our endeavor is identical to Selten's, who axiomatized measures of predictive power in Selten (1991, 1998) and applied them to experimental static and repeated (random matching) games in Selten and Chmura (2008) and Chmura, Goerg, and Selten (2011).

Let us be clear about terminology. Predictive power is the ability of a theory to make predictions that match observations; the predictions are made before observations occur. Descriptive power is the ability of a theory to make predictions that match observations; the predictions are made after observations occur, giving the possibility to build the theory from the data. Of course, descriptive power based on a particular data set does not imply predictive power (i.e., good predictions about other data sets). To put it in another way, in-sample predictability does not imply out-of-sample forecastability. Finally, explanatory power is the ability of a theory to offer causal relationships that explain how and why a phenomenon occurs. If explanations are made before observations, then great explanatory power implies great predictive power, but the converse is not true.

Our methodological stance is simple: predictive power is by no means a necessary virtue, but excluding it from the reasonable criteria seems unreasonable. In particular, claims that game theory is inexact by nature and that empirical difficulties are unavoidable are weak arguments for the exclusion of predictive success. A theory need not make perfect predictions to be useful from a predictive standpoint. We also recognize the importance of other criteria, such as elegance, simplicity, internal coherence, delivering an insightful message (as a fable; see Rubinstein (2006)),

[1] See, for example, Hausman (1998) for a survey.

[2] The multiplicity of causes behind economic phenomena (Mill (1843)) and the importance of subjective factors (Knight (1940), von Mises (1949)) make empirical difficulties unavoidable in economics. Thus, it is ill-advised to seek empirical validation.

and explanatory power; but they are not the focus of this paper.

We ask how a researcher who cares about predictive power should rank theories according to this criterion. We answer by axiomatizing an ordering relation that we apply to the field of repeated interactions. This choice is driven by several reasons.

There exist numerous theories and approaches of repeated interactions, ranging from equilibrium theory and its refinements (Mailath and Samuelson (2006)) to many learning models (Fudenberg and Levine (1998)). In this context where different theories may have different purposes, it is a clarifying experience to establish an objective ranking based on a single criterion.

We also want to draw attention to a form of paradox: the theoretical literature on repeated interactions is characterized by extreme multiplicity, but not the data. This extreme multiplicity appears in two forms. In equilibrium analyses, which use variants of (subgame perfect) Nash equilibrium, extreme multiplicity is a well-known fact that culminates with folk theorems: every individually rational payoff can be an equilibrium payoff. In non-equilibrium analyses, multiplicity takes a different form. There are so many learning models[3] that the combination of their predictions causes extreme multiplicity. In general, learning models make parsimonious predictions, although randomness and changes in initial conditions still result in multiple outcomes. But extreme multiplicity is due to the divergence in predictions across models, and to the explosion of predictions once we match different learning rules against each other.[4] Despite these theoretical results describing a form of chaos, our experimental data show strong regularities, even when players are extremely patient.

Our previous point encourages the creation of new models and approaches. This gives us the opportunity to test recent theories by Ioannou and Romero (2014) and Mathevet (2014). Our emphasis on these papers is not only partisan but also methodological, as they advocate different directions for future research. In general, a researcher can either model boundedly rational behaviors precisely, for example, by specifying learning algorithms, or she can describe the implications or the outcomes of boundedly rational behaviors. Ioannou and Romero (2014) explore the first route and prescribe the use of comprehensive models. The idea is that, by combining various learning rules, we might improve the theory. They propose a learning model a la Camerer and Ho (1999) — a hybrid model combining several learning rules — in which learning occurs over repeated-game strategies and not stage-game actions. By doing so, they advocate the construction of more detailed and complete learning models. Mathevet (2014) explores the second route, prescribing the use of an abstract (model-free) approach. Instead of proposing precise models of bounded rationality, he proposes to formulate plausible recipes for less than rational play. These recipes take the form of sets of axioms on the infinite sequences of play. The author presents it as a framework in which different combinations of axioms lead to different solutions, and hence different representations of

[3]For example, adaptive learning (Milgrom and Roberts (1990)), reinforcement learning (Er'ev and Roth (1998)), aspiration-based learning (Karandikar, Mookherjee, Ray, and Vega-Redondo (1998)), pattern recognition learning (Sonsino (1997)), regret-based learning (Hart and Mas-Colell (2000)), etc.

[4]There is no reason to believe that different players should have the same learning rule. Therefore, studying the outcome of situations in which different learning models face each other is justified.

boundedly rational play. Ioannou and Romero (2014) and Mathevet (2014) meet in their pursuit of sharp predictions, but they differ on other dimensions. For example, abstract approaches favor parsimony and flexibility, while descriptive approaches favor explanatory power.

Our main theoretical result characterizes a binary relation from four axioms. These axioms aim to impose reasonable conditions on the comparison of theories based on predictive power. For example, *success monotonicity* requires the predictive power of a theory to increase if that theory is modified by adding one correct prediction (correct w.r.t. some data set). Similarly, *error monotonicity* requires the predictive power of a theory to decrease if that theory is modified by adding one incorrect prediction (incorrect w.r.t. some data set). Selten (1991)'s measure satisfies error monotonicity, but violates success monotonicity. Given the use of Selten's measure in applied works (e.g., Beatty and Crawford (2011)), we will discuss the implications of this violation.

The experimental results in this paper are important for two reasons. First, we wanted to collect a large, clean data set that shows how individuals behave in these simple repeated games with fixed matching and long supergames. On this front, despite inevitable noise in the data, a quick look at the data shows that there are clear regularities that present themselves. Moving forward, this is promising for both current and future work suggesting that there are likely more regularities in these interactions than standard equilibrium theory would suggest. Second, and more important for this paper, this experimental data set allows us to do a thorough comparative analysis of a wide range of theories using the measures described above. The result of this comparative analysis is an aggregate ranking of theories in terms of predictability across the eight games studied. Among the different rankings studied in this paper, the axiomatic approach to repeated games (Mathevet (2014)) and the strategic EWA model (Ioannou and Romero (2014)) rank at the top in terms of predictability when compared to other theories. This is not an indictment of the other theories presented, as the goal of the other theories is not necessarily predictability. However, this does provide strong support for the two theories mentioned above.

While there is a large experimental literature on repeated games, our paper is the first comparative study of repeated game theories. The previous experimental literature has focused on the qualitative aspects of cooperation, such as the extent, the evolution, or the robustness of cooperation in repeated games, especially in the Prisoners' Dilemma and related environments (e.g., Holt (1985), Roth and Murnighan (1978), Palfrey and Rosenthal (1994), Dal Bo (2005), Dal Bo and Frechette (2011), Aoyagi and Frechette (2009), and Duffy and Ochs (2009)).

Arifovic, McKelvey, and Pevnitskaya (2006) and Chmura, Goerg, and Selten (2011) are the closest papers to ours. Both papers are comparative studies of learning theories in repeated game environments. Arifovic, McKelvey, and Pevnitskaya (2006) study repeated games without rematch (i.e., what is commonly known as a repeated game), while Chmura, Goerg, and Selten (2011) study repeated interactions with a random matching protocol. Since we are interested in repeated game effects, we do not rematch subjects after each round. Moreover, Arifovic, McKelvey, and Pevnitskaya (2006) have a broader than comparative objective, and hence their analysis features four learning models. We provide a full-blown comparative analysis: we develop a measure of predictive power and rank ten theories, including equilibrium and learning theories and an axiomatic approach.

There have been extensive comparative studies based on predictive power, but these studies have compared generalizations of expected utility theory (e.g., Harless and Camerer (1994) and Hey and Orme (1994)) or solution concepts in static games (e.g., Selten and Chmura (2008) and Brunner, Camerer, and Goeree (2011)).

2 Measures of Predictive Power

In this section, we axiomatize a measure of predictive power and compare it to Selten (1991)'s measure.

Let $\mathcal{G}$ be the set of all finite normal-form games. For any game $G \in \mathcal{G}$, let $O(G) = \{1, \ldots, n\}$ or simply O be the set of all observable outcomes. For example, (a discretized version of) the set of feasible payoffs, the set of frequencies of play, or the set of sequences of action profiles can be chosen as sets of observable outcomes. The reason why O is assumed to be finite is that there can only be finitely many experimental observations; therefore, theories that make infinitely many predictions will be at a disadvantage (by making infinitely many false predictions) unless we group their predictions into finitely many cells.

According to Selten (1991), we can distinguish three types of theories. First, a *point theory* is a function τ that assigns to each G a point $\tau(G) \in O$. Second, an *area theory* is a function τ that assigns to each G a set $\tau(G) \subset O$. Third, a *distribution theory* is a function τ that assigns to each G a probability distribution $\tau(G)$ on O. Distribution theories are more complete and informative.

Suppose that experimental data are available. For any $G \in \mathcal{G}$, let $E(G) \subset O$ be the set of outcomes that have been observed in the experiments for game G. For each $x \in E(G)$, let $f(x)$ be the relative frequency with which x has been observed in the experiments. Let $E^*(G) = (f(1), \ldots, f(n))$ be the true distribution of outcomes, i.e., the vector of observed frequencies gathered from experiments.

2.1 Area Theories

In this section, we derive a measure that allows us to compare area theories according to predictive power. It is important to recognize that ranking theories is equivalent to developing a set-based preference ordering. This topic has been studied extensively in the literature (see Barbera, Bossert, and Pattanaik (2004) for an excellent survey).

Our objective is to characterize a binary relation $\succeq$ by a function $\mu : 2^O \to \mathbb{R}$ such that given G and $E(G)$,

$$\tau_1(G) \succeq \tau_2(G) \Leftrightarrow \mu(\tau_1(G)) \geq \mu(\tau_2(G)) \tag{1}$$

for any $\tau_1(G)$ and $\tau_2(G)$. We propose several axioms.

Axiom 1. *(Success Monotonicity)*
If $\tau_1(G) = \tau_2(G) \cup \{x\}$ *for* $x \in E(G) \cap \tau_2(G)^c$*, then* $\tau_1(G) \succ \tau_2(G)$*.*

Axiom 2. *(Error Monotonicity)*
If $\tau_1(G) = \tau_2(G) \cup \{x\}$ *for* $x \in E(G)^c \cap \tau_2(G)^c$*, then* $\tau_2(G) \succ \tau_1(G)$*.*

Axiom 3. *(Equal Treatment of Errors)*
For any distinct $x, y \notin E(G)$*, if* $\{x, y\} \subset \tau(G)$*, then* $\tau(G)\backslash\{x\} \sim \tau(G)\backslash\{y\}$*.*

The first three axioms are rather innocuous. The monotonicity axioms say that adding a correct prediction increases predictive power, while adding an incorrect prediction decreases it. Not all correct predictions appear with the same frequency, and hence carry the same positive weight, which suggests an obvious way to compare correct predictions. However, it is not obvious to quantify how bad an error, i.e. an incorrect prediction is, unless we consider the distance between an error and a correct guess to be a meaningful indicator. Axiom 3 considers all errors to be equally bad.

Let $(\tau_1(G), \ldots, \tau_k(G)) =^* (\tau_1'(G), \ldots, \tau_k'(G))$ mean that each $\tau_i(G)$ and $\tau_i'(G)$ are in 2^O, and for each $x \in O$, the number of $\tau_i(G)$'s that contain x is equal to the number of $\tau_i'(G)$'s that contain x.

Axiom 4. *(Transferability) For all* k*,* $\tau_i(G)$ *and* $\tau_i'(G)$*, if* $(\tau_1(G), \ldots, \tau_k(G)) =^* (\tau_1'(G), \ldots, \tau_k'(G))$ *and* $\tau_i(G) \succeq \tau_i'(G)$ *for all* $i < k$*, then* $\tau_k(G) \succ \tau_k'(G)$ *does not hold.*

This axiom can be found in Fishburn (1986) in the context of subjective probabilities, but it also has meaning our context. Start with any theories $\{\tau_1(G), \ldots, \tau_k(G)\}$ and re-arrange the errors and the correct predictions across theories 1 to $k-1$ in a way that improves each of them at least weakly. When carrying out these "trades" across theories 1 to $k-1$, we can produce unassigned objects. For example, if each theory contains at least one error, then a trivial way of improving their predictive power is to remove these errors (assuming the relation agrees with Axiom 2) so that we end up with a set of unassigned objects consisting of one error for each theory. The axiom says that if these unassigned x's are added to $\tau_k(G)$, then it cannot improve the predictive power of that theory. In short, if we trade errors and correct predictions across theories in a way that does not necessarily balance trades but improves the predictive power of each theory, then the unassigned elements, when added to another theory (here $\tau_k(G)$), necessarily decrease its predictive power. This is represented in Figure 2.1. These transfers leave one error and one correct prediction unassigned. Assuming that τ_1 and τ_2 are improved by these transfers, the set of unassigned elements must have a negative impact on predictive power, regardless of the theory to which they are added.

Proposition 5. *Suppose that Axioms 1-4 hold. Then there are real numbers* $\{p_x\}_{x \in O}$ *such that*

$$\tau_1(G) \succeq \tau_2(G) \Leftrightarrow \sum_{x \in \tau_1(G)} p_x \geq \sum_{x \in \tau_2(G)} p_x \tag{2}$$

where for each $x \in \tau(G)$*,* $p_x > 0$ *if* $x \in E(G)$ *and* $p_x = -\kappa < 0$ *if* $x \notin E(G)$*.*

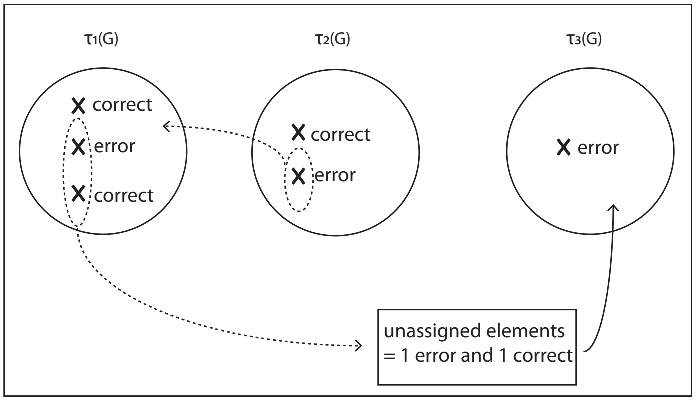


Figure 1: The Transferability Axiom (The dashed lines represent transfers across theories: one error is transferred from τ_2 to τ_1; one error and one correct prediction are transferred from τ_1 to the unassigned set).

This result immediately implies the following simple measure of predictive power

$$\mu^*(\tau(G)) = \sum_{x\in E(G)\cap\tau(G)} p_x - \kappa|E(G)^c \cap \tau(G)|. \tag{3}$$

This formulation is interesting for at least two reasons. First, it has a nice interpretation. Correct and incorrect predictions are clearly separated: every correct prediction carries a value or utility and every error carries a cost or disutility.

Second, this formulation leaves considerable freedom in the definition of p_x and κ. Since p_x represents the valued added by a correct prediction, the frequency $f(x)$ with which x is observed in the experiments should be a reasonable choice. Therefore, we set $p_x = f(x)$. As for the cost associated with an incorrect prediction, several options are possible. For example, we can choose $\kappa = \min_{x\in E(G)} f(x)$, $\kappa = 1/|E(G)|$, or $\kappa = \max_{x\in E(G)} f(x)$. The first and the last option are extreme choices for which a mistake is either nearly costless (any correct prediction would compensate for an error) or extremely costly (not even the most frequently observed prediction would compensate for an error). We make an intermediate choice, $\kappa = 1/|E(G)|$,[5] as it seems reasonable.

Selten (1991, 1998) has axiomatized several measures of predictive power that have been used in the applied literature, one of which concerns area theories (see Beatty and Crawford (2011) for an application of this measure). Selten's measure is defined as

$$\mu_S(\tau(G)) = \sum_{x\in E(G)\cap\tau(G)} f(x) - \frac{|\tau(G)|}{|O|} \tag{4}$$

[5]Note that $\max_{x\in E(G)} f(x) \geq 1/|E(G)| \geq \min_{x\in E(G)} f(x)$.

which is equivalent to

$$\mu_S(\tau(G)) = \sum_{x \in \tau(G)} \left(f(x) - \frac{1}{|O|} \right) \tag{5}$$

On the rhs of (4), the first expression is usually denoted by $r(\tau)$ and called the *hit rate* of theory τ. The negative part is usually denoted by $a(\tau)$ and called the *accuracy* of theory τ. Define by $\succeq_S$ the binary relation induced by μ_S. It is clear from (5) that $\succeq_S$ satisfies strong additivity, equal treatment of errors, and error monotonicity. But it is also clear that Selten's measure can violate success monotonicity. Adding a correct prediction x with a small $f(x)$ can damage predictive power.

Our main disagreement with Selten's measure is about accuracy. In Selten's measure, accuracy has nothing to do with errors and take the size of the predicted set into account. In the unfortunate scenario where many outcomes appear in an experiment, a theory that predicts most of these outcomes is penalized for it, as its size would necessarily be large. In our opinion, only errors are relevant and not the size per se.

2.2 Distribution Theories

In this section, we describe a measure of predictive power for distribution theories. For any game $G \in \mathcal{G}$, a distribution theory $\tau(G) = (p_1(\tau), \ldots, p_n(\tau))$ where $p_x(\tau)$ is the frequency with which x is predicted to occur by τ.

The binary relation $\succeq$ that we use to rank distribution theories is characterized by a measure of predictive power μ. Given G and $E^*(G)$,

$$\tau_1(G) \succeq \tau_2(G) \Leftrightarrow \mu(\tau_1(G)) \leq \mu(\tau_2(G)) \tag{6}$$

for any $\tau_1(G)$ and $\tau_2(G)$, where

$$\mu(\tau(G)) = \sum_{x=1}^{n} (p_x(\tau_1) - f(x))^2$$

is the squared of the Euclidean distance between $\tau(G)$ and the true distribution. This choice not obvious a priori. We refer the reader to Selten (1998) and Selten and Chmura (2008) for further detail. The first paper provides axioms that characterize this measure and the second uses it in an experimental study of static games. Selten (1998)'s axiomatic analysis is concerned with scoring rules, which evaluate predictive power by associating a score to every component of a distribution theory. The axioms result in the characterization of the quadratic scoring rule. If we define the expected score loss as the difference between the score of the true distribution and that of a theory, then our measure μ is the expected score loss of the theory under the quadratic scoring rule.

3 The Theories

In this paper, we compare the predictive power of ten theories: standard equilibrium theory, equilibrium theory with automaton strategies and cost of complexity (Abreu and Rubinstein (1988)), weak renegotiation-proofness (Farrell and Maskin (1989) and Bernheim and Ray (1989)), internal renegotiation-proofness (Ray (1994)), Nash equilibrium with tests (Spiegler (2005)), the axiomatic approach of repeated interactions (Mathevet (2014)), the self-tuning Experience-Weighted Attraction learning model (Camerer, Ho, and Chong (2002)), modified EWA, weighted fictitious play, and reinforcement learning. Most of these theories are well-known in the literature, and so we will only describe the recent theories by and Mathevet (2014).

3.1 Theory Description

3.1.1 The Axiomatic Approach of Repeated Interactions

Mathevet (2014) proposes an axiomatic approach to study repeated interactions between two boundedly rational players. In general, a researcher can either model boundedly rational behaviors precisely, for example, by specifying learning algorithms, or she can model the implications or the outcomes of boundedly rational behaviors. The learning literature adopts the former approach. While this leads to detailed models of behavior, these models often have to be specific to remain tractable. In the end, the literature presents a plethora of models and might appear fragmented. Mathevet (2014) suggests to adopt an abstract approach. Instead of proposing a precise model of bounded rationality in repeated settings, he proposes a plausible recipe for less than rational play. In particular, he proposes a model-free approach, and the infinite sequences of action profiles are natural primitives for this exercise. He formulates three axioms on the infinite sequences of play: individual security, collective intelligence, and internal stability; and then he characterizes solutions.

One virtue of this approach is that it allows us to build a multifaceted picture of boundedly rational play by combining various independent principles. Since bounded rationality is by nature a multifaceted phenomenon, the axiomatic approach seems well-adapted. Moreover, the axiomatic approach makes it possible to explore the many ways to be less than rational within the same framework, by comparing different sets of axioms and their solutions

Main Result. Mathevet (2014)'s main solution payoffs contain all payoffs v such that: v is individually secure (i.e., individually rational); v must lie on a line segment connecting two Pareto unranked pure payoff profiles or v must be a pure payoff profile; v must lie on a line segment starting from the origin $(0, 0)$ and whose slope is determined by players' observed minmaxes (see Mathevet (2014)).

3.1.2 Strategic EWA

Ioannou and Romero (2014) extend the experience weighted attraction model to allow for repeated game strategies (two-state automata in particular). With the new model, they run simulations in four commonly studied two-by-two games. The simulations results show that extending the model to allow for repeated game strategies leads to some commonly observed outcomes that can't be obtained with action-based learning models. For example, the simulations of the strategic EWA model lead high levels of cooperation in the prisoners' dilemma, alternations between the two pure-strategy Nash equilibria in the battle of the sexes game, mutual concilliation in the chicken game, and the Pareto dominant outcome in a stag-hunt game.

4 Experimental Results and Comparative Analysis

4.1 Experimental Design

Experiments were run at the Vernon Smith Experimental Economics Laboratory at Purdue University. The subjects were drawn from a pool of Purdue University undergraduate students that were signed up to participate in experiments. Subjects interacted on computers using an interface that was programmed with the z-Tree software (Fischbacher, 2007) (See appendix for screenshots). Upon entering the lab, subjects were randomly assigned to a computer and given a handout containing instructions. Once all subjects had been seated, the instructions were read aloud as the subjects read along. After the instructions had been read, each subject completed a quiz to make sure that they understood the instructions (See appendix for instructions and quiz). The experiment did not begin until all subjects had successfully answered all questions on the quiz. Each session, including the instructions, lasted about one or two hours (depending on number of periods). During the experiment, all payoffs were displayed in experimental Francs, and the exchange rate to U.S. dollars was announced during the instructions. Subjects' final payoff was the sum of earning of all rounds from the experiment. After the experiment, subjects privately received their payments in cash, with the average payoff totaling 16.21$.

There were a total of 16 sessions consisting of 314 subjects. In each session, subjects played a total of three supergames[6]. At the beginning of each supergame, each subject was randomly paired with one other subject and they played a 2×2 game repeatedly with the same partner. The experimental interface was setup to allow each pair to progress at their own rate so subjects only had to wait for all other pairs to finish at the end of each supergame rather than the end of each round. After all pairs had finished their supergames, subjects were rematched with a new partner that they had not been matched with before, and played a game that they had not played before. The length of the supergame was determined using the following termination rules[7]:

- **T30** - 30 rounds were played with certainty, and then starting in the 31^{st} round, the continuation probability 0.9, so the expected length of each supergame was 40.
- **T0** - Starting in the first round, the continuation probability was 0.99, so the expected length of each supergame was 100.

[6] Subjects never played the same game twice, and were never matched with another subject twice.

[7] The goal was to have long lasting supergames. One way to have a long lasting supergame is to ensure that a minimum of T periods are played (first case). Another way is to have a high continuation probability (second case).

	A	B
A	3,3	1,4
B	4,1	2,2

Prisoners' Dilemma

	A	B
A	1,1	2,4
B	4,2	1,1

Battle of the Sexes

	A	B
A	3,3	0,2
B	2,0	1,1

Stag Hunt

	A	B
A	3,3	1,4
B	4,1	0,0

Chicken

	A	B
A	3,3	1,2
B	2,1	0,0

Common Interest

	A	B
A	2,2	2,2
B	5,3	3,5

Samaritan's Dilemma

	A	B
A	2,2	2,2
B	3,1	0,0

Ultimatum

	A	B
A	4,1	1,2
B	1,2	2,1

Unique Mixed

Figure 2: 2×2 Games Tested

In each round, subjects were asked to select a choice (Y or W) as well as a guess as to what their partner would chose. For every correct guess a subject made they earned a raffle ticket. At the end of each supergame, one raffle ticket was randomly selected, and that subject received a bonus of $1. The payoffs were displayed in a table (as seen in the screenshots), and players roles were always the same (i.e. they didn't have to switch between row and column player).[8] A total of eight 2×2 games were tested. The games tested along with the number of observation pairs collected are displayed in Figure 2.[9] The A summary of the data collected is provided in Figure 3.

4.2 Experimental Results

To get an idea what the data looks like, Figure 4 provides a heat map for each of the eight game displaying the relative frequency of each possible payoff over the final 20 periods of the supergame. Based on these heat maps, it is clear that although the data is noisy, there are some clear regularities. In the prisoners' dilemma the most common outcomes are mutual cooperation and mutual defection. In the battle of the sexes, the most common outcome is alternations between the two pure strategy

[8] The choices were also randomized across supergames meaning that for example, in some supergames Y was the Nash equilibrium of a Prisoner's Dilemma and in others W was the Nash equilibrium.

[9] These games were chosen because these are some of the most commonly studied games in the game theory literature.

Game	Total	T30	T0
Prisoner's Dilemma	70	33	37
Battle of the Sexes	70	31	39
Stag Hunt	50	30	20
Chicken	70	32	38
Common Interest	18	18	0
Samaritan's Dilemma	60	0	60
Ultimatum	60	0	60
Unique Mixed	36	0	36

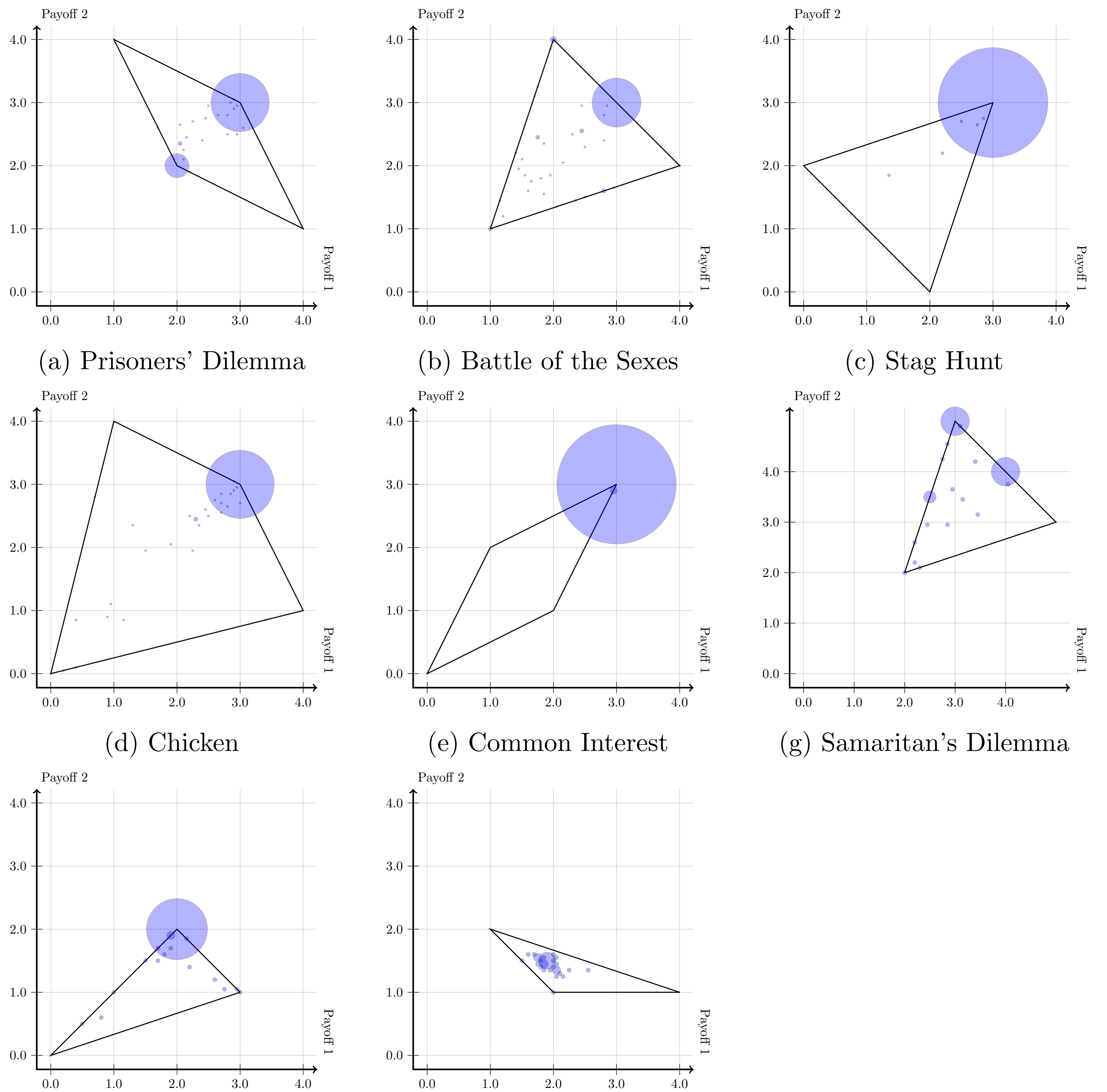
Figure 3: Summary of Data

Nash equilibria, though a few pairs converge to one of the two pure-strategy Nash equilibria. In the stag-hunt, common interest, chicken, and ultimatum games, players tend to converge upon a single point. In the game with the unique mixed strategy Nash equilibrium, players tend to play close to the mixed Nash equilibrium. Finally, in the Samaritan's dilemma, players tend to play either the Nash equilibrium or alternate between payoffs of $(3, 5)$ and $(5, 3)$. Despite the noise, the regularity of the data is promising suggesting that some outcomes are more likely than others.

The next step is to take the measures developed in Section 2 and the predictions made in Section 3 and combine them into a ranking of the theories for each game. Before computing the measures however, several difficulties need to be cleared up. First, this paper analyzes two classes of theories: area theories (which predict an area in $\mathbb{R}^2$) and distribution theories (which predict a distribution over points in $\mathbb{R}^2$). Section 2 describes a measure that can be used for each type of theory, however, these measures are not directly comparable. Since the goal of the papers is to compare the predictability of all theories, we need some way to describe an area theory as a distribution theory, and vice versa. To describe an area theory as a distribution theory, we assume that the area theory predicts equal weight on all points. To describe a distribution theory as an area theory, we assume that all points that are predicted with positive probability are predicted in the corresponding area theory. These transformations allow us to compare all of the theories as area theories using the area measure, as well as compare all theories as distribution theories using the distribution measure.

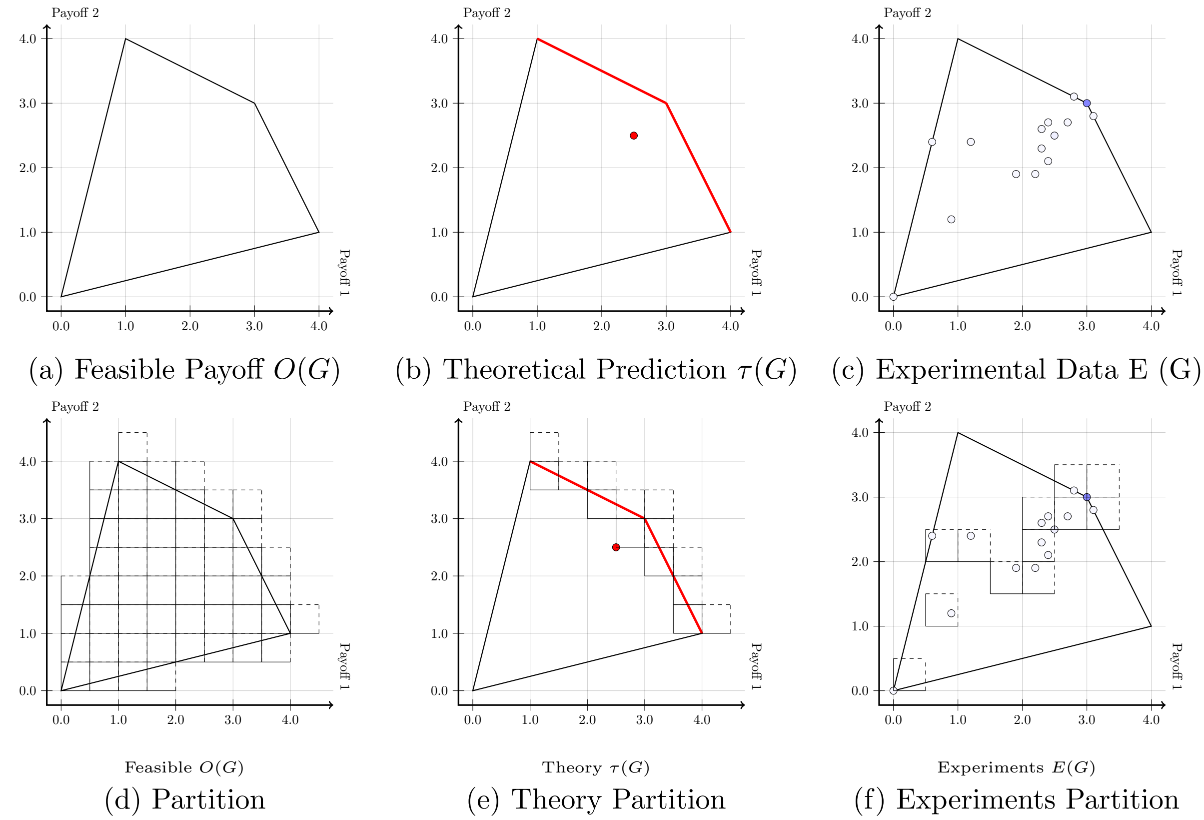
The next difficulty comes in determining how to compare the theory and the data. For example, what constitutes an error? Previous measures of predictability Selten (1991, 1998) have looked at discrete outcome spaces. Discrete outcomes spaces have the desirable property that a prediction is either correct or incorrect. In a continuous outcome space (as the payoff space in $\mathbb{R}^2$ we are examining), a prediction may be close to a payoff but not exactly the same.[10] To solve this problem,

[10]For example, in the last 20 periods of a prisoners' dilemma game in Figure 2 players may play the cooperative



(a) Prisoners' Dilemma (b) Battle of the Sexes (c) Stag Hunt (d) Chicken (e) Common Interest (g) Samaritan's Dilemma (h) Ultimatum (i) Unique Mixed

Figure 4: Experimental data showing average payoffs over the last twenty rounds of a supergame (data combined for both **T0** and **T30**).



(a) Feasible Payoff $O(G)$ (b) Theoretical Prediction $\tau(G)$ (c) Experimental Data E (G)

(d) Partition (e) Theory Partition (f) Experiments Partition

Figure 5: This shows an example of the partition technique used for $\varepsilon = 0.5$. Panel (a) shows the set of feasible payoffs $O(G)$ for the chicken game. Panel (b) shows a set of theoretical predictions. Panel (c) shows the experimental data $E(G)$ (though frequencies can't be determined from plot). Panel (d)-(f) show the corresponding partitions. To calculate the measure for this theory given the data, notice that there are 29 total experimental points, 19 of which overlap with the theory. In addition there are a total of 12 squares that have experimental points contained in them, and a total of 10 squares that the theory predicts with no experimental points in them. Therefore, the measure for this theory would be $\mu(\tau) = \frac{19}{29} - \frac{10}{12}$.

$\mathbb{R}^2$ is partitioned into squares of width ε (open on the top and right, and closed on the bottom and left)[11]. Then any prediction or observation inside a square is considered to be equivalent. The smaller ε, the finer the partition of $\mathbb{R}^2$, the more accurate the predictions have to be. An example of how the calculation is made using the partition is given in Figure 5.

This partition allows us to calculate the measure for each value of epsilon for each game. Plots of rankings for various values of ε are displayed in Figures 6-13. For the area measure, the higher the value, the better the prediction. For the distribution measure, the lower the distance, the better the prediction.

Once we have developed the ranking for each game, then next step is to aggregate the individual game rankings into a single aggregate ranking across all games. There are two ways that this can

outcome every period except one, which would lead to a payoff close but not equal to $(3, 3)$.

[11]The partition starts at the feasible payoff that is closest to the origin.

Rank	Theory	Score	PD	BO	SH	CH	OB	SD	UL	MX
1	axiomatic	68	2	1	1	2	2	8	2	2
2	strategicEWA	61	1	7	8	1	1	7	1	1
3	internalRenegotiatio	55	7	5	1	4	2	1	6	7
4	reinforcementErevRo	51	8	8	1	5	2	3	8	2
5	actionSampling	50	8	8	1	5	2	3	9	2
6	weightedFP	50	8	8	1	5	2	3	9	2
7	weakRenegotiationPro	50	4	2	9	9	2	2	3	7
8	selfTuningEWA	50	8	8	1	5	2	3	9	2
9	Spiegler	38	4	2	10	9	2	10	3	10
10	abreuRubinstein	35	3	5	7	3	10	9	6	10
11	folkTheorem	32	4	2	10	9	11	10	3	7

Table 1: Aggregate Rankings for Area Measure with $\varepsilon = 0.5$ using the Borda Count.

be done. First, for a given value of ε, we can rank the theories for all games. Let T be the total number of theories, and $r^j(\tau)$ be the rank of theory τ in game j. Then we can use the Borda count to give each theory a score based on,

$$S(\tau) = \sum_j \left(T + 1 - r^j(\tau)\right)$$

Therefore, if a given theory has the highest rank for game j then it receives a count of T, if it has the second highest rank it receives a count of $T-1$, all the way down to the lowest rank receiving a count of 1. Once the scores are summed up over all games, then we can use those to rank the theories. The results for the ranking using the Borda count are displayed in Tables 1 and 2 for the area and distance measures respectively.

The second way that we can rank the theories is to assume that the measures are cardinal measures, and simply sum up the measure across each game. In this case, let $\mu^j(\tau)$ be the measure for theory τ in game j.

$$S(\tau) = \sum_j \mu^j(\tau)$$

The score is again the sum over all games. In this case, the ranking is determined by the theory with the highest summed score over all games for the area measure, and lowest summed score over all games with the distance measure. The results for the cardinal aggregate measure are displayed in Tables 3 and 4.

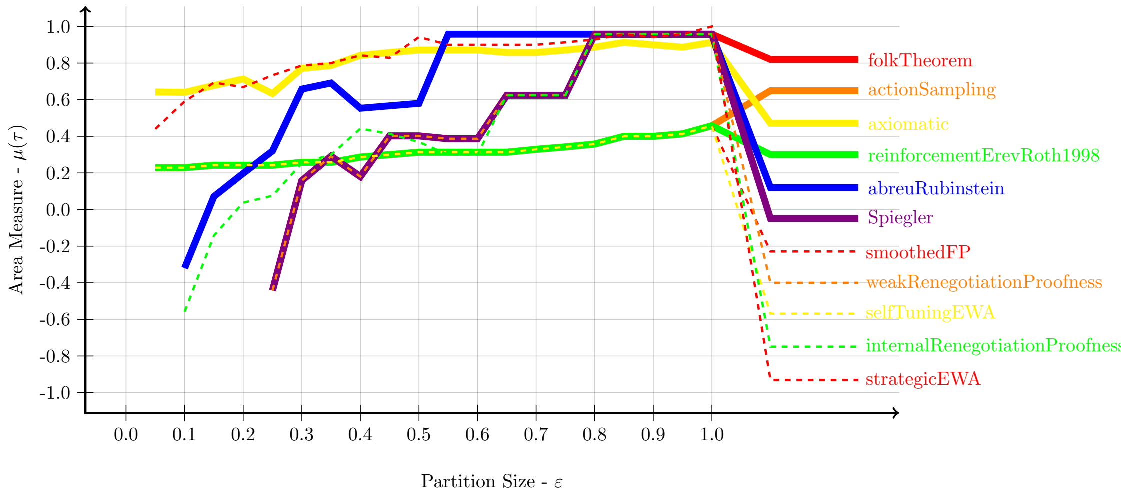
Examining Tables 1-4, we can see that the axiomatic model (Mathevet (2014)) and the strategic EWA model (Ioannou and Romero (2014)) are consistently near the top of the rankings regardless of which ranking is used, or how the rankings are aggregated.

Rank	Theory	Score	PD	BO	SH	CH	OB	SD	UL	MX
1	weakRenegotiationPro	63	4	1	9	4	1	1	4	1
2	axiomatic	62	1	6	3	3	1	4	3	5
3	strategicEWA	56	2	7	1	1	9	7	1	4
4	internalRenegotiatio	53	7	4	3	7	1	5	7	1
5	Spiegler	52	4	1	10	4	1	2	4	10
6	folkTheorem	51	4	1	10	4	11	2	4	1
7	reinforcementErevRo	45	8	8	3	8	1	8	2	5
8	selfTuningEWA	39	8	8	2	8	1	8	9	5
9	actionSampling	38	8	8	3	8	1	8	9	5
10	abreuRubinstein	38	3	4	8	2	10	6	7	10
11	weightedFP	38	8	8	3	8	1	8	9	5

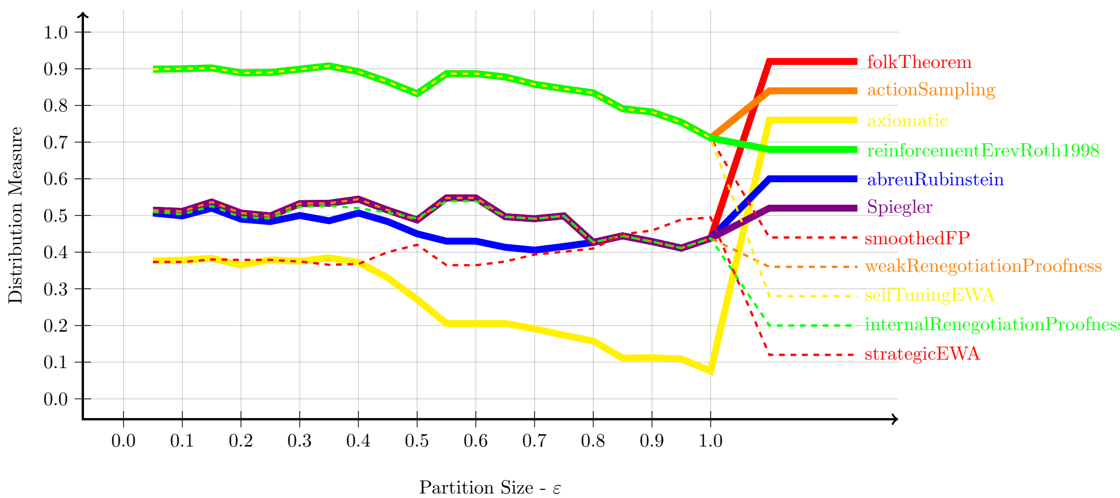
Table 2: Aggregate Rankings for Distance Measure with $\varepsilon = 0.5$ using the Borda Count.

Rank	Theory	Score	PD	BO	SH	CH	OB	SD	UL	MX
1	strategicEWA	4.65	0.94	0.10	0.61	0.49	1.00	0.19	0.74	0.58
2	axiomatic	4.48	0.87	0.46	0.88	0.13	0.94	0.15	0.67	0.39
3	internalRenegotiatio	3.51	0.37	0.10	0.88	-0.00	0.94	0.36	0.58	0.27
4	reinforcementErevRo	3.25	0.31	0.09	0.88	-0.12	0.94	0.20	0.55	0.39
5	actionSampling	2.71	0.31	0.09	0.88	-0.12	0.94	0.20	0.02	0.39
6	weightedFP	2.71	0.31	0.09	0.88	-0.12	0.94	0.20	0.02	0.39
7	selfTuningEWA	2.71	0.31	0.09	0.88	-0.12	0.94	0.20	0.02	0.39
8	weakRenegotiationPro	2.25	0.40	0.22	-0.08	-0.50	0.94	0.35	0.63	0.27
9	abreuRubinstein	0.37	0.58	0.10	0.79	0.08	0.50	-0.26	0.58	-2.00
10	folkTheorem	-0.56	0.40	0.22	-0.69	-0.50	-0.50	-0.40	0.63	0.27
11	Spiegler	-1.39	0.40	0.22	-0.69	-0.50	0.94	-0.40	0.63	-2.00

Table 3: Aggregate Rankings for Area Measure with $\varepsilon = 0.5$ summing values.

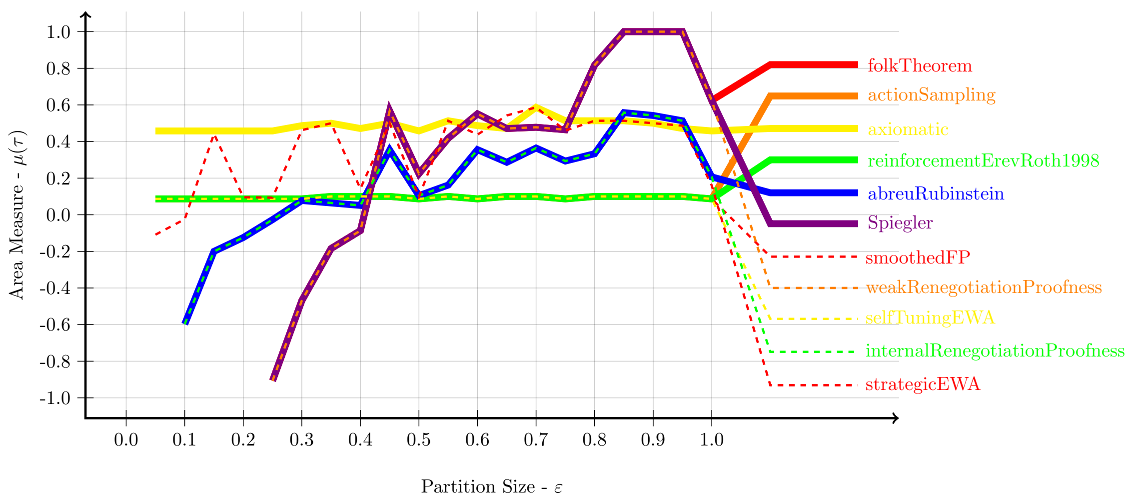


(a) Predictive Measure m as a function of ε.

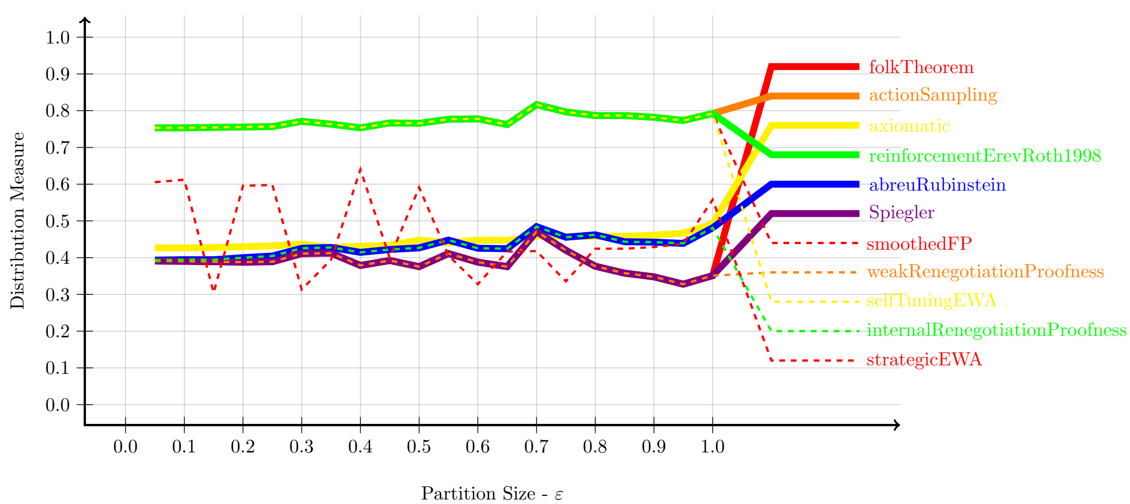


(b) Distance between theory and experiments as a function of ε.

Figure 6: Comparison of theories using predictive measure m and distance measure for the Prisoners' Dilemma.

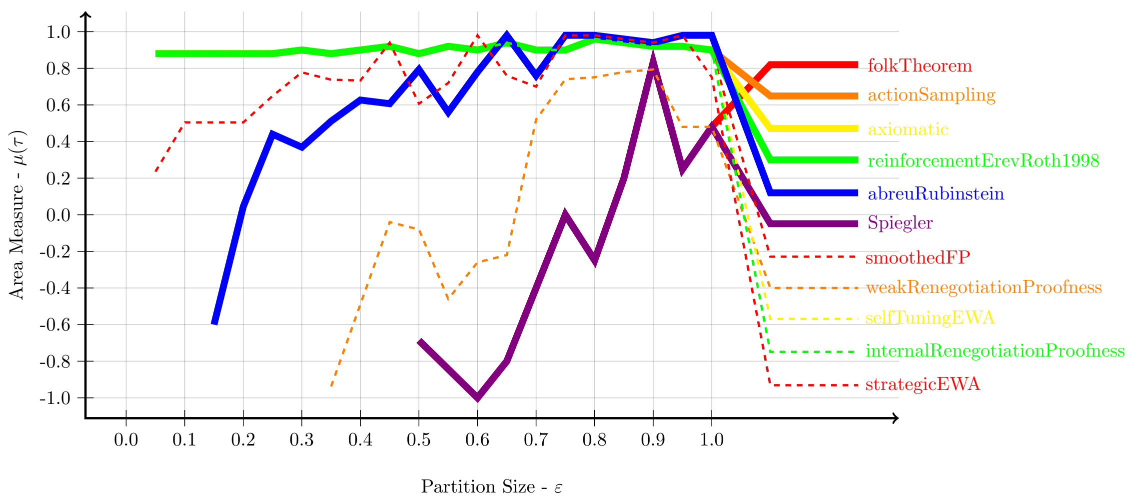


(a) Predictive Measure m as a function of ε.

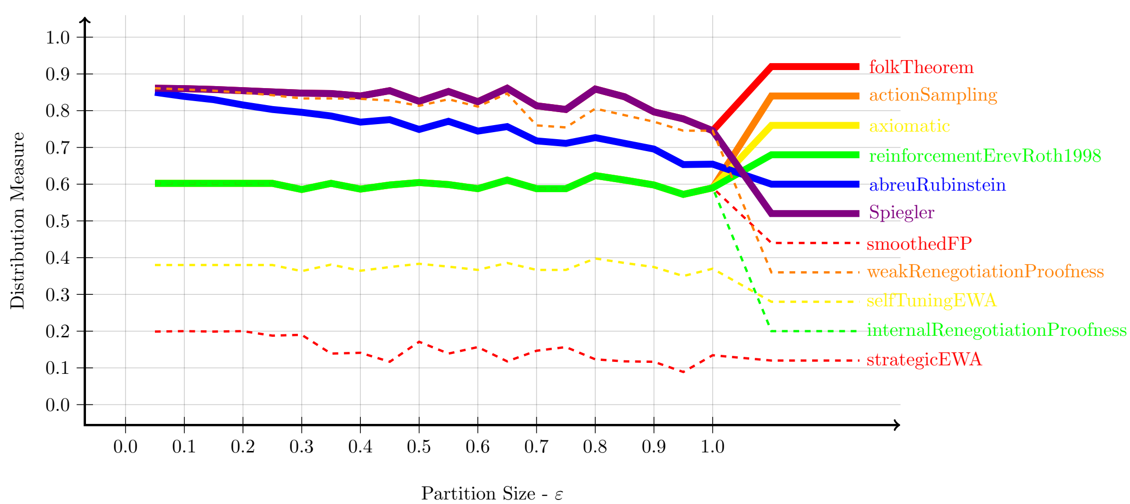


(b) Distance between theory and experiments as a function of ε.

Figure 7: Comparison of theories using predictive measure m and distance measure for the Battle of the Sexes.

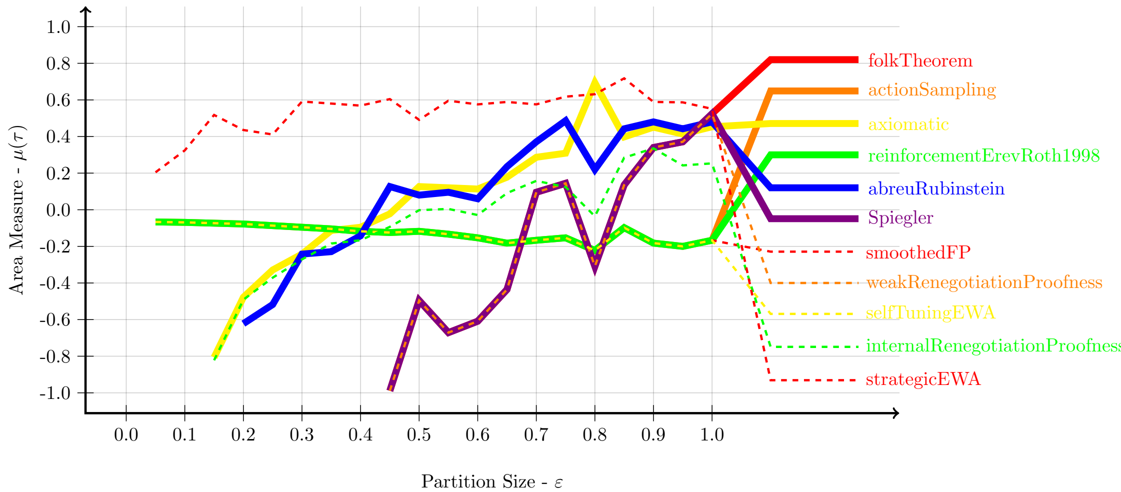


(a) Predictive Measure m as a function of ε.

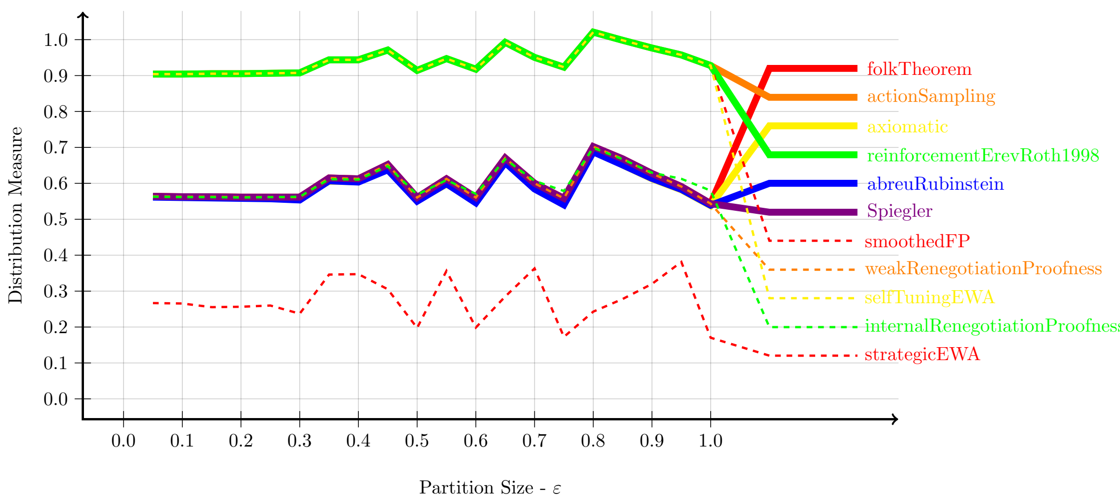


(b) Distance between theory and experiments as a function of ε.

Figure 8: Comparison of theories using predictive measure m and distance measure for the Stag-Hunt Game.

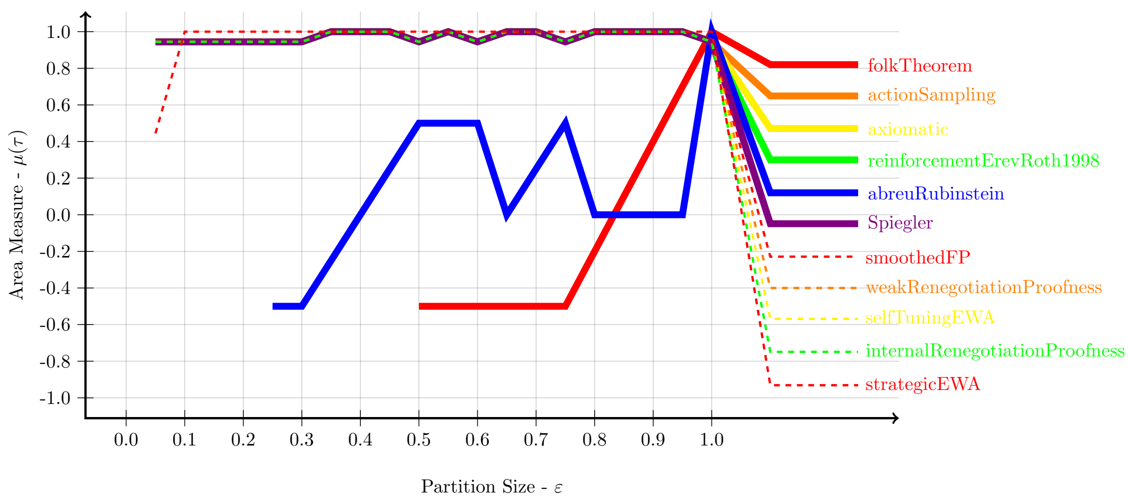


(a) Predictive Measure m as a function of ε.

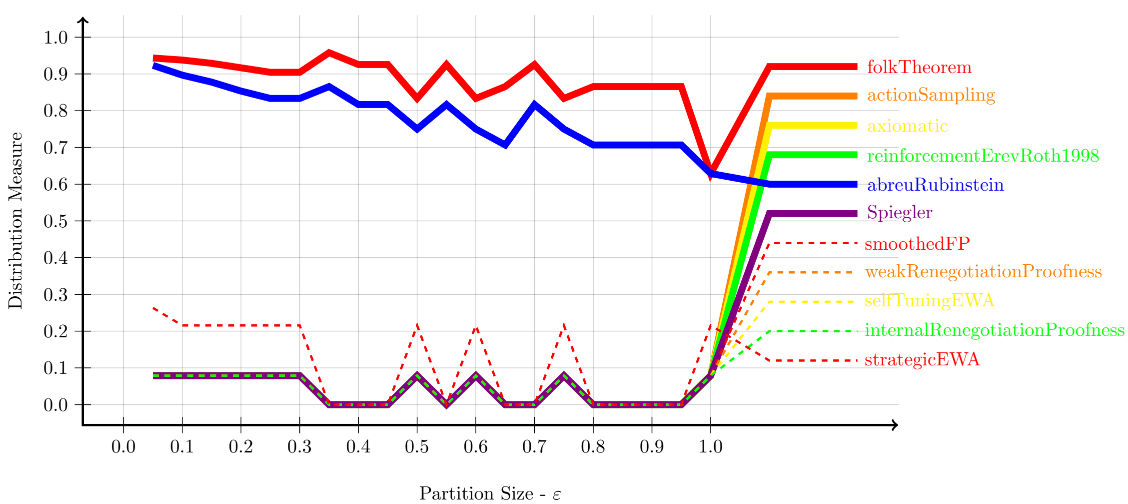


(b) Distance between theory and experiments as a function of ε.

Figure 9: Comparison of theories using predictive measure m and distance measure for the Chicken Game.

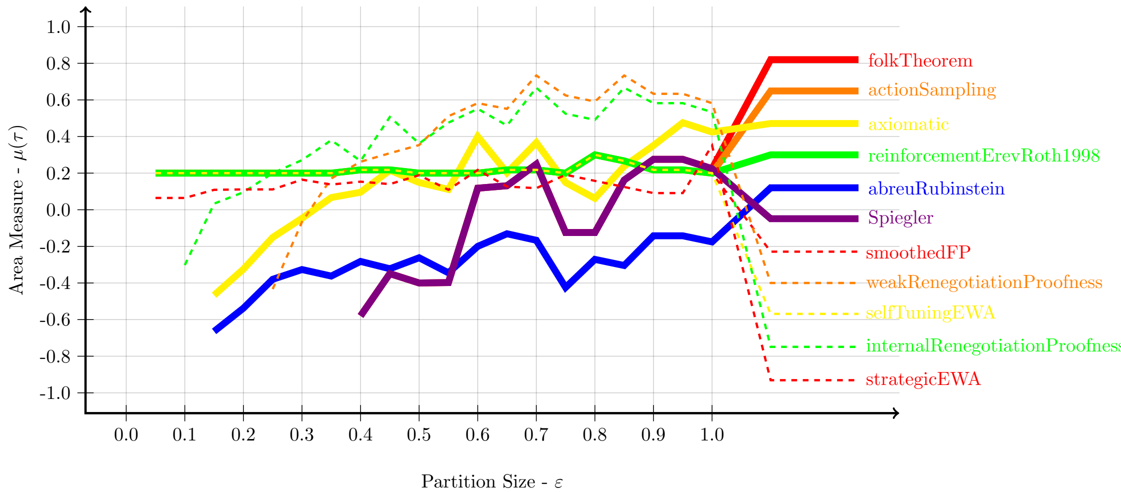


(a) Predictive Measure m as a function of ε.

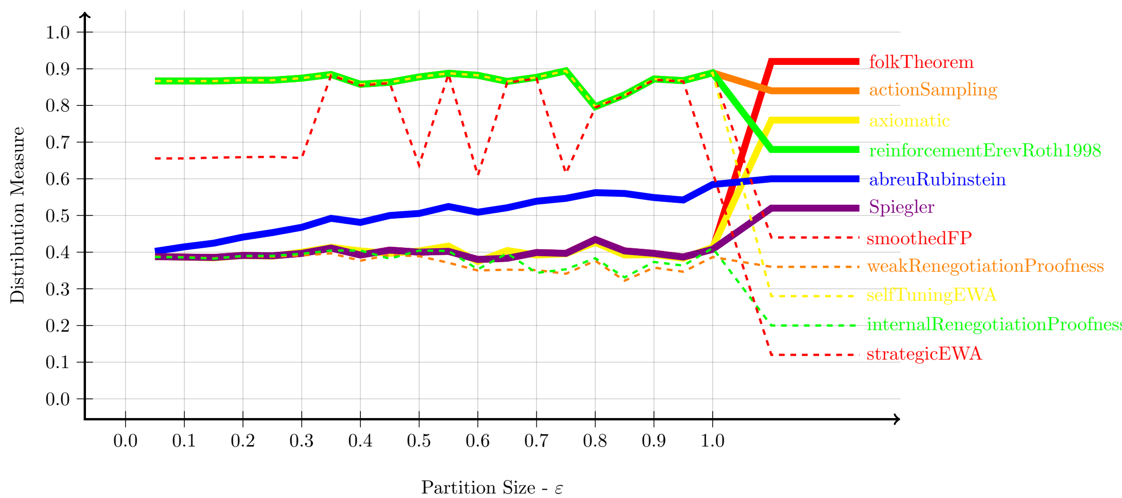


(b) Distance between theory and experiments as a function of ε.

Figure 10: Comparison of theories using predictive measure m and distance measure for the Common Interest Game.

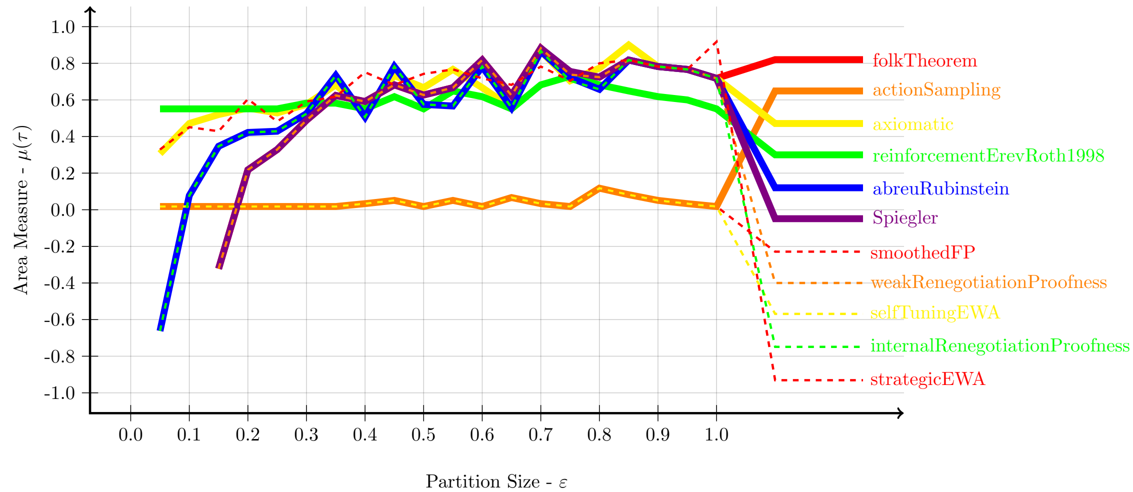


(a) Predictive Measure m as a function of ε.

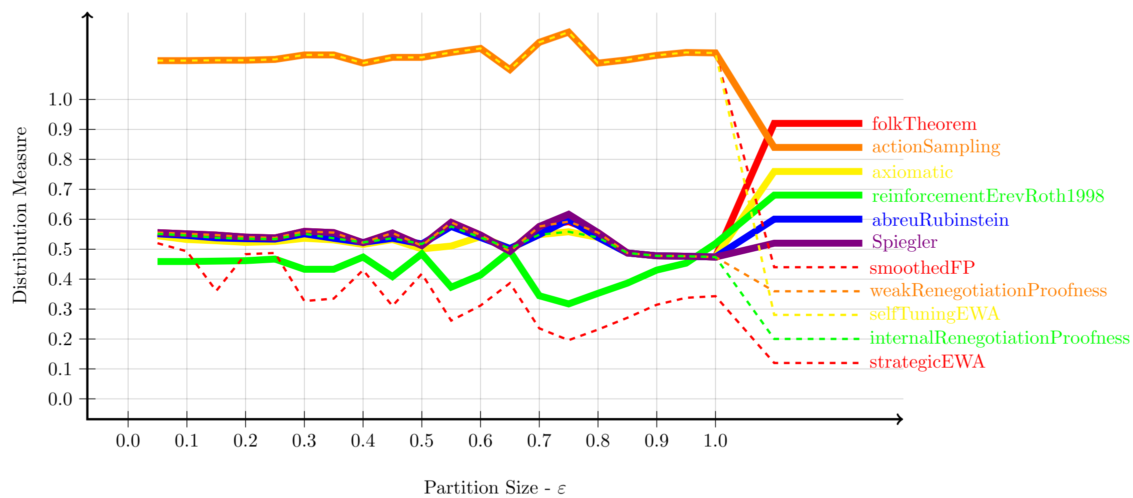


(b) Distance between theory and experiments as a function of ε.

Figure 11: Comparison of theories using predictive measure m and distance measure for the Samaritan's Dilemma.

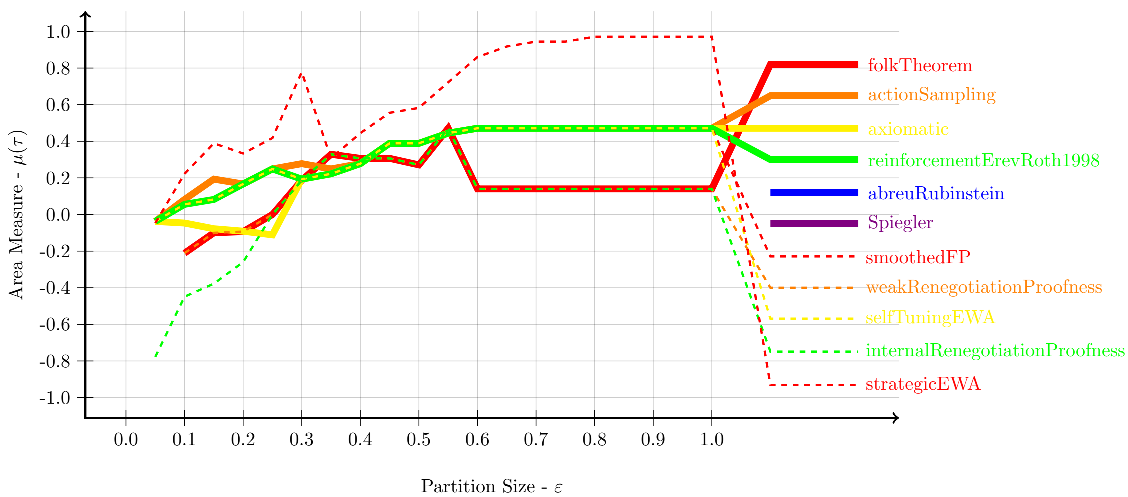


(a) Predictive Measure m as a function of ε.

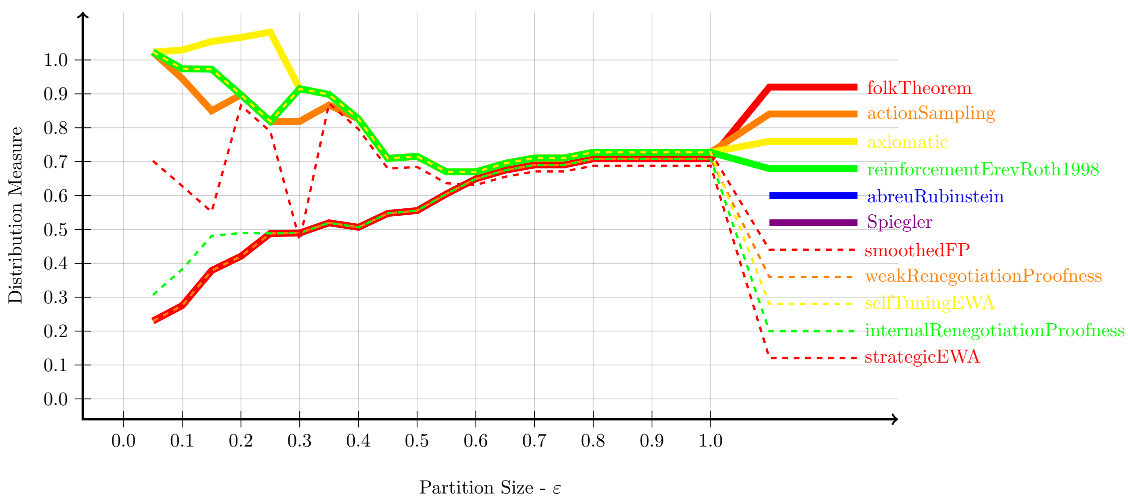


(b) Distance between theory and experiments as a function of ε.

Figure 12: Comparison of theories using predictive measure m and distance measure for the Ultimatum Game.



(a) Predictive Measure m as a function of ε.



(b) Distance between theory and experiments as a function of ε.

Figure 13: Comparison of theories using predictive measure m and distance measure for the Unique Mixed Game.

Rank	Theory	Score	PD	BO	SH	CH	OB	SD	UL	MX
1	strategicEWA	3.33	0.42	0.59	0.17	0.20	0.22	0.64	0.42	0.68
2	axiomatic	3.57	0.27	0.45	0.60	0.55	0.08	0.40	0.50	0.72
3	internalRenegotiatio	3.65	0.49	0.43	0.60	0.57	0.08	0.40	0.52	0.56
4	weakRenegotiationPro	3.77	0.49	0.38	0.81	0.56	0.08	0.39	0.51	0.56
5	folkTheorem	4.55	0.49	0.38	0.83	0.56	0.83	0.40	0.51	0.56
6	reinforcementErevRo	5.27	0.83	0.77	0.60	0.91	0.08	0.88	0.48	0.72
7	selfTuningEWA	5.71	0.83	0.77	0.38	0.91	0.08	0.88	1.14	0.72
8	actionSampling	5.93	0.83	0.77	0.60	0.91	0.08	0.88	1.14	0.72
9	weightedFP	5.93	0.83	0.77	0.60	0.91	0.08	0.88	1.14	0.72
10	Spiegler	6.24	0.49	0.38	0.83	0.56	0.08	0.40	0.51	3.00
11	abreuRubinstein	6.95	0.45	0.43	0.75	0.55	0.75	0.51	0.52	3.00

Table 4: Aggregate Rankings for Distance Measure with $\varepsilon = 0.5$ summing values.

5 Conclusion

This paper started with a simple question: What model has the most predictive power over a class of 2×2 repeated games? To address this question, we first characterized a measure of predictive power based on four axioms. Then we ran experiments over eight 2×2 repeated games and calculated predictions from 11 models. Finally, we used our model of predictive power to rank the theories using a variety of measures.

From this analysis, we can make several conclusions. First, the experimental data is relatively clean in the payoff space. Second, we find that Mathevet (2014)'s axiomatic approach, and Ioannou and Romero (2014)'s learning model perform well across both measures and both types of rankings (finishing first and second in all except one in which they are second and third). Again, it should be reiterated that though this does provides strong support for the predictability of these two models, it does not necessarily undermine the other models as predictability is not the primary concern of all of the models, and some of the models are not meant to be used for fixed matching repeated games. What this does suggest is further work trying create models that provide sharp predictions in these and other games to match what we see in the experimental lab.

A Proofs

Proof. It follows from Fishburn (1986, 1992) that Axiom 4 implies the existence of real numbers $\{p_x\}_{x\in O}$ such that

$$\tau_1(G) \succeq \tau_2(G) \Leftrightarrow \sum_{x\in\tau_1(G)} p_x \geq \sum_{x\in\tau_2(G)} p_x. \tag{7}$$

Axiom 1 implies that for any $x \in E(G)$ and any $\tau_2(G)$ such that $x \notin \tau_2(G)$,

$$\tau_2(G) \cup \{x\} \succ \tau_2(G)$$

and thus, by (7), $p_x > 0$. Similarly, Axiom 2 implies that for any $x \notin E(G)$ and any $\tau_2(G)$ such that $x \in \tau_2(G)$,

$$\tau_2(G) \backslash \{x\} \succ \tau_2(G)$$

and thus, by (7), $p_x < 0$. Finally, Axiom 3 implies that for any distinct $x, y \notin E(G)$ and any $\tau(G)$ such that $\{x, y\} \subset \tau(G)$,

$$\tau(G)\backslash\{x\} \sim \tau(G)\backslash\{y\}$$

and thus, by (7), $p_x = p_y$. $\square$

B The Twelve Theories

In this paper, we compare the predictive power of twelve theories: standard equilibrium theory, equilibrium theory under bounded rationality (Abreu and Rubinstein (1988)), weak renegotiation-proofness (Farrell and Maskin (1989) and Bernheim and Ray (1989)), internal renegotiation-proofness (Ray (1994)), Nash equilibrium with tests (Spiegler (2005)), axiomatic bargaining (Nash (1950) and Kalai and Smorodinsky (1975)), the axiomatic approach of repeated interactions (Mathevet (2014)), self-tuning Experience-Weighted Attraction (EWA) dynamics, modified EWA, weighted fictitious play, and reinforcement learning. In the interest of being self-contained, we first describe the essence of each of the twelve theories.

B.1 Theory Description

B.1.1 Equilibrium Theory and Bounded Rationality

The reader is assumed to be familiar with the basics of standard equilibrium theory in repeated games (see, e.g., Mailath and Samuelson (2006)). We will associate equilibrium theory with the use of subgame-perfect Nash equilibrium.

In Abreu and Rubinstein (1988), players play an infinitely repeated game, but they are boundedly rational. That is, they are only able to use strategies that are representable by finite automata. A finite automaton can be thought of as a set of (psychological) states with transitions between these states. The machine transitions from one state to another as a function of the opponent's action. There corresponds an action to each state. Moreover, there exists a small cost of maintaining each state. Therefore, a player prefers changing automaton if she can receive at least the same utility with another automaton having fewer states. Abreu and Rubinstein (1988) assume lexicographic preferences over automata: payoffs matter first and then the number of states.

Main Results. Standard folk theorems (e.g., Fudenberg and Maskin (1986)) establish that every individually rational payoff is supported as a subgame perfect equilibrium payoff, provided the discount factor is sufficiently close to one. Abreu and Rubinstein (1988) show that in a Nash equilibrium, both automata have the same number of states, and in any two periods of the cycle,[12] if an automaton plays the same stage game action, then it must be true of the other machine as well. As a consequence, equilibrium payoffs in 2×2 games must usually lie on line segments. These line segments connect two pure payoff profiles that correspond to action profiles, say a and b, such that $a_i \neq b_i$ for $i = 1, 2$.

B.1.2 Weak Renegotiation-Proofness

Weak renegotiation-proofness is a refinement of subgame perfection due to Farrell and Maskin (1989) and Bernheim and Ray (1989). The notion supposes that players agree ex-ante to play a subgame perfect equilibrium, but they are able to renegotiate the continuation after every period. Farrell and Maskin (1989) write: "players are unlikely to play, or to be deterred by, a proposed continuation equilibrium (on or off the equilibrium path) that is strictly Pareto dominated by another equilibrium that they believe is available to them." Therefore, a subgame-perfect equilibrium is renegotiation-proof if there does not exist any continuation equilibrium that strictly Pareto dominates any other. By restricting off-equilibrium behaviors, renegotiation-proofness can constrain the equilibrium paths that can be supported, leading to selection.

Main Result. Let A_i be player i's mixed action set. Say that a feasible payoff v is sustainable by a punishment-reward scheme if for each i, there is an action profile a^i that satisfies $u_i(a^i) \geq v_i$ and $\max_{a_j \in A_j} u_j(a_j, a^i_i) < v_j$. Farrell and Maskin (1989) demonstrate that every individually rational payoff that is sustainable by a punishment-reward scheme is a weakly renegotiation-proof equilibrium payoff, provided the discount factor is sufficiently close to one. In other words, a payoff profile v can be supported in equilibrium if there exist punishments that can deter any deviator (j) while rewarding the punisher (i) beyond v_i. If so, we can construct the appropriate continuations.

B.1.3 Internal Renegotiation-Proofness

Internal renegotiation-proofness, proposed by Ray (1994), strengthens the requirement of renegotiation-proofness in response to the permissiveness of the weak version. As already said, any *weakly* renegotiation-proof equilibrium relies on the existence of punishments that deter any deviator while rewarding the punisher. But these same punishments could be used to enforce more advantageous payoffs (to both players) than what this particular equilibrium offers. If this opportunity exists,

[12] Finite automata necessarily produce cyclic sequences of play.

then it is plausible that the players would renegotiate this equilibrium away. Therefore, it is important to look at the set of *all* payoffs that can be supported by some punishments and rewards, which is what internal renegotiation-proofness prescribes. Given the set of all supportable payoffs, no two of them can be strictly Pareto ranked, for otherwise the dominated one would not survive renegotiation.

Main Result. Ray (1994) proves that as the discount factor goes to one, every internally renegotiation-proof equilibrium payoff must be individually rational and either (i) on the Pareto frontier of the set of feasible payoffs or (ii) a singleton.

B.1.4 Nash Equilibrium with Tests

Spiegler (2005) introduces the concept of Nash Equilibrium with Tests (NEWT). This equilibrium concept, which is not a refinement of Nash equilibrium, is motivated by the fact that in standard equilibrium concepts, such as subgame perfection, the equilibrium path can be supported by players' beliefs for which there is no actual evidence. Spiegler (2005) writes: "Players validate their equilibrium beliefs by invoking counterfactual events that are never observed in equilibrium ... I construct an equilibrium concept that requires players to *justify their equilibrium beliefs by events that take place on the equilibrium path itself.*" Both players use finite automata as strategies, and the equilibrium concept requires that the first time player j visits a state, say q_j', player i reacts to it by playing the myopic best-responding action. However, there must be a time after which, every time that q_j' is visited again, player i plays the action prescribed by her dynamic best-response to q_j'. In other words, in a NEWT, players first play myopically against every state on their path, and in this sense explore their opponent's automaton, in order to justify their long-term behavior: each player eventually plays a dynamic best-response to each of her opponent's continuation automaton. Hence, the continuation automata form a Nash equilibrium.

Main Results. Spiegler (2005) proves the following results. Let br_j be player j's best-response in the stage game. First, suppose that the stage game has a unique Nash equilibrium and that it strictly Pareto-dominates any other stage-game outcome. Then, only infinite repetition of it is consistent with NEWT. Second, in coordination games — i.e. for every i and pure action a_i, $(a_i, br_j(a_i))$ is a pure Nash equilibrium of the stage game — every individually rational discounted payoff profile can be approximated (arbitrarily closely) by some NEWT, provided that the discount factor is sufficiently close to one.

B.1.5 Axiomatic Bargaining: the Nash (1950) and the Kalai and Smorodinsky (1975) Solutions

A pure bargaining problem is one where two (or n) individuals have to agree on a feasible outcome within a set, and in the event that no agreement is reached, a given disagreement point is the result. The traditional formulation, following Nash (1950), is in utility-space: a bargaining game consists of a set of feasible utilities and of a disagreement utility point. Axiomatic treatments propose various axioms that lead to different solutions. The Nash (1950) and the Kalai and Smorodinsky

(1975) solutions are among the most well-known solutions. The latter solution results from using a monotonicity axiom as a substitute for the axiom of independence of irrelevant alternatives used by Nash (1950). Both solutions predict a unique point in the set of feasible utilities. The main reason for studying these solutions in this context is that the Nash bargaining solution emerges naturally in models of repeated interactions (e.g., intertemporal bargaining models of Kambe (1999) and Abreu and Pearce (2007)).[13]

Main Result. Let $u^d = (u_1^d, u_2^d)$ be the disagreement point. The Nash (1950) solution is the point u (on the Pareto frontier) that maximizes the product $(u_1 - u_1^d)(u_2 - u_2^d)$. Let $u^* = (u_1^*, u_2^*)$ where u_i^* is the maximal feasible utility for player i. The Kalai and Smorodinsky (1975) solution lies at the intersection of the Pareto frontier and the line segment connecting u^d to u^*.

B.1.6 The Axiomatic Approach of Repeated Interactions

Mathevet (2014) proposes an axiomatic approach to study repeated interactions. This approach involves producing axioms on the infinite sequences of action profiles to capture essential features of repeated interactions, and then deriving solutions from them. Despite its abstraction, this approach is rather pragmatic, because the basic axioms of the theory apply to observable data, i.e., the sequences of play. Mathevet (2014)'s conceptual claim is that this method is rich enough to capture important dynamic phenomena through various axioms, and yet parsimonious enough to study their implications in a flexible fashion. In particular, since (intertemporal) utilities come from sequences of play, the class of solutions available in sequence-space is inevitably larger than in utility-space, as in axiomatic bargaining (Section B.1.5). Mathevet (2014) suggests several axioms that a solution should satisfy. In this paper, we study the solution that results from his three main axioms: individual rationality, collective intelligence, and internal stability.

Main Result. According to Mathevet (2014)'s characterization, the solution payoffs contain all payoffs v satisfying the following conditions. First, v must be individually rational. Second, v must lie on the Pareto frontier of a line segment connecting two pure payoff profiles (trivially, this Pareto frontier is the segment itself when the extremities are Pareto unordered). Third, v must lie on a line segment starting from the origin $(0, 0)$ and whose slope is determined by players' observed bargaining power (see Mathevet (2014)).

B.2 Adaptive Learning Models

The adaptive learning models studied in this paper are all related to the experience weighted attraction (EWA) model presented by Camerer and Ho (1999). Though the standard EWA model is not tested here, it is presented to simplify the presentation of the models that are tested here. The standard EWA model combines the two common learning theories; belief-based learning and reinforcement learning; into a hybridized model. The players attractions are updated based on two

[13] These bargaining models assume more structure than our experiments have, e.g. players must be able to sign enforceable contracts. In this sense, our experiments are not meant to test these models, yet these models make the axiomatic solutions a relevant landmark in this context.

equations. The first equation is the experience in period t which determines the relative weight placed on previous attractions and last periods play:

$$N(t) = \rho N(t-1) + 1$$

The second equation is the attraction of player i for strategy j in period t, which is determined by the attraction in period $t-1$, the hypothetical forgone payoffs, the period t experience $N(t)$, and two parameters ϕ and δ as follows:

$$A_i^j(t) = \frac{\phi N(t-1) A_i^j(t-1) + \left[\delta + (1-\delta) I\left(s_i^j, s_i(t)\right)\right] \pi_i\left(s_i^j, s_{-i}(t)\right)}{N(t)},$$

where $s_i(t)$ is the joint-strategy played by all other players than i in period t. The parameter ϕ is the weight put on previous attractions, and the parameter δ is the weight given to hypothetical forgone payoffs. Once the attractions are determined, the probability that player i plays strategy j in period t is determined by the logistic choice function with parameter λ,

$$p_i^j(t) = \frac{e^{\lambda A_i^j(t)}}{\sum_{k \in S_i} e^{\lambda A_i^k(t)}}. \tag{8}$$

Therefore, the standard EWA model has four parameters: ρ, ϕ, δ, and λ. In addition, the initial level of experience $N(0)$ and initial attractions $A_i^j(0)$ need to be set.

B.2.1 Weighted Fictitious Play

Weighted fictitious play was introduced by Cheung and Friedman (1997). From the standard EWA model if $\delta = \rho = \phi = 1$ and $N(0) = 0$, then we are left with,

$$A_i^j(t) = \frac{(t-1) A_i^j(t-1) + \pi_i\left(s_i^j, s_{-i}(t)\right)}{t},$$

The attractions are therefore simply the average payoff that player i would have received if they had played s_i^j in every period. Then the action is chosen using the logistic choice function with parameter λ. This differs from the original model of fictitious play proposed by Brown (1951) in that players make choices with the logistic choice function rather than simply choosing the action with the highest attraction.

B.2.2 Reinforcement Learning

From the standard EWA model if $\delta = \rho = 0$ and $N(0) = 1$, then we are left with,

$$A_i^j(t) = \phi A_i^j(t-1) + I\left(s_i^j, s_i(t)\right) \pi_i\left(s_i^j, s_{-i}(t)\right),$$

In each period all attractions are discounted by ϕ, and the attraction for the action that was played gets the payoff added as well.

B.2.3 Self-tuning EWA

The standard EWA has been widely successful as it has proven to fit experimental data in a variety of games. One potential drawback of EWA is the large number of free parameters. Ho, Camerer, and Chong (2007) develop the self-tuning EWA model which reduces EWA down to a one parameter model by changing several parameters to self-tuning functions. The centerpiece of the model is the two self-tuning functions: the attention function $\delta_{i,j}$ and the change detector function $\phi_i(t)$. The attention function $\delta_{ij}(t)$ replaces the parameter δ, and takes the form,

$$\delta_{ij}(t) = \begin{cases} 1 & \text{if } u_i\left(s_i^j, s_{-i}(t)\right) \geq u_i(t) \\ 0 & \text{otherwise} \end{cases} .$$

The attention function is based on the idea that individuals are more aware of actions that would have lead to a higher payoff than those that would have lead to a lower payoff. The change-detector function $\phi_i(t)$ replaces the parameter ϕ. First, define the cumulative history vector for strategy k for player i is,

$$h_i^k(t) = \frac{\sum_{\tau=1}^{t} I\left(s_{-i}^k, s_{-i}(\tau)\right)}{t},$$

which tells the frequency with which a strategy has been played up to period t. Next, the recent history vector is 0's and 1's, $r_i^k(t) = I\left(s_{-i}^k, s_{-i}(t)\right)$. The surprise index is then,

$$S_i(t) = \sum_{k=1}^{m_{-i}} \left(h_i^k(t) - r_i^k(t)\right)^2 \in [0, 2)$$

Finally, the change-detector function is

$$\phi_i(t) = 1 - \frac{1}{2} S_i(t).$$

The idea behind the change-detector function is that when play is relatively stationary (low surprise index), then players base their attraction more on previous attractions than current play. On the other hand, when play is changing (high surprise index), then players put less weight on their previous attractions, and weight the current payoffs more heavily. In addition to these self-tuning function, the following additional assumptions are made: parameter ρ is always equal to ϕ, $\phi = \rho$; initial experience is set to $N(0) = 1$; and initial attractions are set using the cognitive hierarchy model (Camerer, Ho, and Chong, 2004) with 1.5 steps of reasoning. This leaves the self-tuning EWA model with one free parameter, λ from the logistic function (8) which maps attractions to probabilities.

B.2.4 Strategic EWA

B.2.5 Action-Sampling Learning

Action-sampling learning is an extension of the action-sampling equilibrium concept introduced in Selten and Chmura (2008). In each period of action-sampling learning, a player makes their choice by randomly choosing n samples with replacement from their opponents vector of past play. The subject then forms beliefs from these n samples, and best responds. That is, suppose in period t player 1 samples n terms $R_1 = \{r_1, \ldots, r_n\}$ from player 2's history from the from the $t-1$ periods. Then player 1 chooses,

$$s_1(t) = \underset{s_1 \in S_1}{\operatorname{argmax}} \sum_{r \in R_1} u_1 \left(s_1, r\right).$$

It is also assumed that player play randomly (uniform probability) over strategies until their opponent has picked all strategies at least once.

C Theory Predictions

References

Abreu, D., and D. Pearce (2007): "Bargaining, Reputation, and Equilibrium Selection in Repeated Games with Contracts," *Econometrica*, 75(3), 653–710.

Abreu, D., and A. Rubinstein (1988): "The Structure of Nash Equilibrium in Repeated Games with Finite Automata," *Econometrica*, 56(6), 1259–81.

Aoyagi, M., and G. Frechette (2009): "Collusion as public monitoring becomes noisy: Experimental evidence," *Journal of Economic Theory*, 144(3), 1135–1165.

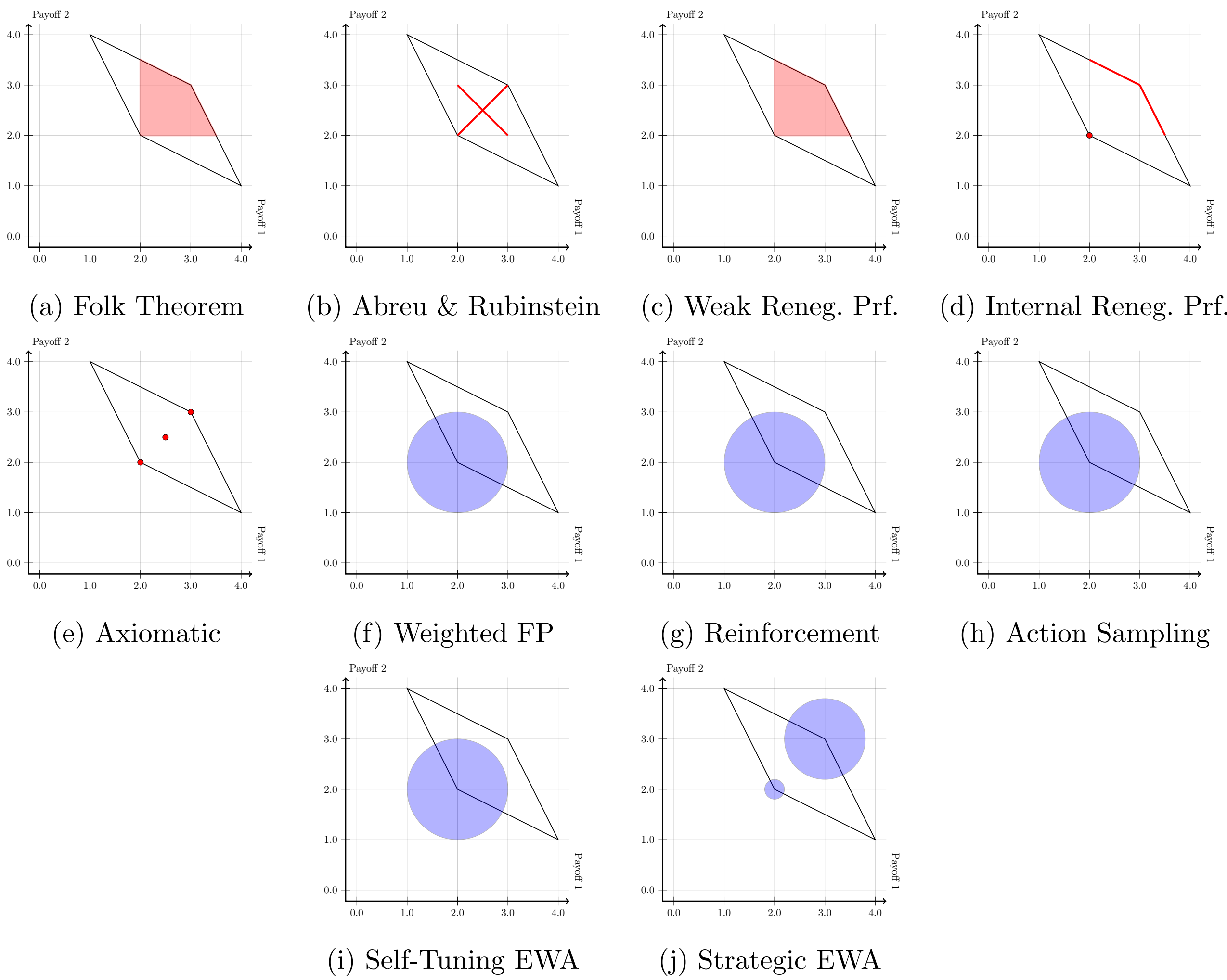
Arifovic, J., R. D. McKelvey, and S. Pevnitskaya (2006): "An initial implementation of the Turing tournament to learning in repeated two-person games," *Games and Economic Behavior*, 57(1), 93–122.

Barbera, S., W. Bossert, and P. K. Pattanaik (2004): "Ranking Sets of Objects," in *Handbook of Utility Theory. Volume II Extensions*, ed. by S. Barbera, P. J. Hammond, and C. Seidl, pp. 893–977. Kluwer Academic Publishers, Dordrecht.

Beatty, T. K. M., and I. A. Crawford (2011): "How Demanding Is the Revealed Preference Approach to Demand?," *American Economic Review*, 101(6), 2782–95.

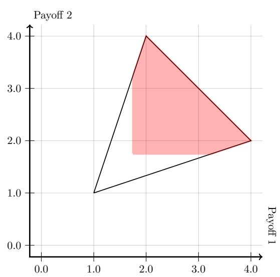
Bernheim, D. B., and D. Ray (1989): "Collective dynamic consistency in repeated games," *Games and Economic Behavior*, 1(4), 295–326.

Brown, G. (1951): "Iterative solution of games by fictitious play," *Activity analysis of production and allocation*, 13, 374–376.

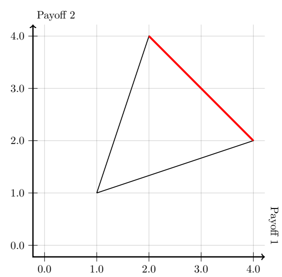


(a) Folk Theorem
(b) Abreu & Rubinstein
(c) Weak Reneg. Prf.
(d) Internal Reneg. Prf.
(e) Axiomatic
(f) Weighted FP
(g) Reinforcement
(h) Action Sampling
(i) Self-Tuning EWA
(j) Strategic EWA

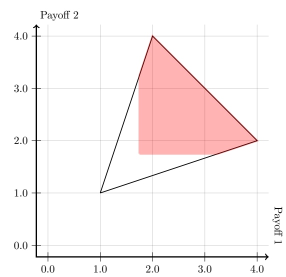
Figure 14: Predictions for Prisoners' Dilemma



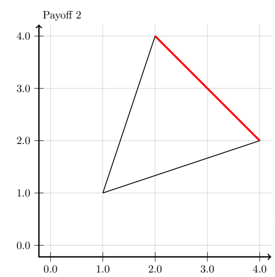
(a) Folk Theorem



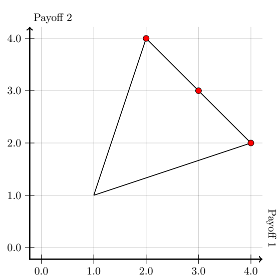
(b) Abreu & Rubinstein



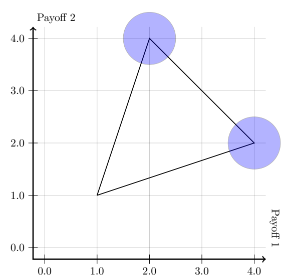
(c) Weak Reneg. Prf.



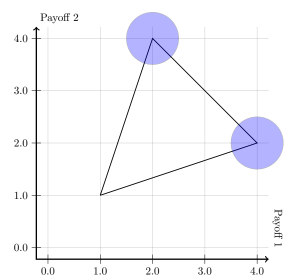
(d) Internal Reneg. Prf.



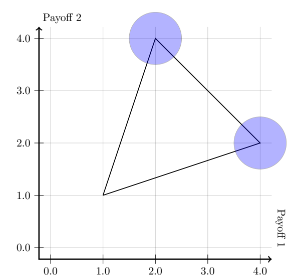
(e) Axiomatic



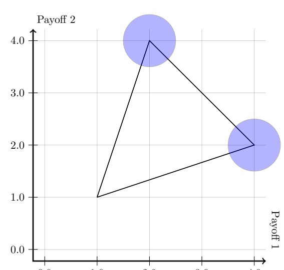
(f) Weighted FP



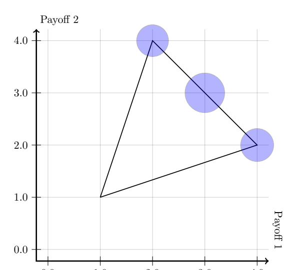
(g) Reinforcement



(h) Action Sampling

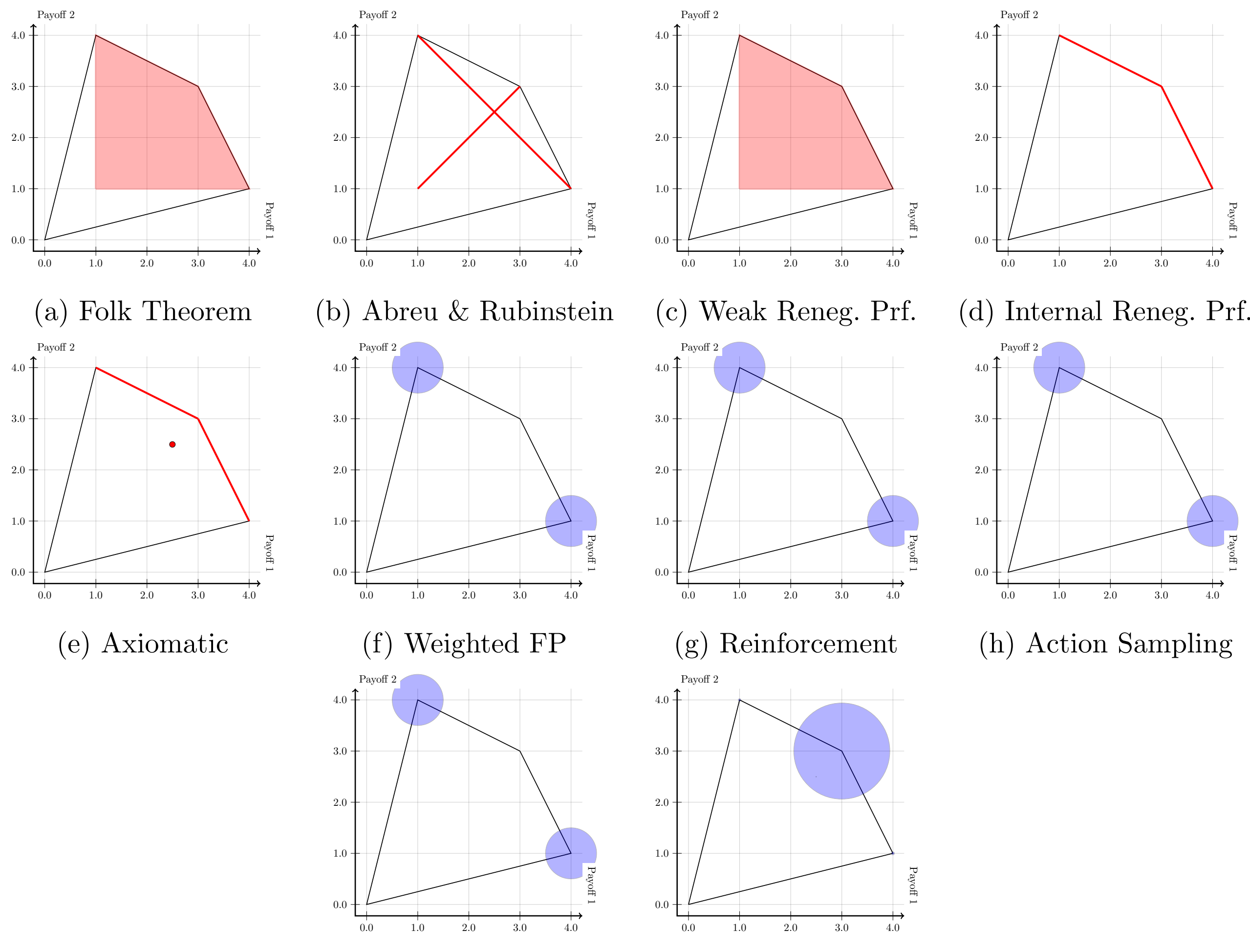


(i) Self-Tuning EWA



(j) Strategic EWA

Figure 15: Predictions for Battle of the Sexes

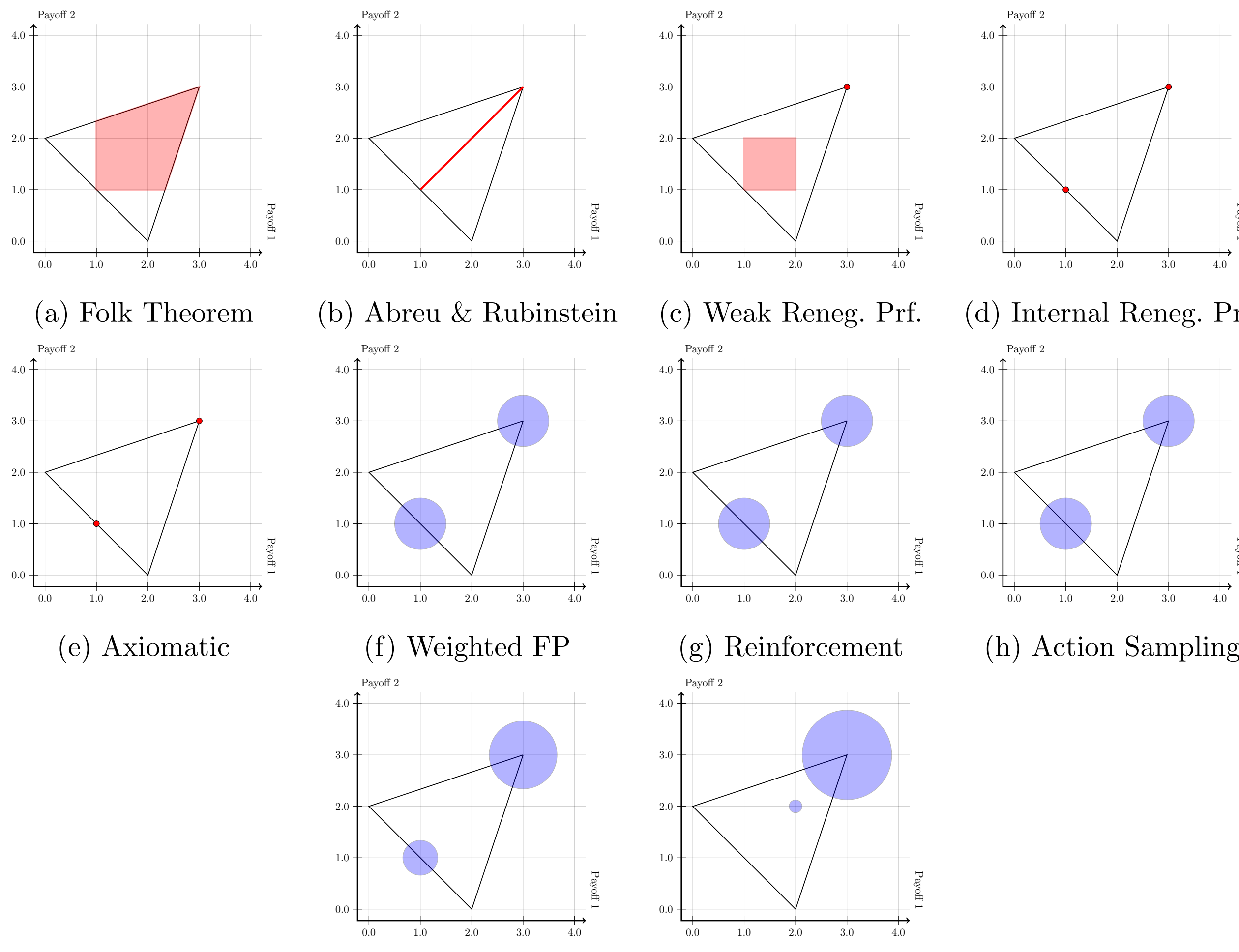


(a) Folk Theorem (b) Abreu & Rubinstein (c) Weak Reneg. Prf. (d) Internal Reneg. Prf.

(e) Axiomatic (f) Weighted FP (g) Reinforcement (h) Action Sampling

(i) Self-Tuning EWA (j) Strategic EWA

Figure 16: Predictions for Chicken

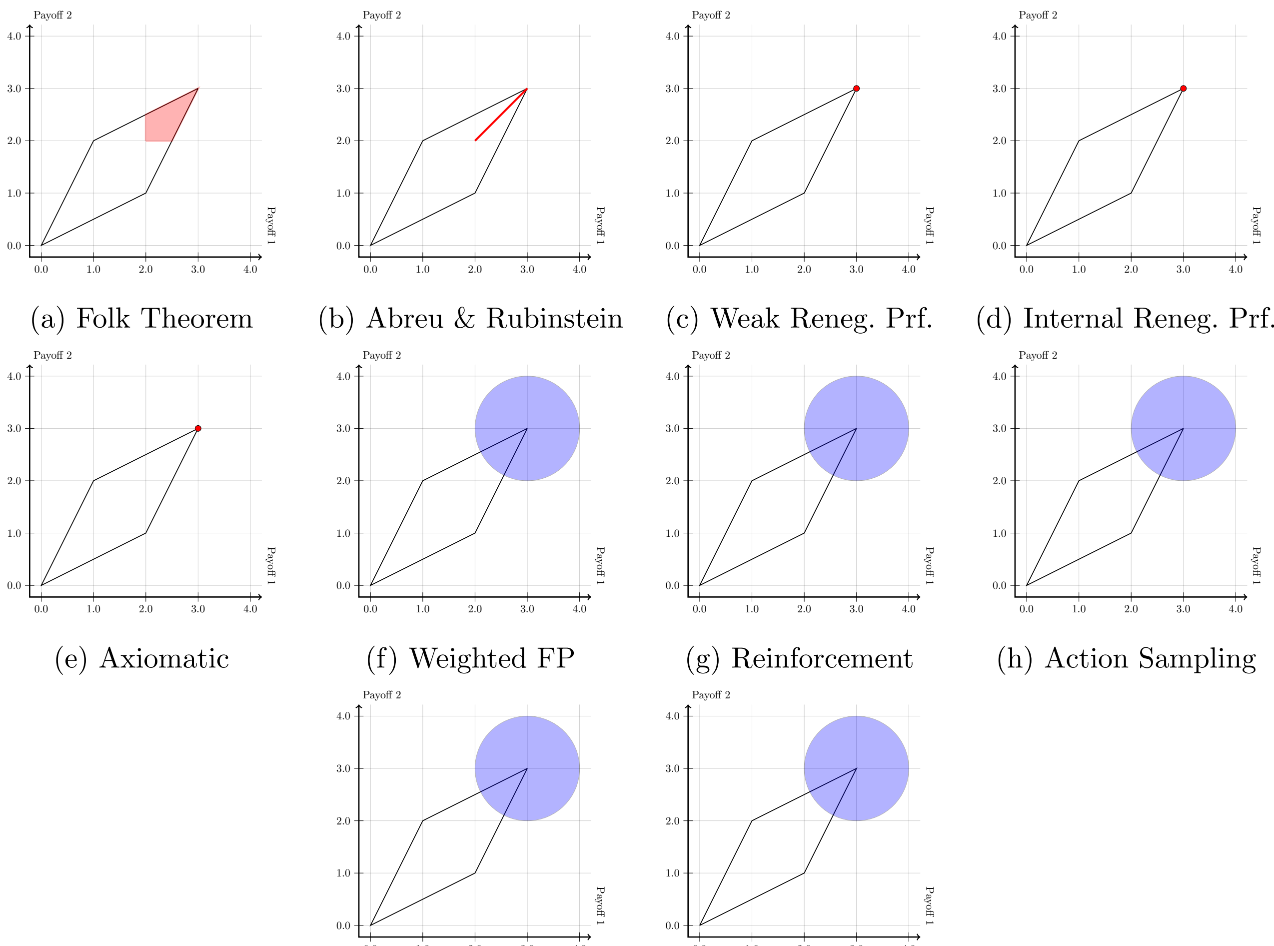


(a) Folk Theorem (b) Abreu & Rubinstein (c) Weak Reneg. Prf. (d) Internal Reneg. Prf.

(e) Axiomatic (f) Weighted FP (g) Reinforcement (h) Action Sampling

(i) Self-Tuning EWA (j) Strategic EWA

Figure 17: Predictions for Stag Hunt

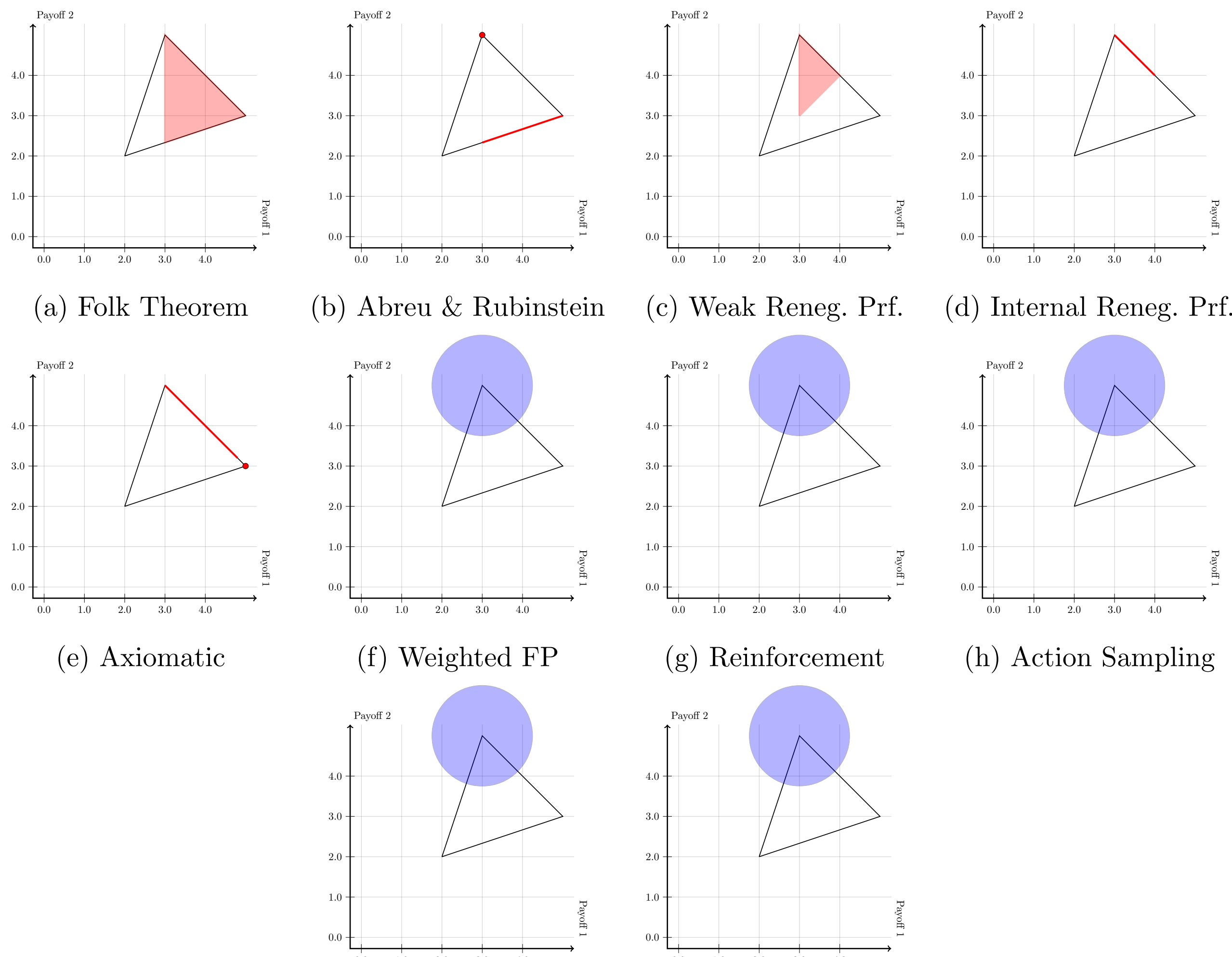


(a) Folk Theorem (b) Abreu & Rubinstein (c) Weak Reneg. Prf. (d) Internal Reneg. Prf.

(e) Axiomatic (f) Weighted FP (g) Reinforcement (h) Action Sampling

(i) Self-Tuning EWA (j) Strategic EWA

Figure 18: Predictions for Common Interest Game

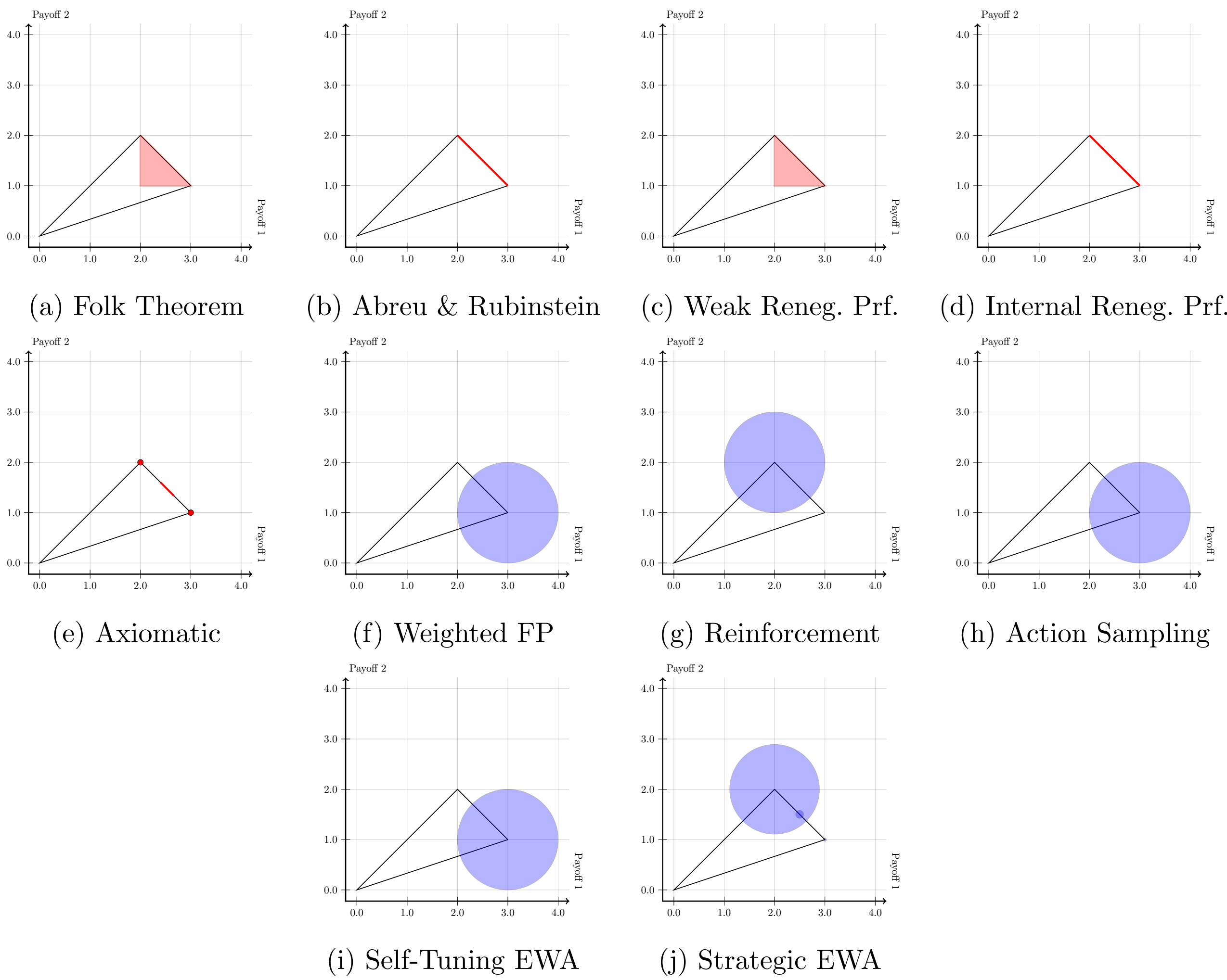


(a) Folk Theorem (b) Abreu & Rubinstein (c) Weak Reneg. Prf. (d) Internal Reneg. Prf.

(e) Axiomatic (f) Weighted FP (g) Reinforcement (h) Action Sampling

(i) Self-Tuning EWA (j) Strategic EWA

Figure 19: Predictions for Stag Hunt

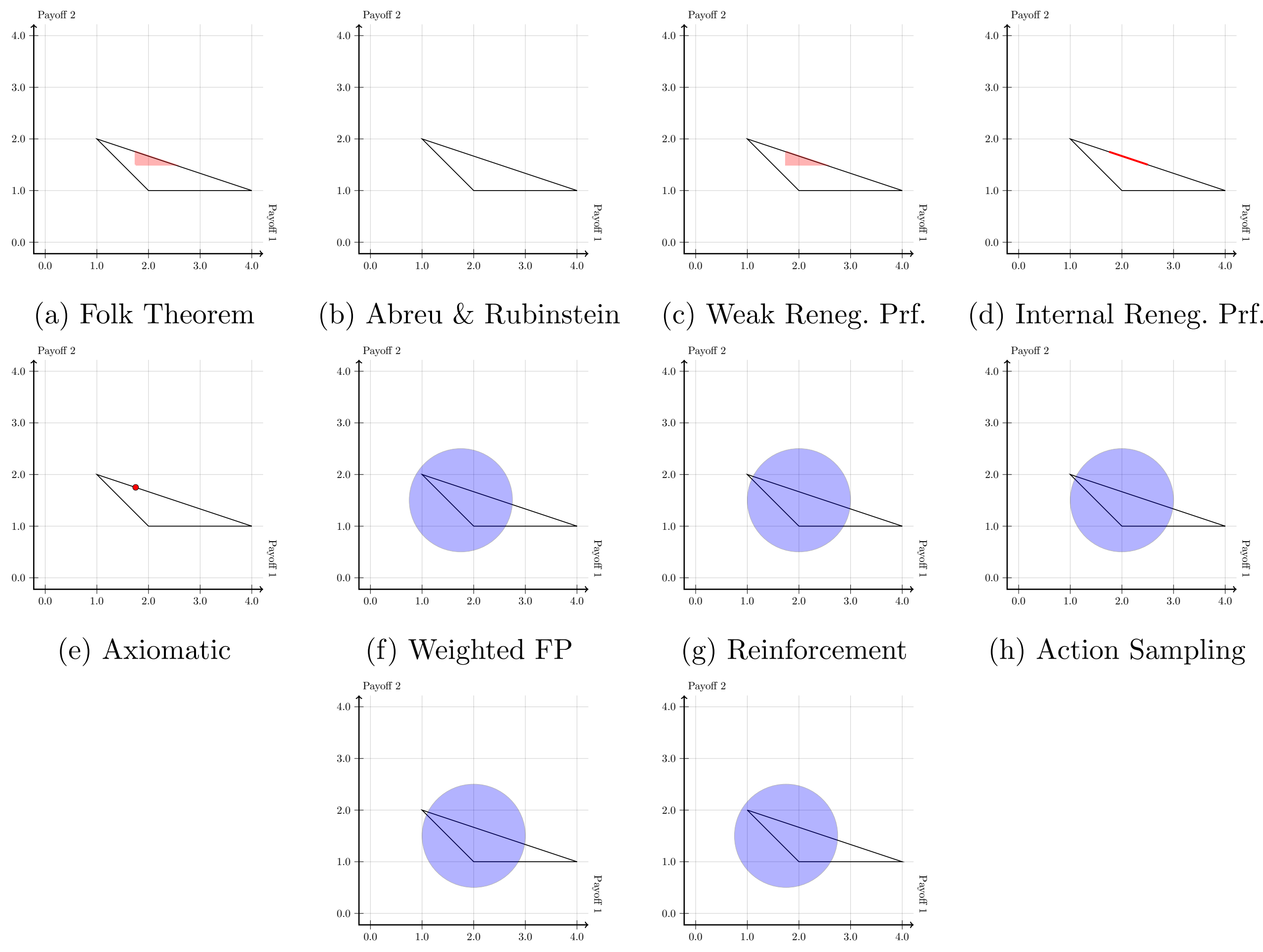


(a) Folk Theorem (b) Abreu & Rubinstein (c) Weak Reneg. Prf. (d) Internal Reneg. Prf.

(e) Axiomatic (f) Weighted FP (g) Reinforcement (h) Action Sampling

(i) Self-Tuning EWA (j) Strategic EWA

Figure 20: Predictions for Ultimatum Game



(a) Folk Theorem (b) Abreu & Rubinstein (c) Weak Reneg. Prf. (d) Internal Reneg. Prf.

(e) Axiomatic (f) Weighted FP (g) Reinforcement (h) Action Sampling

(i) Self-Tuning EWA (j) Strategic EWA

Figure 21: Predictions for Unique Mixed Equilibrium Game

Brunner, C., C. Camerer, and J. Goeree (2011): "Stationary Concepts for Experimental 2 X 2 Games: Comment," *American Economic Review*, 101(2), 1029–40.

Camerer, C., and T.-H. Ho (1999): "Experience-weighted Attraction Learning in Normal Form Games," *Econometrica*, 67(4), 827–874.

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2002): "Sophisticated Experience-Weighted Attraction Learning and Strategic Teaching in Repeated Games," *Journal of Economic Theory*, 104(1), 137–188.

——— (2004): "A Cognitive Hierarchy Model of Games*," *Quarterly Journal of Economics*, 119(3), 861–898.

Cheung, Y.-W., and D. Friedman (1997): "Individual Learning in Normal Form Games: Some Laboratory Results," *Games and Economic Behavior*, 19(1), 46 – 76.

Chmura, T., S. Goerg, and R. Selten (2011): "Learning in repeated 2 x 2 games," *Games and Economic Behavior*, 76, 43–73.

Commons, J. R. (1934): *Institutional Economic*. The McMillan Company, New-York.

Dal Bo, P. (2005): "Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games," *American Economic Review*, 95(5), 1591–1604.

Dal Bo, P., and G. Frechette (2011): "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence," *American Economic Review*, 101(1), 411–29.

Duffy, J., and J. Ochs (2009): "Cooperative behavior and the frequency of social interaction," *Games and Economic Behavior*, 66(2), 785–812.

Er'ev, I., and A. E. Roth (1998): "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, 88(4), 848–81.

Farrell, J., and E. Maskin (1989): "Renegotiation in repeated games," *Games and Economic Behavior*, 1(4), 327–360.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10(2), 171–178.

Fishburn, P. C. (1986): "The axioms of subjective probability," *Statistical Science*, 1(3), 335–345.

——— (1992): "Signed orders and power set extensions," *Journal of Economic Theory*, 56(1), 1–19.

Friedman, M. (1953): *Essays in Positive Economics*. University of Chicago Press, Chicago, Illinois.

Fudenberg, D., and D. K. Levine (1998):. MIT Press, Cambridge, Massachusetts.

Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54(3), 533–54.

Harless, D. W., and C. Camerer (1994): "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, 62(6), 1251–89.

Hart, S., and A. Mas-Colell (2000): "A Simple Adaptive Procedure Leading to Correlated Equilibrium," *Econometrica*, 68(5), 1127–1150.

Hausman, D. M. (1998): "The Philosophy of Economics," in *Routledge Encyclopedia of Philosophy*, ed. by E. Craig, vol. 3, pp. 211–22. Routledge, London.

Hey, J. D., and C. Orme (1994): "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62(6), 1291–1326.

Ho, T. H., C. F. Camerer, and J.-K. Chong (2007): "Self-tuning experience weighted attraction learning in games," *Journal of Economic Theory*, 133(1), 177 – 198.

Holt, C. A. (1985): "An Experimental Test of the Consistent-Conjectures Hypothesis," *American Economic Review*, 75(3), 314–25.

Homan, P. T. (1931): "The Institutionalist School," in *Encyclopedia of the Social Sciences*, vol. 5. The Macmillan Company.

Ioannou, C. A., and J. Romero (2014): "A generalized approach to belief learning in repeated games," *Games and Economic Behavior*, 87(0), 178 – 203.

Kalai, E., and M. Smorodinsky (1975): "Other Solutions to Nash's Bargaining Problem," *Econometrica*, 43(3), 513–18.

Kambe, S. (1999): "Bargaining with Imperfect Commitment," *Games and Economic Behavior*, 28(2), 217–237.

Karandikar, R., D. Mookherjee, D. Ray, and F. Vega-Redondo (1998): "Evolving Aspirations and Cooperation," *Journal of Economic Theory*, 80(2), 292–331.

Knight, F. (1940): "What is "truth" in economics?," *Journal of Political Economy*, 48, 1–32.

Mailath, G., and L. Samuelson (2006): *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, New-York.

Mathevet, L. (2014): "Axiomatic Behavior in Repeated Games," .

Milgrom, P., and J. Roberts (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58(6), 1255–1277.

MILL, J. S. (1843): *A System of Logic, Ratiocinative and Inductive*, vol. 1. John W. Parker, London.

NASH, J. (1950): "The Bargaining Problem," *Econometrica*, 18(2), 155–162.

PALFREY, T., AND H. ROSENTHAL (1994): "Repeated Play, Cooperation and Coordination: An Experimental Study," *Review of Economic Studies*, 61(3), 545–65.

RAY, D. (1994): "Internally Renegotiation-Proof Equilibrium Sets: Limit Behavior with Low Discounting," *Games and Economic Behavior*, 6, 162–177.

ROTH, A., AND J. K. MURNIGHAN (1978): "Equilibrium behavior and repeated play of the prisoners' dilemma," *Journal of Mathematical Psychology*, 39(1), 83–96.

ROTHBARD, M. (1957): "In Defense of 'Extreme Apriorism'," *Southern Economic Journal*, 23, 314–320.

RUBINSTEIN, A. (2006): "Dilemmas of an Economic Theorist," *Econometrica*, 74(4), 865–883.

SELTEN, R. (1991): "Properties of a measure of predictive success," *Mathematical Social Sciences*, 21(2), 153–167.

——— (1998): "Axiomatic Characterization of the Quadratic Scoring Rule," *Experimental Economics*, 1(1), 43–61.

SELTEN, R., AND T. CHMURA (2008): "Stationary Concepts for Experimental 2x2-Games," *American Economic Review*, 98(3), 938–66.

SONSINO, D. (1997): "Learning to Learn, Pattern Recognition, and Nash Equilibrium," *Games and Economic Behavior*, 18(2), 286–331.

SPIEGLER, R. (2005): "Testing threats in repeated games," *Journal of Economic Theory*, 121(2), 214–235.

VON MISES, L. (1949): *Human Action. A Treatise on Economics*. Yale University Press, New Haven.