

Machine Games: Theory and Experimental Evidence

Christos A. Ioannou ^{*†}

Université Paris 1 Panthéon - Sorbonne

This draft: June 25, 2025

Abstract

We propose an equilibrium model of behavior with Moore machines. The machines are subjected to some small likelihood of committing implementation errors. We analyze the machine game, where each player chooses in the beginning and commits thereafter to a machine to play an infinitely-repeated game. A pair of machines induces a sequence of action profiles, which are modelled by a Markov process. We derive predictions on a set of 2×2 games consisting of payoff variations of the Prisoner's Dilemma, Stag Hunt and Battle of the Sexes. Crucially, predicted pairs differ across payoff variations. We also conduct experiments with a high continuation probability (specifically, $\delta = 0.99$), and contrast the frequent patterns mined from the experimental data with the equilibrium game plays and the action profiles induced by the equilibrium machine pairs. We find that almost all predicted plays and machines find support in the data.

JEL: C72, C81, C63, D83

Keywords: Automata, Markov Process, Repeated Games, Experiments

^{*}I would like to thank Christian Hilbe, Anastasios Magdalinos, Laurent Mathevet, Martin A. Nowak, Yaroslav Rosokha, Dominik Schmidt, Doron Sonsino for helpful suggestions, Benjamin Levy and Irene Locci for excellent research assistance. This work has been funded by a French government subsidy managed by the Agence Nationale de la Recherche under the framework of the Investissements d'Avenir Programme (reference ANR-17-EURE-001). Finally, I am indebted to the editor, Stefan Penczynski, and two anonymous referees for their insightful comments, which significantly improved the paper. The usual disclaimer applies.

[†]Centre d'Économie de la Sorbonne, Université Paris 1 Panthéon - Sorbonne. Email: christos.ioannou@univ-paris1.fr

1 Introduction

Folk Theorems establish, under a variety of assumptions about the players' preferences, that a multitude of payoff profiles is compatible with the Nash equilibrium in an infinitely-repeated game.¹ Yet the abundance of Nash equilibrium payoff profiles is problematic as the characterizations often predict profiles with no empirical validity (see [Friedman \(1971\)](#), [Tirole \(1988\)](#), [Fudenberg and Maskin \(1993\)](#)). A model of behavior that predicts instead *only* empirically-valid profiles would therefore be more credible. Another desideratum for any model of behavior is that it delivers changes (in the right direction) in predicted behavior for variations in the payoffs within a game. Indeed, experimental literature provides compelling evidence that variations in the payoffs may influence behavior. In the Prisoner's Dilemma, [Blonski, Ockenfels, and Spagnolo \(2011\)](#) point out that cooperation rates depend on the payoff earned by cooperating when the other player defects (see also [Dal Bó and Fréchette \(2018\)](#)). In the Stag Hunt, [Battalio, Samuelson, and Van Huyck \(2001\)](#) argue that the payoff-dominant equilibrium is more likely to emerge the smaller the optimization premium is.² Lastly, in the Battle of the Sexes, [Sonsino and Sirota \(2003\)](#) find that the proportion of pairs alternating on the two pure-strategy Nash equilibria varies considerably across different payoff matrices (see also [Arifovic and Ledyard \(2018\)](#)).

In the present study, we propose an equilibrium model of behavior in infinitely-repeated games that tackles with the above challenges. Our theoretical framework assumes first a generalized strategy subspace. The set of possible strategies in repeated games is infinite (uncountable) making it unrealistic and impractical to expect a player to thoroughly explore such an infinite set. We thus restrict the space drastically by considering only strategies defined by finite-state automata called Moore machines ([Moore \(1956\)](#)). A *Moore machine* consists of a finite set of internal states (one of which is specified to be the initial state), an output function and a transition function. We assume further that the machines are subjected to some small likelihood of committing errors in the implementation of actions along the lines of Selten's trembling hand ([Selten \(1975\)](#)). We thus study the players' behavior by analyzing the

¹The Folk Theorem in repeated games with discounting states that, for every feasible and individually rational stage-game payoff profile, there is a discount factor large enough such that the payoff profile can be generated by an equilibrium of the repeated game.

²The optimization premium is the difference between the payoffs of the best response to the other player's strategy and the inferior response. The relationship between the optimization premium and the payoff-dominant equilibrium is corroborated in [Dal Bó, Fréchette, and Kim \(2021\)](#).

machine game, where each player chooses in the beginning and commits thereafter to a machine to play an infinitely-repeated game.

In the proof-of-concept application, we first specify a limited number of archetypal machines consisting of no more than two states. Although this subspace is relatively small, it is still sufficiently rich to support sophisticated strategic behavior, such as punishments, triggers as well as anticipation of punishments and triggers (see [Stahl \(1996, 1999\)](#), [Stahl and Haruvy \(2012\)](#)). A pair of machines induces a sequence of action profiles, which are modelled by a Markov process. The occurrence of errors ensures that the process is ergodic, which enables us to compute the invariant distribution of the action profiles for each machine pair. We derive predictions on a set of 2×2 games consisting of three payoff variations of the Prisoner’s Dilemma and Stag Hunt, and two payoff variations of the Battle of the Sexes. In each game type, we allow for both symmetric and asymmetric versions.³

We next conduct online experiments with a high continuation probability (specifically, $\delta = 0.99$), where subjects are paired in a fixed matching protocol to play one of the aforementioned games with perfect monitoring and complete information. Finally, we adopt the methods from [Ioannou et al. \(2025\)](#) (inspired from the Winepi algorithm of [Mannila, Toivonen, and Verkamo \(1997\)](#)) to mine the experimental data for frequent patterns in each game,⁴ and validate them against the equilibrium game plays and the action profiles induced by the equilibrium machine pairs. Crucially, our validation procedure is markedly different from studies that utilize maximum likelihood estimation to find the proportion of the population that uses a particular strategy/machine. Here, we determine what percentage of the frequent patterns detected in each game *could* have been generated by a specific equilibrium machine (see Subsection 5.3 for more details).⁵

³Asymmetry in the payoffs may complicate the environment for the players or, under suitable parametrizations, simplify it. The former will happen if payoff-asymmetry causes a normative conflict amongst such principles as efficiency and egalitarianism ([Nikiforakis, Noussair, and Wilkening \(2012\)](#)). In contrast, payoff-asymmetry may simplify the environment if it makes an efficient and egalitarian profile salient relative to the other asymmetric ones. We test here both environments.

⁴[Ioannou, Mathevet, Romero, and Zhang \(2025\)](#) show that in games with a sufficiently high continuation probability, oftentimes participants reach a predictable way of playing the game (i.e. a frequent pattern). Therefore, it is only natural to compare the frequent patterns extracted from the experimental data to the induced action profiles of the predicted machines.

⁵A pattern, in general, cannot be identified with one single machine (see the inference problem of [McKelvey and Palfrey \(2001\)](#)), but with a set of machines. Consequently, even though the pattern can be extracted (given that the history of play is publicly observed), we can only attribute it to a set of possible machines.

We make four key contributions in this paper. Though we are neither the first to conduct an experiment that utilizes a high continuation probability⁶ nor the first to use asymptotic theory to derive predictions, we are the first to integrate the two together. The close proximity of the continuation probability to 1, allows a *direct* comparison between the asymptotic predictions and the frequent patterns mined from the experimental data. Our approach is thus holistic in the sense that it combines theoretical predictions with an empirical investigation by means of a controlled experiment. Furthermore, the high continuation probability induces a sufficiently long interaction between the same two participants (in expectation 100 periods of game play), which gives us insights into what happens in long repeated games (for example, whether cooperation/coordination will be sustained).

The second contribution is that the proposed theoretical framework is versatile enough to predict *different* machine pairs for variations in the payoffs within a game. Importantly, the predicted pairs are also in line with the comparative statics drawn from the existing literature. For example, holding everything else constant, [Dal Bó and Fréchette \(2018\)](#) argue that, in the Prisoner’s Dilemma, if (i) the gain from defection when the other player cooperates, and (ii) the loss from cooperation when the other player defects both go down, then the frequency of cooperative plays will go up. Though, in the first version of the Prisoner’s Dilemma, the proposed framework does not predict cooperative plays, but defecting ones, in the second version, where both aforementioned conditions go down relative to the first version, the framework predicts both cooperative and defecting plays.

Our third contribution is substantive. We find that almost all of the equilibrium game plays and action profiles induced by the equilibrium machine pairs find support in the data. This finding is important for two reasons. First, it provides strong evidence to support the empirical validity of the proposed theoretical framework. Second, it allows researchers to take a step forward by identifying plausible machines deployed by participants, which can help navigate future theoretical work towards refinements or conditions that lead to the use of these machines.

The fourth contribution is a statistical model that allows us to determine the games where the theoretical model performs (relatively) better. In our application,

⁶Very few studies exist that utilize such a high continuation probability. We are aware of one study that utilizes a $\delta = 0.99$ and 0.995 ([Ioannou et al. \(2025\)](#)), one study that utilizes a $\delta = 0.98$ ([Romero and Rosokha \(2019\)](#)), and two studies that utilize a $\delta = 0.95$ ([Romero and Rosokha \(2018, 2023\)](#)).

equilibrium game plays have a point mass distribution. Standard, statistical hypothesis testing cannot be used on this occasion as a single, non-compliant (to the equilibrium predictions) play is a sufficient datapoint to provide evidence against the theoretical prediction. We address this issue by modifying the statistical approach of Ioannou and Makris (2019) to assess the closeness (in a sense explained in Subsection 5.4) of the empirical distribution to the theoretical one in each game.

The paper is structured as follows. We first situate the current study in the relevant literature. In Section 3, we present the theoretical framework and a proof-of-concept application. In Section 4, we present the games, the experimental protocol, and the theoretical predictions. In Section 5, we first explain briefly the approach utilized to mine the frequent patterns. We then proceed to contrast the frequent patterns detected in the experimental data with the predicted game plays, and the action profiles induced by the predicted machine pairs. Lastly, we provide a statistical approach to evaluate the theoretical model proposed across our selected games. In Section 6, we offer concluding remarks.

2 Related Literature

The paper relates and contributes to a few strands of the literature. First, the present study adds to an extensive literature on equilibrium selection. In this strand, we can identify five themes. In the first theme, we find the seminal notions of risk and payoff dominance due to Harsanyi and Selten (1988). The second theme utilizes structural models to select equilibria.⁷ The third theme centers on the prevailing common view that in order to escape a prediction of equilibrium indeterminacy, the environment needs to possess a sufficiently large degree of heterogeneity and/or of asymmetric information (see Ioannou (2022)). Of special interest in this theme (and relevant to our study as well) are the coordination games. The theory of Global games of Carlsson and van Damme (1993) constitutes the most popular approach to escape the prediction of equilibrium indeterminacy by means of deploying uncertainty about the economic fundamentals (see also Morris and Shin (1998, 2003)).⁸ The fourth

⁷For instance, Zhang and Hofbauer (2016) use Quantal Response equilibrium as a homotopy method for equilibrium selection to study 2×2 coordination games (see also Turocy (2005)).

⁸An alternative approach, the theory of Poisson games, models the number of actual players as a Poisson random variable to capture population uncertainty in large coordination games (Makris

theme includes theories that rely on evolutionary approaches as a mechanism to coordinate on certain equilibrium outcomes (e.g. [Ellison \(1993\)](#), [Kandori, Mailath, and Rob \(1993\)](#), [Cooper \(1996\)](#), [Binmore and Samuelson \(1997, 1999\)](#)). Finally, the fifth theme in this strand, and the one more closely related to this study, includes theoretical contributions that rely on some boundedly rational learning process to select an equilibrium outcome (e.g. [Ellison \(1997\)](#), [Jéhiel \(2001\)](#)).

Our paper also pertains to the strand that studies the set of Nash equilibrium payoffs and the structure of the equilibria in the machine game. Constraining the players in the infinitely-repeated game to choose machines in lieu of (repeated-game) strategies does not alter the set of equilibrium payoffs (i.e. we obtain the same conclusion as that of the Folk Theorem). However, including in the ordering of players' preferences a dislike towards the machine complexity (secondary to the payoffs) improves predictability. Specifically, [Abreu and Rubinstein \(1988\)](#) show that, if agents' preferences are increasing in repeated-game payoffs and decreasing in the complexity of the machines employed, the set of Nash equilibrium payoffs that can occur is dramatically reduced from the Folk Theorem result. Yet the authors indicate that a wide variety of payoffs still remain consistent with equilibrium behavior even in the presence of complexity costs.⁹ Existing literature also characterizes the structure of the equilibria in the machine game. [Abreu and Rubinstein \(1988\)](#) show that in any equilibrium of the machine game there is one-to-one correspondence between the actions of each of the two players (see also [Piccione \(1992\)](#)).

A third strand that relates to this study is the recent literature investigating the strategies experimental subjects use in infinitely-repeated games. Two general approaches have been used to tackle with the issue of strategy inference. The first approach involves *direct* observation of strategies using the strategy elicitation method ([Selten \(1967\)](#)). Specifically, subjects are asked to directly design their strategy, which is then deployed to play the game in lieu of the subject (e.g. [Romero and Rosokha \(2018, 2019, 2023\)](#), [Dal Bó and Fréchette \(2019\)](#)). The second approach

(2008)).

⁹[Abreu and Rubinstein \(1988\)](#) define the complexity of a machine as the number of states of the minimal automaton implementing it, while [Banks and Sundaram \(1990\)](#) argue that the traditional number-of-states measure neglects some essential features, such as informational requirements at a state. They propose instead a criterion of complexity that takes into account both the size (number of states) and transitional structure of a machine. Under this proposition, they prove that the resulting Nash equilibria of the machine game are now trivial: the machines recommend actions in every period that are invariably stage-game Nash equilibria.

involves *indirect* inference of strategies from the observed realization of behavior (the so-called ‘direct-response’ method). The most common method of indirect inference is to use a finite mixture model to estimate the proportion of the population that uses a particular strategy (e.g. Dal Bó and Fréchette (2011), Fudenberg, Rand, and Dreber (2012), Vespa (2020)).¹⁰

The disadvantage of the indirect-inference approach is that, at the estimation stage, the researcher has to specify the set of all available strategies. The advantage of using the indirect-inference approach over the strategy-elicitation approach is that the researcher does not interfere nor restrict subjects’ behavior during the experiment. Indeed, many researchers warned of this caveat in the strategy-elicitation approach. Roth (1995) points out on p. 323 that “having to submit entire strategies forces subjects to think about each information set in a different way than if they could primarily concentrate on those information sets that arise in the course of the game.” Rand, Greene, and Nowak (2012) and Rand et al. (2014) corroborate Roth’s point as they find evidence that making people think deliberately lowers their cooperative behavior in the repeated Prisoner’s Dilemma and Public Goods games.¹¹ These reasons seemed compelling enough to lead us away from the strategy-elicitation method, and towards implementing a non-intrusive environment where subjects play the infinitely-repeated game directly without being asked to construct strategies.

3 Theory

3.1 Preliminaries

To simplify exposition, we start with some notation. The stage game is represented in standard strategic (normal) form. The set of players is denoted by $I = \{1, \dots, n\}$. Each player $i \in I$ has an *action set* denoted by $\mathcal{A}_i$. An *action profile* $a = (a_i, a_{-i})$ consists of the action of player i and the actions of the other players, denoted by

¹⁰Alternatively, Engle-Warnick and Slonim (2006) determine the best-fitting set of strategies based on a fitness function and an increasing cost for larger sets of strategies.

¹¹Brandts and Charness (2011) survey twenty-nine papers to investigate similarities and differences in the outcomes of the direct-response method versus the strategy-elicitation method. The authors find no difference in sixteen papers, find differences in four papers, and find mixed evidence in nine papers.

$a_{-i} = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n) \in \mathcal{A}_{-i}$. In addition, each player i has a real-valued, stage-game, payoff function $g_i : \mathcal{A} \rightarrow \mathbb{R}$, which maps every action profile $a \in \mathcal{A}$ into a payoff for i , where $\mathcal{A}$ denotes the cartesian product of the action spaces $\mathcal{A}_i$, written as $\mathcal{A} \equiv \prod_{i=1}^I \mathcal{A}_i$. In an infinitely-repeated game with perfect monitoring, the stage game in each time period $t = 0, 1, \dots$ is played with the action profile chosen in period t publicly observed at the end of that period. The *history* of play at time t is denoted by $h^t = (a^0, \dots, a^{t-1}) \in \mathcal{H}^t$ where $a^r = (a_1^r, \dots, a_n^r)$ denotes the actions taken in period r . The set of histories is given by

$$\mathcal{H} = \bigcup_{t=0}^{\infty} \mathcal{H}^t,$$

where we define the initial history to the null set $\mathcal{H}^0 = \{\emptyset\}$. A *strategy* $s_i \in S_i$ for player i is a function $s_i : \mathcal{H} \rightarrow \mathcal{A}_i$, where the strategy space of i consists of K_i discrete strategies; that is, $S_i = \{s_i^1, s_i^2, \dots, s_i^{K_i}\}$. Furthermore, we denote a strategy combination of the n players except i by $s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$. The set of joint-strategy profiles is denoted by $S = S_1 \times \dots \times S_n$. Each player i has a payoff function $\pi_i^t : S \rightarrow \mathbb{R}$, which represents the average payoff per period when the joint-strategy profile is played for t periods.

3.2 Moore Machines and Errors

The number of possible (repeated-game) strategies in a repeated game increases exponentially with the number of periods, making it impractical to deal with such a large number of strategies. We thus restrict the strategy space drastically by considering only strategies defined by finite-state automata called Moore machines (Moore (1956)). A *Moore machine* for player i , M_i , in an infinitely-repeated game $G = (I, \{S_i\}_{i \in I}, \{\pi_i\}_{i \in I})$ is a four-tuple $(Q_i, q_i^0, f_i, \tau_i)$ where Q_i is a finite set of internal states of which q_i^0 is specified to be the initial state, $f_i : Q_i \rightarrow \mathcal{A}_i$ is an output function that assigns an action to every state, and $\tau_i : Q_i \times \mathcal{A}_{-i} \rightarrow Q_i$ is the transition function that assigns a state to every state and other players' actions. This formalization for the transition function fits the natural description of a strategy as i 's plan of action in all possible circumstances that are consistent with i 's plans.¹²

¹²In contrast, the notion of a game-theoretic strategy for i requires the specification of an action for every possible history including those that are inconsistent with i 's plan of action. To formulate

Importantly, we assume that players' behaviors are limited by the constraints of their inherent nature.¹³ Such limitations have been treated by researchers under the notion of 'errors.' Consequently, we assume that the machines are subjected to some small likelihood of committing errors in the implementation of their actions à la Selten's trembling hand.¹⁴ Formally, a player i 's machine in an infinitely-repeated game G commits an implementation error with probability ϵ when, for any given state q , the output function returns the action $f_i(q)$ with probability $1 - \epsilon$ and draws another action " $f_i(q)$ " where $f_i(q) \neq f_i(q)$ otherwise.¹⁵

3.3 Equilibrium Machines

We study the players' behavior in the infinitely-repeated game G by analyzing the machine game where each player chooses and commits to a machine to play the game. In the following definition, we merely assume that each player's preferences are positively sensitive to the payoffs.

Definition 1. A *machine game* of the infinitely-repeated game G is a strategic game $G_{\mathcal{M}} = (I, \{\mathcal{M}_i\}_{i \in I}, \{\mathcal{P}_i\}_{i \in I})$ in which $\mathcal{M}_i$ is the set of all available machines of player i in the infinitely-repeated game, and $\mathcal{P}_i$ is the expected payoff of player i when the machines induce the sequence $(a^t(M_i, M_{-i}))_{t=1}^{\infty}$ of action profiles and $(q^t(M_i, M_{-i}))_{t=1}^{\infty}$ of states defined as follows: for i and $t \geq 1$, we have

- $q_i^1(M_i, M_{-i}) = q_i^0$
- $a_i^t(M_i, M_{-i}) = f_i(q_i^t(M_i, M_{-i}))$
- $q_i^{t+1}(M_i, M_{-i}) = \tau_i(q_i^t(M_i, M_{-i}), a_{-i}^t(M_i, M_{-i}))$.

The machine game is a strategic one; to incorporate Subgame Perfect equilibrium considerations, the solution concept needs to be modified to require that after every

the game-theoretic notion of a strategy, one would have to construct the transition function so that $\tau_i : Q_i \times \mathcal{A} \rightarrow Q_i$ instead of $\tau_i : Q_i \times \mathcal{A}_{-i} \rightarrow Q_i$.

¹³Conventional game theory rests on the foundation of hyper-rational agents with full ability to select the most-preferred action. Yet such an assumption is usually defended on its methodological merits rather than its empirical validity.

¹⁴For an experimental investigation of repeated Prisoner's Dilemma games in the presence of implementation errors see [Fudenberg, Rand, and Dreber \(2012\)](#).

¹⁵Similar approaches have been used in [Kraines and Kraines \(2000\)](#), [Brandt and Sigmund \(2005\)](#) and [Ioannou \(2014\)](#).

history in the repeated game, the machines constitute an equilibrium of the machine game.¹⁶ We next state formally our solution concept.

Definition 2. Machines (M_i^*, M_{-i}^*) constitute a Nash equilibrium in the machine game $G_{\mathcal{M}}$ of the infinitely-repeated game G , if for every player i ,

$$\mathcal{P}_i(M_i^*, M_{-i}^*) \geq \mathcal{P}_i(M'_i, M_{-i}^*)$$

for every machine M'_i of player i .¹⁷

3.4 A Proof-of-Concept Application

We proceed next to operationalize the aforementioned framework using the Markov process. To simplify exposition, let us consider a 2×2 repeated game.¹⁸ In addition, let us constrain further the set of Moore machines by imposing a bound of no more than two states in a machine, which gives us a total of 26 machines (see Figure 1). We denote the set of 26 machines by $\mathcal{M}$, omitting the subscript i given that the players' strategy sets are the same by assumption. Restricting our attention to the proposed strategy subspace is appropriate for two reasons. First, the specific subspace accommodates elements of bounded rationality as envisioned by [Simon \(1947\)](#). Indeed, in this context, bounded rationality suggests that a player may not consider all feasible strategies but limit herself to simpler ones.¹⁹ Second, the proposed subspace is still rich enough to allow for sophisticated strategic behavior, which incorporates not only

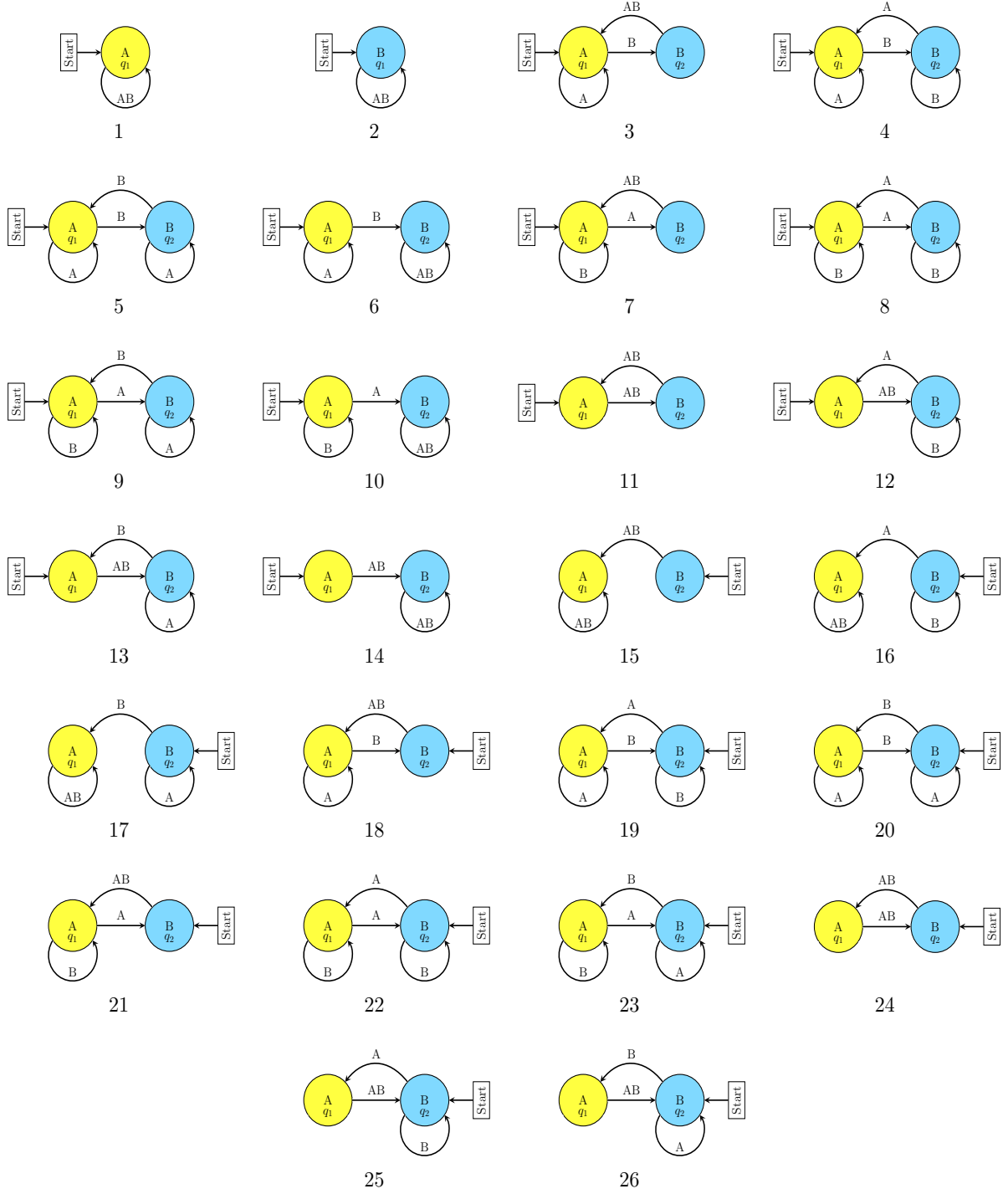
¹⁶As [Osborne and Rubinstein \(1994\)](#) point out, such a modification implies that the play of the machines does not have an introductory phase – only a cycling one. In the application in Subsection 3.4, where we apply the Markov process, the expected payoffs are independent of the initial condition, hence our notion of equilibrium can be extended to Subgame Perfection.

¹⁷In the application in Subsection 3.4, we restrict attention to pure equilibrium machine pairs in line with our theoretical framework, where a player chooses (once) in the beginning and commits thereafter to a machine; looking at mixed equilibrium machine pairs is beyond the scope of this exposition.

¹⁸An application of the Markov process on repeated, 2-player games with arbitrary actions can be found in [Stewart, Parsons, and Plotkin \(2016\)](#). An application of the Markov framework on repeated, n -player games with two actions can be found in [Hauert and Schuster \(1997\)](#), [Hilbe et al. \(2014\)](#), and [Hilbe et al. \(2017\)](#).

¹⁹A more complex machine is more likely to break down. It is also more difficult to learn and implement. For these reasons, a player values simpler machines. [Kuzmics, Palfrey, and Rogers \(2014\)](#) provide theoretical and empirical evidence to suggest that simplicity is a natural criterion for strategy choice.

Figure 1: ONE-STATE AND TWO-STATE MOORE MACHINES ($\mathcal{M}$)



punishments and triggers, but also anticipation of punishments and triggers. Let us also point out that the choice of a cap at two states finds support in the literature in evolutionary games and learning models (e.g. [Hanaki et al. \(2005\)](#), [Imhof, Fudenberg, and Nowak \(2007\)](#), [Ioannou and Romero \(2014\)](#) and [Dal Bó and Fréchet \(2018\)](#)).

In the 2×2 repeated game, each period of play leads to an action profile j ($j = 1, 2, 3, 4$): (A, A) , (A, B) , (B, A) , and (B, B) . Note that the first position denotes the action taken by player i and the second position that of player $-i$. The transition rules are labeled by quadruples $(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$ of zeros and ones, where σ_j is 1 if the (transition) rule plays A, and 0 if the rule plays B after profile j is realized. Thus, there are $2^4 = 16$ possible transition rules.²⁰ Furthermore, as highlighted earlier, we will assume that the machines are error-prone; let a machine have transition rule $\mathbf{p} = (p_1, p_2, p_3, p_4)$ where p_j is strictly positive denoting the probability of choosing A after the corresponding profile of the previous period. The space of all such rules is the four-dimensional unit cube; the corners are just the degenerate transition rules.

A transition rule $\mathbf{p} = (p_1, p_2, p_3, p_4)$ that is matched against a transition rule $\mathbf{q} = (q_1, q_2, q_3, q_4)$ yields a Markov process where the transitions between the four possible states²¹ are given by the matrix

$$\begin{pmatrix} p_1 q_1 & p_1(1 - q_1) & (1 - p_1)q_1 & (1 - p_1)(1 - q_1) \\ p_2 q_3 & p_2(1 - q_3) & (1 - p_2)q_3 & (1 - p_2)(1 - q_3) \\ p_3 q_2 & p_3(1 - q_2) & (1 - p_3)q_2 & (1 - p_3)(1 - q_2) \\ p_4 q_4 & p_4(1 - q_4) & (1 - p_4)q_4 & (1 - p_4)(1 - q_4) \end{pmatrix}.$$

If $\mathbf{p}$ and $\mathbf{q}$ are in the interior of the strategy cube, then all entries of this stochastic matrix are strictly positive; that is, the presence of errors ensures that the chain is ergodic. Consequently, there exists a unique stationary distribution $\phi^{\mathbf{p}/\mathbf{q}} = (\phi_1, \phi_2, \phi_3, \phi_4)$ such that $p_j^{(n)}$ is the probability of being in state j in the n^{th} period, and converges to ϕ_j for $n \rightarrow \infty$ ($j = 1, 2, 3, 4$). The components ϕ_j are strictly positive and sum up to 1. The stochastic vector ϕ is a left eigenvector of the matrix above for the eigenvalue 1. It follows that the expected payoff for player i using $\mathbf{p}$ against player $-i$ using $\mathbf{q}$ is given by

²⁰In addition, there are two possible initial conditions for a total of 32 transition rules. However, as indicated shortly, the unique stationary distribution and the expected payoffs are independent of the initial conditions.

²¹To be consistent with the conventional notation in Markov chains, the word ‘profile’ is replaced with the word ‘state.’

$$\mathcal{P}_i(\mathbf{p}, \mathbf{q}) = g_i(A, A)\phi_1 + g_i(A, B)\phi_2 + g_i(B, A)\phi_3 + g_i(B, B)\phi_4, \quad (1)$$

where the coefficients are the stage-game payoffs. Notice that ϕ_j and the expected payoffs are independent of the initial conditions.

The payoff obtained by a machine using transition rule $\mathbf{p}$ against a machine using transition rule $\mathbf{q}$ can be computed via (1). Computing the ergodic distribution of the realized actions for each machine pair enables us to obtain the 16×16 payoff matrix. However, if one wants to compute the limit value of the payoffs for $\epsilon = 0$, the aforementioned argument cannot be applied as the transition matrix is no longer irreducible hence the stationary distribution ϕ is no longer uniquely defined. Nevertheless, using the approach in Nowak, Sigmund, and El-Sedy (1995), one can show, without loss of generality, that when a machine using a transition rule $\mathbf{p}$ is paired with a machine using transition rule $\mathbf{q}$, the transitions *between* their stationary regimes occur with a frequency proportional to ϵ ; hence, do not affect the expected payoff as $\epsilon \rightarrow 0$. This allows us to also compute the 16×16 payoff matrix in the case of $\epsilon \rightarrow 0$. In the Supplementary Appendix, we provide (for expositional purposes) the payoff matrix for ‘Always-Defect’ (ALLD), ‘Tit-For-Tat’ (TFT), ‘Grim-Trigger’ (GT) and ‘Win-Stay, Lose-Shift’ (WSLS) in the presence of a *small* implementation error level affecting the interaction. We also provide the 16×16 payoff matrix in the case of $\epsilon \rightarrow 0$.

3.4.1 Moore Machines, Transition Rules and β -Machines

Let us conclude this section by clarifying the differences across some key components of our framework. In Subsection 3.4, we indicated that there are $2^4 = 16$ transition rules. Yet in Figure 1, we display 26 (one-state and two-state) Moore machines. This discrepancy between transition rules and machines is due to the inclusion in the machine definition of the initial state q_i^0 . However, recall further that the invariant distribution, and thereby the expected payoffs in the 16×16 payoff matrix, are independent of the initial conditions. This independence from the initial conditions makes the initial state in the formalization of a Moore machine redundant.

In this context, we introduce the notion of a *behavioral machine*²² (henceforth, for brevity, referred to as β -machine). Specifically, a β -machine is a three-tuple $(Q_i, f_i,$

²²Ioannou and Romero (2014) alternatively refer to them as rules of behavior.

τ_i) where Q_i , f_i and τ_i are defined as before. With this definition in place, we can identify 15 β -machines. In Figure 2, we provide the 15 β -machines. Each of the 15 β -machines is matched with a transition rule.²³ In addition, a β -machine (or transition rule) corresponds to either one or two Moore machines in the sense that the latter implement the *same* underlying behavior as the β -machine. In Figure 2, we also match the β -machines with the equivalent Moore machine(s) where the numerical labels of Moore machines are taken from Figure 1. For the reader’s convenience, we also include either the commonly-used name or a brief description of the respective behavior. Finally, notice that the 15 β -machines cover the entire $\mathcal{M}$ subspace. Henceforth, we will only be referring to β -machines.

4 Experimental Design & Theoretical Predictions

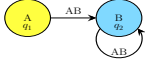
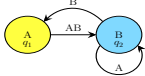
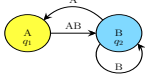
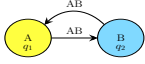
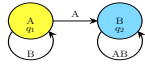
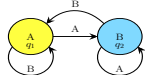
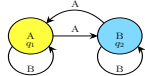
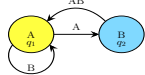
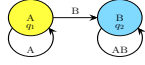
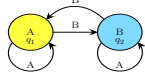
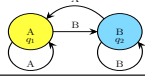
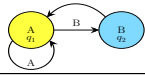
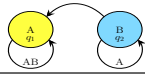
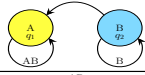
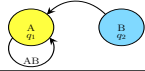
4.1 Stage Games

Our choice of games is not coincidental. We targeted the Prisoner’s Dilemma, Stag Hunt and Battle of the Sexes, where it is well documented that variations in their payoffs may influence behavior. Furthermore, these stage games cover a wide variety of strategic situations, characterized by strategic dominance, Pareto dominance and coordination motives.

The Prisoner’s Dilemma (PD) has become the workhorse to investigate social interactions. The PD is characterized by the relation $T > R > P > S$ where the letters arise from the generic payoff matrix in Table 1. Furthermore, it is often required that $2R > T + S$ so that alternations between cooperation and defection are not more profitable than joint cooperation. Standard equilibrium theory predicts that, in the single-shot PD game, the Pareto-dominated action profile of mutual defection will emerge. Adding more periods to the game does little to improve the theoretical outlook as long as the last period is common knowledge. If the last period is not known, subject to the assumptions of the Folk Theorem, nearly any path of play can be supported in equilibrium. Yet laboratory behavior is susceptible to both

²³Transition rule (S^{1100}) cannot be captured by a β -machine. This rule chooses A forever if action profiles (A, A) or (A, B) occur, and chooses B forever if (B, A) or (B, B) occur. Yet it cannot be depicted by a β -machine because its two (absorbing) states are disconnected.

Figure 2: EQUIVALENCIES

Transition Rule	β -Machine	Moore Machines	Description
S^{0000}		2,14	Quick A, then B
S^{0001}		13,26	BA or A Once
S^{0010}		12,25	BB or A Once
S^{0011}		11,24	Alternations
S^{0100}		10	AA Triggers B
S^{0101}		9,23	Tat-For-Tit
S^{0110}		8,22	Win-Shift,Lose-Stay
S^{0111}		7,21	AB or B Once
S^{1000}		6	Grim-Trigger
S^{1001}		5,20	Win-Stay,Lose-Shift
S^{1010}		4,19	Tit-For-Tat
S^{1011}		3,18	AA or B Once
S^{1101}		17	BB Triggers A
S^{1110}		16	BA Triggers A
S^{1111}		1,15	Quick B, then A

Notes: The first column indicates the transition rule, where each rule is denoted with its respective quadruple but without the commas. The second column displays the respective β -machine. The third column indicates the equivalent Moore machines in the sense that they implement the same underlying behavior as the corresponding β -machine, and the fourth column provides either the commonly-used name or a brief description of the behavior.

Table 1: GENERIC PRISONER’S DILEMMA PAYOFF MATRIX

	Cooperate	Defect
Cooperate	R	S
Defect	T	P

Notes: Players either *Cooperate* (C) or *Defect* (D). The payoffs are those of the row player. If both players choose ‘C,’ then each earns the *Reward* payoff (R). If both players choose ‘D,’ then each earns the *Punishment* payoff (P). If one player chooses ‘D’ and the other one chooses ‘C,’ then the former player earns the *Temptation* payoff (T) and the latter earns the *Sucker’s* payoff (S).

the relative payoffs and the horizon of the game (see Embrey, Fréchette, and Yuksel (2018), Dal Bó and Fréchette (2018)).²⁴

The Stag Hunt (SH) is a coordination game, which can be seen as a parable of social situations in which coordination can be pursued either to increase mutual rewards (which requires necessarily collaboration) or to earn lower rewards (which does not require collaboration). Using the letters from the generic PD matrix for convenience, the SH is characterized by the relation $T < R > P > S$. In SH games, typically, there is one mixed Nash equilibrium and two pure-strategy Nash equilibria that are Pareto-ranked: a payoff-dominant (or Pareto efficient) equilibrium and a risk-dominant equilibrium. Battalio, Samuelson, and Van Huyck (2001) indicate that, in experimental SH games, the sensitivity of subjects to the history of the other player’s play is greater in games with a larger optimization premium. Thus, the risk-dominant equilibrium is more likely to emerge the higher the optimization premium is.

The Battle of the Sexes (BoS) is another coordination game. It is characterized by the relation $T > S > R = P$. Similar to SH, there is one mixed Nash equilibrium and two pure-strategy Nash equilibria. However, in contrast to SH, the two pure-strategy Nash equilibria in this game create a tension between the two players due to

²⁴Dal Bó and Fréchette (2018) in their meta-analysis of PD games, reduce the number of payoff parameters to g (the gain from defection when the other player cooperates) and l (the loss from cooperation when the other player defects) to conclude that cooperation is decreasing in both g and l . Blonski, Ockenfels, and Spagnolo (2011) propose an axiomatic approach for equilibrium selection that revolves around providing conditions for when cooperation is more likely. Specifically, they characterize a selection criterion that results in a critical discount factor $\delta^* = \frac{T-S-R+P}{T-S}$, which is strictly greater than the minimum discount factor $\underline{\delta}$ required to support mutual cooperation. Blonski, Ockenfels, and Spagnolo (2011) show that when $\Delta\underline{\delta}$ and $\Delta\delta^*$ change in opposite directions, δ^* fares much better as a tool for predicting changes in the frequency of cooperation than the standard discount factor.

opposing preferences: each player earns more in a different equilibrium than the other. Nevertheless, repeated interactions allow players to reconcile their conflicting interests intertemporally. Specifically, theoretical analysis has shown that players can achieve the efficient outcome using an alternating strategy (see [Bhaskar \(2000\)](#), [Kuzmics, Palfrey, and Rogers \(2014\)](#)), which also finds compelling experimental support (e.g. [Cason, Lau, and Mui \(2013\)](#), [Ioannou and Romero \(2014\)](#), [Arifovic and Ledyard \(2018\)](#)).

Figure 3: STAGE GAMES

	A	B
A	32,32	12,50
B	50,12	25,25

(a) Prisoner's Dilemma (V1)

	A	B
A	48,48	12,50
B	50,12	25,25

(b) Prisoner's Dilemma (V2)

	A	B
A	32,48	12,50
B	50,12	25,25

(c) Prisoner's Dilemma (V3)

	A	B
A	45,45	0,35
B	35,0	40,40

(d) Stag Hunt (V1)

	A	B
A	45,45	0,42
B	42,0	12,12

(e) Stag Hunt (V2)

	A	B
A	45,45	0,42
B	35,0	40,12

(f) Stag Hunt (V3)

	A	B
A	0,0	9,15
B	15,9	0,0

(g) Battle of the Sexes (V1)

	A	B
A	3,3	9,17
B	20,10	3,3

(h) Battle of the Sexes (V2)

The payoff matrices of the PD games are indicated in the top row of Figure 3. The cooperative action is denoted with the letter 'A' and the action of defection

is denoted with the letter ‘B.’ Each player’s dominant action is to play ‘B.’²⁵ The payoff matrices of the SH games are indicated in the middle row. ‘A’ is a strict best response to any mixture that attaches a probability greater than 0.8 to ‘A,’ while ‘B’ is a strict best-response to any mixture attaching a lower probability to ‘A.’ In this game, there are two pure-strategy Nash equilibria $\{(A,A), (B,B)\}$ and a mixed Nash equilibrium in which ‘A’ is played with probability 0.8.²⁶ Action profile (A,A) is the payoff-dominant Nash equilibrium and (B,B) is the risk-dominant Nash equilibrium.²⁷ Finally, the payoff matrices of the BoS games, are indicated in the bottom row. The two pure-strategy Nash equilibria are $\{(A,B), (B,A)\}$. Each player receives a higher payoff in the equilibrium in which she plays ‘B.’ Alternating between the two pure-strategy Nash equilibria leads to the only Pareto optimal outcome. In the mixed Nash equilibrium, ‘A’ is played with probability 3/8 in the symmetric BoS (V1) game and, in the asymmetric BoS (V2) game, ‘A’ is played with probability 1/3 for the row player and with probability 6/23 for the column player.

Our intention is to provide a tight test of the theoretical framework proposed. We thus include an asymmetric variation in each class of games alongside the symmetric ones. The symmetric variations of the PD games are taken from [Dal Bó and Fréchette \(2011\)](#), and the symmetric ones of the SH games are taken from [Battalio, Samuelson, and Van Huyck \(2001\)](#). The asymmetric matrix in the PD and SH combines the two respective, symmetric variations. Both the symmetric variation and the asymmetric one of the BoS games are taken from [Arifovic and Ledyard \(2018\)](#).

The asymmetry in the payoffs of the PD game causes tension amongst two normative principles: efficiency and egalitarianism. This tension complicates the environ-

²⁵The minimum discount factor required to support mutual cooperation δ in a Subgame Perfect Equilibrium is 0.72 in PD (V1) and 0.083 in PD (V2). A similar conclusion (i.e. more cooperation in PD (V2)) is also drawn using the criterion of [Blonski, Ockenfels, and Spagnolo \(2011\)](#): the δ^* is $\frac{31}{38}$ in PD (V1) and $\frac{15}{38}$ in PD (V2).

²⁶[Dal Bó, Fréchette, and Kim \(2021\)](#) in their meta-analysis of SH, reduce the number of payoff parameters to Λ (the loss arising from an unilateral deviation from the efficient equilibrium) and λ (the loss arising from an unilateral deviation from the inefficient equilibrium). [Dal Bó, Fréchette, and Kim \(2021\)](#) find that as the size of the basin of attraction of ‘A’ increases and/or the optimization premium decreases, subjects become more likely to choose the payoff-dominant equilibrium. The basin of attraction is the maximum probability of the other player choosing ‘B’ that still makes ‘A’ a best response. Given these parameters, the basin of attraction of ‘A’ is equal to $\frac{\Lambda}{\Lambda+\lambda}$ and the optimization premium is equal to $\Lambda + \lambda$. Thus, in all our SH games, the basin of attraction of ‘A’ is 0.2, while in the SH (V1) game the optimization premium is 10, and in the SH (V2) game the optimization premium is $\frac{15}{33}$.

²⁷(A,A) is a risk-dominant profile if its basin of attraction is greater than 0.5, which is not the case in our selection of SH games.

ment considerably as players need to recognize and conform to the *preferred* principle of the other player. In contrast, in the symmetric PD variations, there exists no tension between efficiency and egalitarianism as the two are perfectly aligned. In SH, the specific asymmetric variation simplifies the environment as the efficient and payoff-dominant equilibrium is also the egalitarian one. As a result, (A,A) is salient relative to the other asymmetric profiles. Finally, the asymmetry in the payoffs in the BoS game inhibits egalitarianism even if players alternate on the pure-strategy Nash equilibria.²⁸ Nevertheless, in the symmetric BoS game, alternating between the two equilibria *does* lead to the only Pareto optimal and egalitarian outcome.

4.2 Experiments

The experiments were conducted on Amazon’s Mechanical Turk (henceforth, referred to as MTurk for brevity). Our initial intention was to conduct the experiments in a standard laboratory setting with university students. However, as it was only relatively recently that university laboratories had opened after the COVID lockdown measures, we felt that it was quite likely that the subject pool coming to the laboratory would be biased. Thus, we chose instead to conduct the experiments on MTurk.

Prior to running the actual experiments, we imposed a restriction that only US-based individuals could sign up for an experimental session. This requirement was essential to reduce the likelihood of dealing with individuals that were not proficient in English – the language of the experimental instructions. At the portal, participants were first asked to provide informed consent. Then, they had to go through the experimental instructions and complete a quiz to ensure their understanding of the game. The instructions emphasized that in order to proceed to the game-play stage,

²⁸In light of recent evidence documenting that egalitarianism and efficiency are strong attractors of behavior (see Ioannou et al. (2025)), one may wonder how many machine states are needed to attain an allocation that is both efficient and egalitarian in the asymmetric BoS game and, thereby, how restrictive is the choice of only one-state and two-state machines in the machine space. To attain an efficient and egalitarian allocation, a pair of machines would need to induce the action profiles $\{(A, B), (A, B), (A, B), (A, B), (A, B), (B, A), (B, A), (B, A), (B, A)\}$. Thus, a minimum of five states would be needed. The action-convergence criterion utilized allowed us to look at all possible patterns of reasonable length, including patterns that induce the above action profiles (for more details on the criterion, see Subsection 5.1). We did not locate a pattern longer than a length of two, which corroborates the findings in Arifovic and Ledyard (2018). Therefore, our choice to put a cap at two states was non-binding.

participants had to answer all quiz questions correctly.²⁹ Upon successful completion of the quiz, participants entered the waiting room where the matching took place. Once two participants were successfully matched, a game from Figure 3 was randomly assigned to them and the pair started their repeated interaction. The values in the payoff matrices denote Experimental Currency Units (ECUs). Participants played one infinitely-repeated game with perfect monitoring and complete information. In the experiment, the two actions of the stage games ‘A’ and ‘B’ (in Figure 3) were labeled as ‘X’ and ‘Y.’ Furthermore, ‘X’ and ‘Y’ were randomized so that, for instance, (X, X) was the Nash equilibrium in some PDs and (Y, Y) was the Nash equilibrium in others. Each player’s payoff in the repeated game consisted of the sum of the ECUs accumulated from each period. The exchange rate, 500 ECUs for \$1, was common knowledge. Initially, for their participation in the experiment, they received a payment of 2,500 ECUs (i.e. \$5). Upon completion of the repeated game, participants were asked to provide some demographic information. The experimental instructions and the quiz questions are provided in the Supplementary Appendix.

We used a high continuation probability (i.e. $\delta = 0.99$) to induce long interactions (in expectation 100 periods) between the same two participants. The termination rule was common knowledge. The actual draw for the length of the repeated games was done *ex ante*. The drawn length of the repeated games was 96. We fixed the length to make consistent comparisons across games without invoking variability in learning and/or fatigue, which could confound our results.

Participants were allowed to participate in only one session. The sessions lasted around 45 minutes. Average earnings per participant were \$9.80. In total, there were 516 participants that formed 258 datapoints across the 8 games. Table 2 displays the number of participants who played in each game.

4.3 Experimental Implementation Vis-a-Vis Theoretical Framework

The theoretical framework assumes that players commit from the beginning of the game to one-state or two-state machines, and locates the Nash equilibrium pairs based on asymptotic-payoff calculations. Strategic behavior is inherent in the Nash

²⁹Participants were told that repeated failure to answer the questions correctly would result in their dismissal.

Table 2: CHARACTERISTICS OF THE SESSIONS

Game	Acronym	# of Participants
Prisoner’s Dilemma (V1)	PD1	64
Prisoner’s Dilemma (V2)	PD2	68
Prisoner’s Dilemma (V3)	APD3	70
Stag Hunt (V1)	SH1	62
Stag Hunt (V2)	SH2	64
Stag Hunt (V3)	ASH3	66
Battle of the Sexes (V1)	BoS1	60
Battle of the Sexes (V2)	ABoS2	62

Notes: The Table displays the number of participants who played the corresponding game. A datapoint consists of a pair of participants (hence the number of datapoints is one half of the above numbers).

equilibrium solution concepts. Underlying strategic behavior are the assumptions of rationality and higher-order rationality.³⁰ In the context of the experimental setting, these assumptions seem quite stringent and far too unrealistic to be satisfied by subjects (at least) from the very beginning of the infinitely-repeated game. However, with *sufficient* experience, subjects are more likely to exhibit strategic behavior in the game in line with the underlying rationality assumptions. In particular, once subjects accumulate enough experience, they can better understand the dynamics of the game (e.g. the long horizon, the implications of their actions, best responding to beliefs, the strategy set) and can thus foresee the outcome of their game play based on their own plan of action and the anticipated plan of action of the other player.

The mapping between the theoretical framework and the experimental implementation begins *only after* subjects accumulated sufficient experience with the infinitely-repeated game. At that point, ‘strategic’ subjects based on the knowledge accrued and introspection decide on how to play the game; that is, *lock* their plan of action

³⁰A player is rational if she best-responds to her beliefs. If she believes others are rational, and if she believes others believe others are rational and so on and so forth, then she satisfies higher-order rationality.

for the remainder of the game. Therefore, we (implicitly) partition the phase of accumulating experience and the phase where subjects commit/lock to a particular game plan, and focus *exclusively* on the last phase which parallels the theoretical setting.

Our approach finds support in many studies. Consider first the study of [Romero and Rosokha \(2018\)](#), where the authors allow subjects to construct strategies by developing a set of rules (i.e. ‘if this, then that’ statements). Romero and Rosokha divide the infinitely-repeated game into two stages: the ‘free’ stage and the ‘lock’ stage. In the free stage, subjects are able to edit their rule set as the game progresses and their understanding of the game’s dynamics improves, whereas in the lock stage, subjects’ rule sets are locked for the remainder of the game. Along similar lines, studies in the literature that model learning in repeated games, separate the so-called ‘pre-experimental’ phase from the ‘experimental’ one (see e.g. [Hanaki, Sethi, Erev, and Peterhansl \(2005\)](#), [Ioannou and Romero \(2014\)](#)); that is, partition the inexperienced, initial behavior of players in the infinitely-repeated game and the experienced behavior after accruing knowledge of the game’s dynamics. The lock stage in [Romero and Rosokha \(2018\)](#) and the experimental phase in the learning literature parallel our last phase where subjects commit to a particular game plan.

4.4 Theoretical Predictions

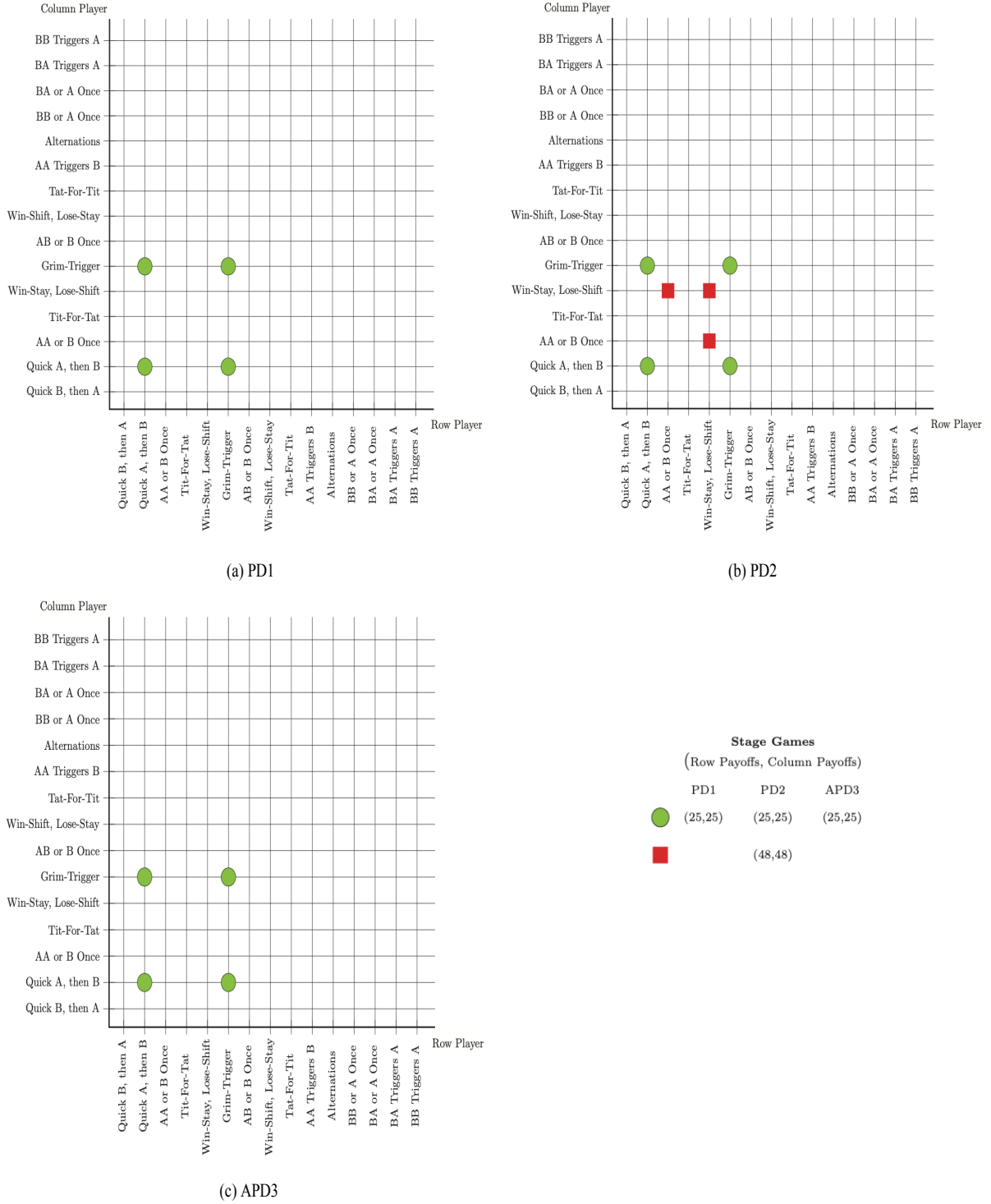
In Figures 4-6, we display the 15×15 possible β -machine combinations in the PD, SH and BoS games, respectively. The legend of the respective figures indicates the expected payoff in equilibrium of the row and column players across all variations. In the Supplementary Appendix, we contrast the predicted payoffs in the proposed framework with the theoretical payoff predictions of the Folk Theorem.

Prisoner’s Dilemma Games

In Figure 4, we display the equilibrium pairs and (expected) payoffs across the three variations. Noticeably, the predictions in all relevant spaces (i.e. payoffs, machine pairs) are quite sharp.

In PD1, we observe four equilibrium pairs. The expected payoff in all combinations is 25 for both the row and column players as the pairs induce plays of (B, B) (i.e. defecting plays). In PD2, we retain the four equilibrium pairs from PD1 that play the

Figure 4: PREDICTED MACHINE PAIRS IN THE PRISONER'S DILEMMA GAMES



Notes: The Figure provides the predicted pairs across the three payoff variations of the Prisoner's Dilemma.

(B, B) action profile. However, given that, in this variation, the gain from defection when the other player cooperates is very small, three equilibrium machines inducing plays of (A, A) (i.e. cooperative plays) resulting in an expected payoff of 48 for both players also arise. Thus, in this PD variation, we should expect both cooperative and defecting plays. Moving into an asymmetric environment that combines the payoffs from the previous two PD variations is detrimental to the possibility of cooperative plays. In APD3, we observe that the predictions with respect to equilibrium payoffs and pairs are identical to those in Figure 4(a).

We formulate next our two PD predictions.

Prediction 1. *In PD1 and APD3, plays within a game consist of (B, B) .*

Prediction 2. *In PD2, plays within a game consist of either (A, A) or (B, B) .*

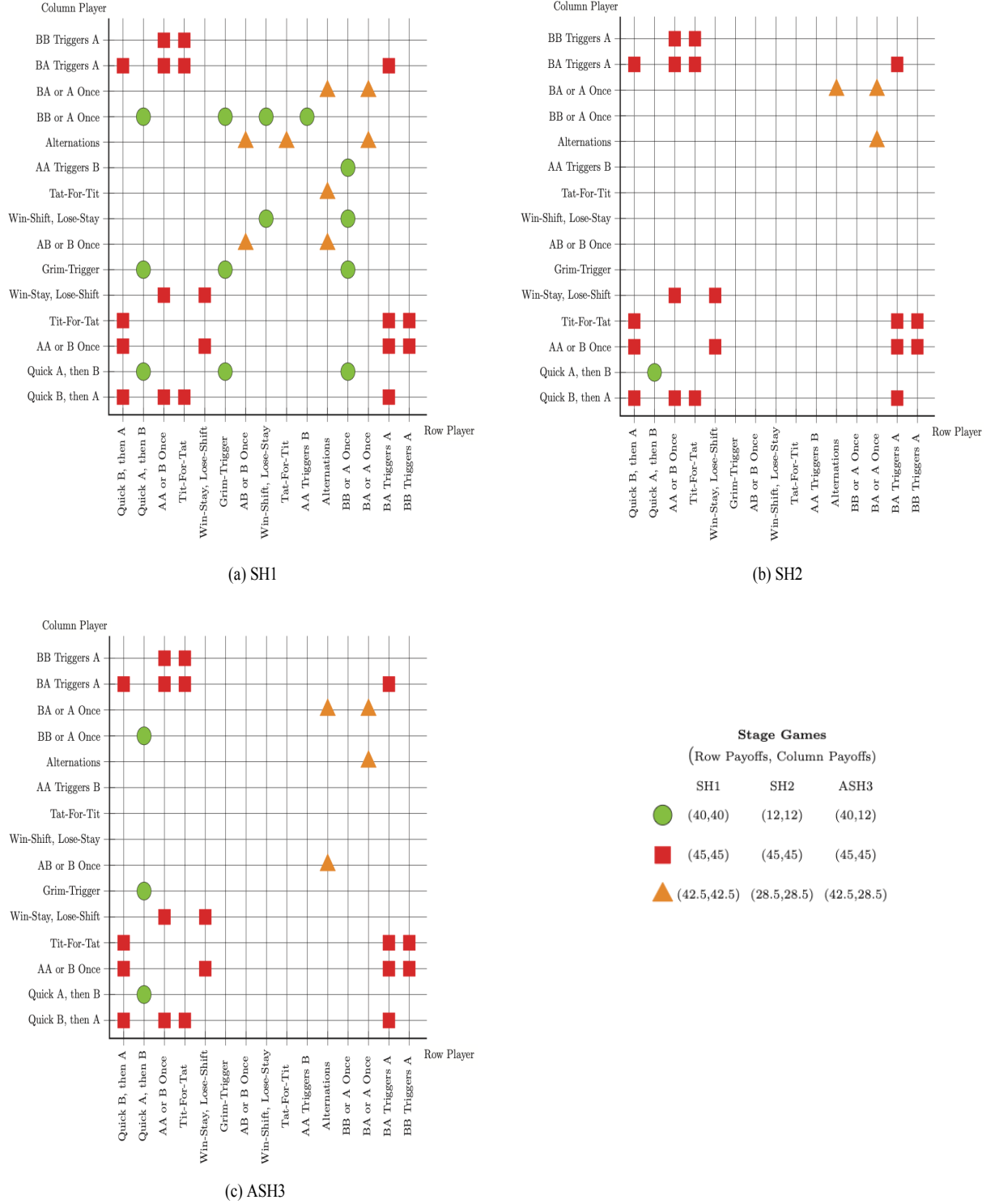
Stag Hunt Games

In Figure 5, we provide the equilibrium pairs and payoffs in the SH games. In sharp contrast to the PD games, in the SH games, machine predictability is compromised. In all SH games, we observe three classes of equilibrium pairs. In the first class, we find machines that play (A, A) and thus achieve the Pareto efficient payoff of 45. Furthermore, in all three variations, the set of machines in this first class is identical. In the second class, we find equilibrium pairs that play the risk-dominant action profile (B, B) . Finally, in the third class, we find machines that alternate between (A, A) and (B, B) . Importantly, the volume of machines included in the latter two classes differs from variation to variation.

In SH1, we observe a very large number of equilibrium pairs. The first class consists of nineteen equilibrium pairs that settle on (A, A) . In the second class, we find equilibrium pairs ‘Quick A, then B,’ ‘Grim-Trigger’ and others that settle on the risk-dominant payoff of 40. Finally, in the third class, equilibrium machines alternate between the payoff-dominant and risk-dominant equilibrium, thus securing an expected payoff of 42.5 for each player.

In SH2, predictions are sharper relative to SH1. Though the first class of equilibrium pairs is identical, in the second class that includes machines coordinating at the risk-dominant equilibrium, we find only one β -machine: ‘Quick A, then B.’ Furthermore, in the third class, we have two β -machines: ‘Alternations’ and ‘BA or A

Figure 5: PREDICTED MACHINE PAIRS IN THE STAG HUNT GAMES



Notes: The Figure provides the predicted pairs across the three payoff variations of Stag Hunt.

Once,’ which produce the three orange triangles in the figure.

In ASH3, predictability is not as sharp as that in SH2. The first class is again identical to the other two variations, but the second class is different from both. In addition to ‘Quick A, then B’ which is an equilibrium choice for both row and column players, we also have another two β -machines that can be utilized by the column player to form an equilibrium pair when matched with the row player’s ‘Quick A, then B:’ ‘BB or A Once’ and ‘Grim-Trigger.’ Finally, in the third class, in addition to the equilibrium pairs in SH2, we have one more equilibrium pair: ‘Alternations’ for the row player and ‘AB or B Once’ for the column player. The asymmetric expected payoffs are 42.5 and 28.5 for the row player and column player, respectively.

In the SH games, we formulate one prediction.

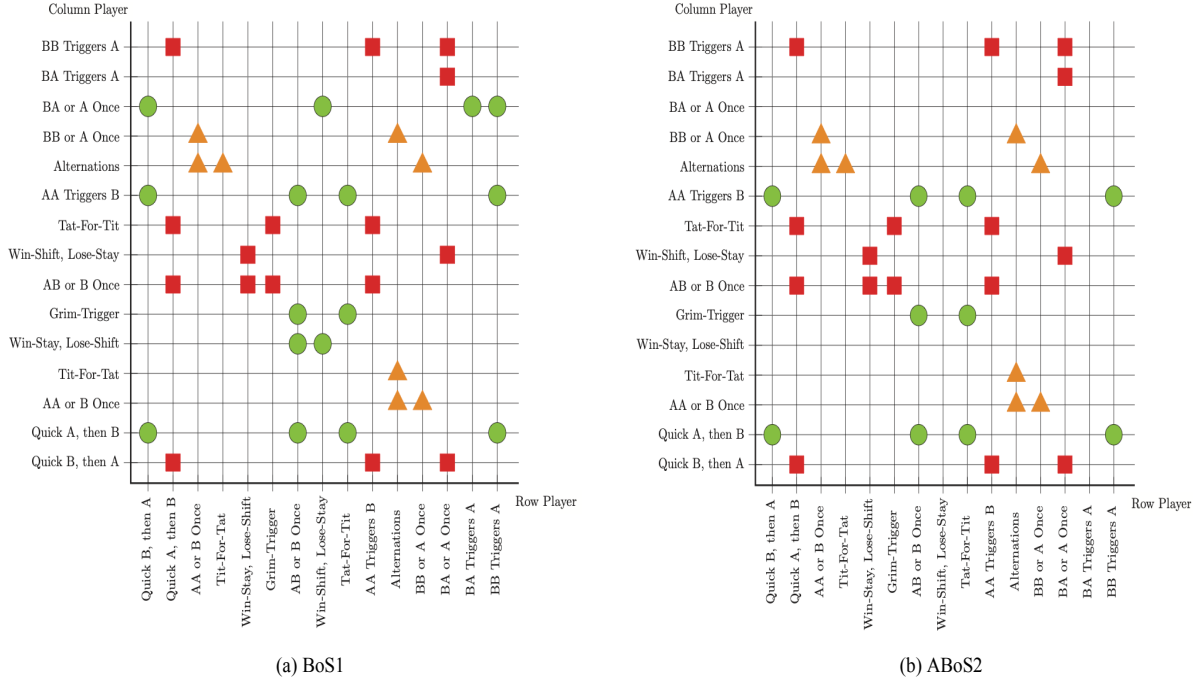
Prediction 3. *In all three SH variations, plays within a game consist of either (A, A) or (B, B) or alternations between (A, A) and (B, B) .*

Battle of the Sexes Games

In Figure 6, we provide the equilibrium pairs and payoffs in the two variations of the BoS games. Given that this is another coordination game, similar to the SH games, machine predictability is compromised. This is particularly true in BoS1 where all machines are equilibrium predictions. Machine predictability is improved in ABoS2 but only slightly. Similar to the SH games, there exist three classes of equilibrium pairs. In the first class, we find machines that coordinate on the row player’s not-preferred Nash equilibrium (i.e. (A, B)), whereas in the second class, we find machines that coordinate on the row player’s preferred equilibrium (i.e. (B, A)). In the third class, we find machines that alternate between the two pure-strategy Nash equilibria of the stage game thus achieving Pareto efficient payoffs. The machines in the second and third classes are identical across the two variations.

In BoS1, in the first class, there are many equilibrium machines resulting in payoffs (9, 15). In the second class, we find, for example, row player’s machines ‘Quick A, then B,’ ‘AA Triggers B’ and ‘Grim-Trigger’ that are matched with column player’s machine ‘Tat-for-Tit.’ The payoffs in this class are (15, 9). It is important to note that given that the payoff matrix is symmetric, machines in the first two classes are

Figure 6: PREDICTED MACHINE PAIRS IN THE BATTLE OF THE SEXES GAMES



Stage Games	
(Row Payoffs, Column Payoffs)	
BoS1	ABoS2
● (9,15)	(9,17)
■ (15,9)	(20,10)
▲ (12,12)	(14.5,13.5)

Notes: The Figure provides the predicted pairs across the two payoff variations of the Battle of the Sexes.

interchangeable.³¹ In the third class, we find ‘Alternations,’ ‘BB or A Once,’ ‘Tit-For-Tat’ and others that result in alternations between the two pure-strategy Nash equilibria of the stage game. In this third class, the payoff of each player is 12.

In ABoS2, the machines in the first class constitute a proper subset of those in the first class of the symmetric variation. The payoffs in this first class are (9,17). The machines in ABoS2 in the second and third classes are identical to those in the symmetric variation. The payoffs in the second class are (20,10) and, in the third

³¹In BoS1, we forewent putting all thirty-two points in the same color to illustrate that the red squares (i.e. (B, A)) in BoS1 are retained in ABoS2.

class, the payoff of the row player is 14.5 and the payoff of the column player is 13.5.

The last prediction pertains to the BoS games.

Prediction 4. *In the two BoS variations, plays within a game consist of either (A, B) (i.e. the row player’s not-preferred Nash equilibrium) or (B, A) (i.e. the row player’s preferred Nash equilibrium) or alternations between (A, B) and (B, A) .*

5 Findings

We first extract frequent patterns from our dataset using the action-convergence criterion from Ioannou et al. (2025). We then contrast the frequent patterns mined with the predicted game plays and the action profiles induced by the predicted machine pairs. Finally, we provide a statistical procedure to evaluate the theoretical model proposed across our games.

5.1 Frequent Pattern Mining

Recently, Ioannou et al. (2025) documented that in games with long interactions between the same participants, oftentimes participants reach a predictable way of playing the game. The authors propose an action-convergence criterion to extract frequent patterns.³² Following their approach, a pattern p is said to be $(1 - \phi)$ -frequent over interval T , if p minimizes the frequency of errors while allowing at most a frequency ϕ of errors over T . Analogously, we say that a sequence $(1 - \phi)$ -converges to p (or is $(1 - \phi)$ -convergent) over T , if it admits a $(1 - \phi)$ -frequent pattern over T . Otherwise, the sequence is $(1 - \phi)$ -divergent. Similar to Ioannou et al. (2025), here, we also consider the last twenty periods of game play and allow a maximum of two errors within the specified interval; that is, $\phi = \frac{2}{20} = 0.1$.³³

The results are displayed in Table 3. Overall, we observe that there are 193 pairs (sequences of game play) that (0.9) -converged (henceforth, for brevity, we omit $(1 - \phi)$ - which is always the same) to a pattern out of a possible 258 pairs; this corresponds

³²The interested reader should refer to the paper for the technical details.

³³In the Supplementary Appendix, we include robustness checks on both the number of errors (1, 2 and 3) and intervals of game-play periods (20 and 25 periods), which confirm the insensitivity of the results to these reasonable choices of errors and period intervals.

Table 3: FREQUENT PATTERNS MINED

	PD1	PD2	APD3	SH1	SH2	ASH3	BoS1	ABoS2
Data Points	32	34	35	31	32	33	30	31
Convergent	24	28	22	25	29	26	20	19
	3	24	3	16	26	22		
	21	4	19	9	3	4		
							7†	3
								6
							13	10
Divergent	8	6	13	6	3	7	10	12

Equal Payoffs
 Pareto Efficient

Notes: Table 3 displays the number of convergent sequences in the last twenty periods based on the action-convergence criterion proposed in Ioannou et al. (2025). The limit for errors is two. The columns indicate the stage games, whereas the rows indicate the frequent patterns mined. Frequent patterns that are Pareto efficient are displayed with a purple background, whereas egalitarian ones are displayed with a border. The † (next to 7) is to highlight that given that the matrix is symmetric, pattern (A,B) and pattern (B,A) in the following row are interchangeable.

to a 75% convergence. The highest percentage of convergence is found in SH2 with 91% (29 pairs out of 32), and the lowest percentage of convergence is found in ABoS2 with 61% (19 pairs out of 31). Strikingly, the 75% overall convergence observed corresponds to only five patterns: (A,A), (B,B), (A,B), (B,A) and ((A,B),(B,A)). Frequent patterns (A,A) and (B,B) are detected in the PD and SH games, whereas (A,B), (B,A) and ((A,B),(B,A)) are mined in the BoS games. Furthermore, it is important to note that there are changes in subjects' behavior for payoff variations

within a given class of games. For instance, in PD1, there are 3 pairs that converged to pattern (A,A) (i.e. the cooperative profile) and 21 pairs that converged to pattern (B,B) (i.e. the defecting profile), whereas in PD2, there are 24 pairs that converged to pattern (A,A) and only 4 pairs that converged to pattern (B,B). Finally, almost all of the frequent patterns mined reflect a payoff across players that is either efficient or egalitarian or both. This result corroborates the findings in Ioannou et al. (2025) who document that Pareto efficiency and egalitarianism are strong attractors of behavior.

5.2 Theory vs. Behavior: Game Plays

We next compare the frequent patterns mined to the equilibrium game plays.³⁴ Here, we hit a big obstacle. Standard, statistical hypothesis testing cannot be used as a single pair settling on a non-compliant (to the equilibrium predictions) play is a sufficient datapoint to provide evidence against the theoretical prediction.³⁵ In Subsection 5.4, we will provide a formal statistical procedure that assesses the closeness (in a sense explained there) of the empirical distribution to the theoretical one in each game. But for now, we will settle to a comparison, by sheer observation, of the frequent patterns mined in Table 3 with the theoretical predictions. We note that each prediction is matched with the corresponding result.

Prediction 1 states that in PD1 and APD3, equilibrium plays within a game consist of (B, B) . In Table 3, we see that in these two games, with the exception of 3 pairs in each game that converged to the cooperative play (A, A) , the rest converged to the defecting play (B, B) in line with the theoretical prediction.

Result 1. *In PD1 and APD3, plays within a game consist almost entirely of (B, B) .*

Prediction 2 pertains to the PD2 game. The theoretical framework predicts either plays of (A, A) or (B, B) . As shown in Table 3, all 28 pairs that converged to a pattern indeed chose either (A, A) or (B, B) .

Result 2. *In PD2, plays within a game consist of either (A, A) or (B, B) .*

³⁴We thus exclude from our analysis sequences of game play that do not correspond to frequent patterns.

³⁵More specifically, standard hypothesis testing cannot be used when the H_0 is that plays follow a distribution that puts all probability on just the equilibrium play(s) because a single datapoint, which is inconsistent with the H_0 (and hence carries zero probability of occurrence under the H_0), is sufficient to reject H_0 for any (positive) level of significance.

Prediction 3 states that, in all three SH variations, plays within a game consist of (A, A) or (B, B) or alternations between (A, A) and (B, B) . We see in Table 3 that there is no single pair that alternated between (A, A) and (B, B) ; nevertheless, all other pairs did converge to either (A, A) or (B, B) in line with the theoretical prediction.

Result 3. *There is no evidence to support alternations between (A, A) and (B, B) in any of the SH games. We find that in all three SH variations, plays within a game consist of either (A, A) or (B, B) .*

Finally, Prediction 4 prescribes that plays within a game consist of either (A, B) (i.e. the row player’s not-preferred Nash equilibrium) or (B, A) (i.e. the row player’s preferred Nash equilibrium) or alternations between (A, B) and (B, A) . Indeed, the results in Table 3 show that all frequent patterns mined were consistent with the theoretical prediction.

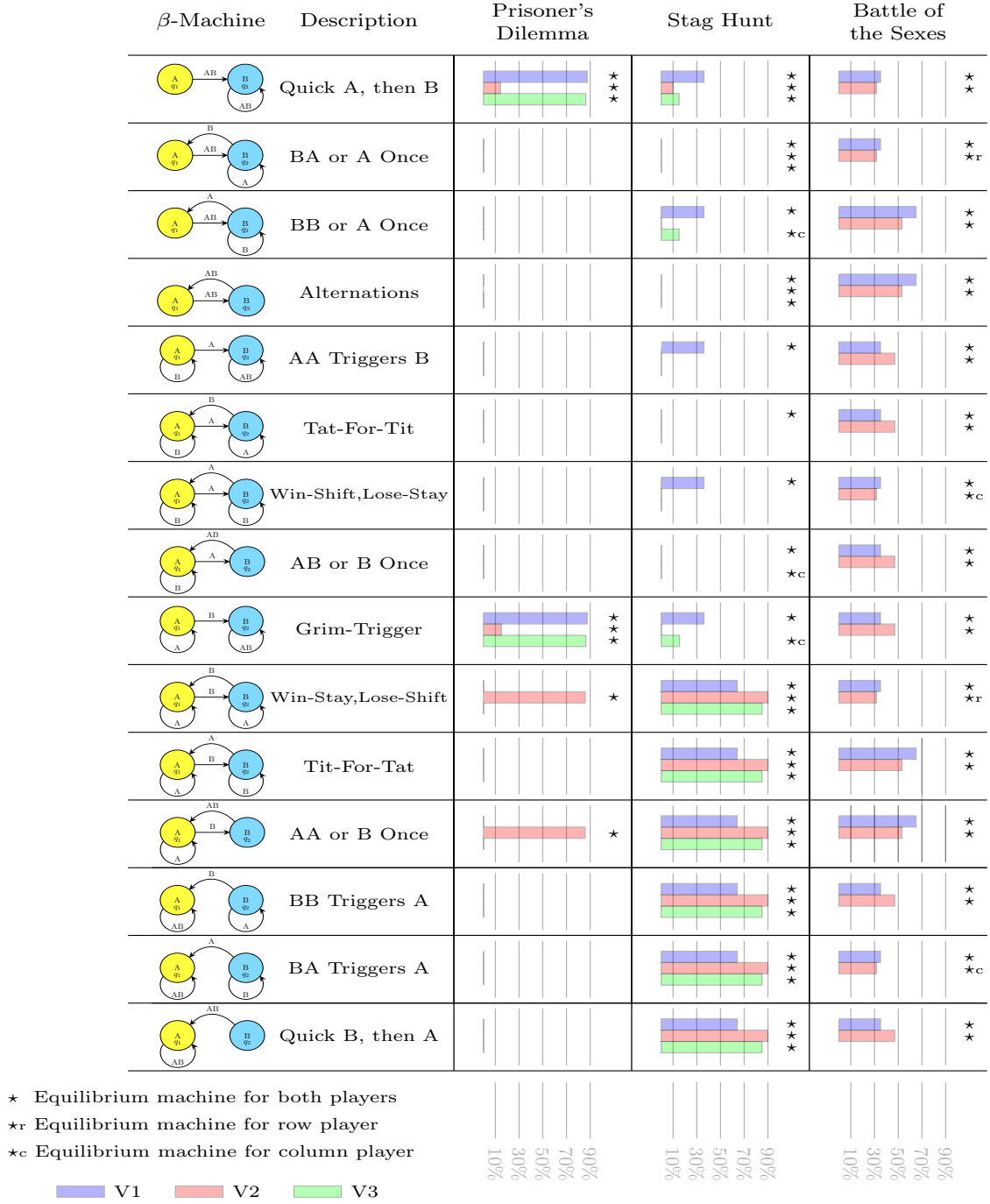
Result 4. *In the two BoS variations, plays within a game consist of either (A, B) or (B, A) or alternations between (A, B) and (B, A) .*

5.3 Theory vs. Behavior: Game Machines

We compare here the frequent patterns mined from the experimental data with the action profiles induced by the equilibrium machines. However, such a comparison is inhibited by a serious hurdle. A frequent pattern extracted reflects a *single* finite history; therefore, there are many different machines that can generate it. For example, pattern (A, A) can be generated by ‘Grim-Trigger’ and ‘Win-Stay, Lose-Shift’ or ‘Tit-For-Tat’ and ‘Quick B, then A.’ To overcome this problem, we use a simple approach: we determine what percentage of the frequent patterns detected in each game *could* have been generated by the specific predicted β -machine. Consequently, we are not so much concerned with whether there are many machines that could fit to the frequent patterns detected – this is a corollary of the approach utilized. We are more interested whether the action profiles induced by the equilibrium machines find support in the data. The results are displayed in Figure 7. For the reader’s convenience, we also include a $\star$ in the bar chart to denote an equilibrium machine.

As mentioned earlier, machine predictions are quite sharp in the PD games. Recall that in PD1 and APD3, the predicted outcome is that pairs will end up defecting.

Figure 7: FREQUENT PATTERNS MINED AND EQUILIBRIUM MACHINES



Notes: The first column corresponds to the β -machine. The second column provides either the commonly-used name or a brief description of the behavior. The remaining three columns (one for each class of games) show with bar charts the percentage of frequent patterns that could have been generated by the specific predicted β -machine in each game variation.

Moreover, β -machines ‘Quick A, then B’ and ‘Grim-Trigger’ are predicted to deliver this stream of defections. Subjects’ behavior in PD1 and APD3, predominantly settled on the defecting profile. In PD1, 21 out of 24 pairs converged to a pattern of defections (i.e. 87.5%), whereas in APD3, 19 out of 22 pairs converged to a defecting pattern (i.e. 86%). These percentages are shown in the bar charts of ‘Quick A, then B’ and ‘Grim-Trigger’ for PD1 and APD3, respectively. In PD2, the predicted outcome is that pairs will either end up defecting or cooperating. In addition to the two aforementioned β -machines (i.e. ‘Quick A, then B’ and ‘Grim-Trigger’), ‘Win-Stay, Lose-Shift’ and ‘AA or B Once’ also constitute equilibrium ones. The latter two machines, deliver the stream of cooperative profiles. In fact, ‘Win-Stay, Lose-Shift’ and ‘AA or B Once’ capture 86% of the frequent patterns detected.

In the SH games, the predicted outcomes reflect one of the following three scenarios: (i) playing the payoff-dominant equilibrium (A, A) , (ii) playing the risk-dominant equilibrium (B, B) , and (iii) alternating between the payoff-dominant and risk-dominant equilibria $((A, A), (B, B))$. In SH1, all β -machines are supported in equilibrium. In SH2, the number of equilibrium machines is drastically reduced as a few machines delivering alternations between the payoff-dominant and risk-dominant equilibria and a few ones that settle on the risk-dominant equilibrium are no longer supported. In ASH3, in addition to the equilibrium machines from SH2, some of the equilibrium machines from SH1 are supported, but only for the column player. Recall that in all SH games, pairs either ended up at the payoff-dominant equilibrium or at the risk-dominant equilibrium. More generally, all equilibrium machines across variations find support in the data with one exception: we do not find any support for the predicted alternating sequence between the payoff-dominant and risk-dominant equilibria.

In the BoS games, the predicted plays either have a player choosing their preferred pure-strategy Nash equilibrium or the two players alternating between the two pure-strategy Nash equilibria of the stage game. Similar to the SH games, in the BoS games, almost all machines can be supported in equilibrium. Specifically, in BoS1, all β -machines are supported in equilibrium, whereas in ABoS2, two β -machines are supported in equilibrium for the row player, another two are supported in equilibrium for the column player, while the remaining β -machines are supported in equilibrium for both players. The results here are even more encouraging than those in the SH games. Recall that in the two BoS variations, subjects either alternated between

the two pure-strategy Nash equilibria or settled on one of the two pure-strategy Nash equilibria. We observe that all equilibrium machines deliver either the former behavior or the latter.

5.4 A Statistical Procedure to Evaluate the Theoretical Model

In Subsection 5.2, we highlighted that standard, statistical hypothesis testing cannot be used on any of our games as even one pair of players choosing a non-compliant (to the equilibrium predictions) play provides enough evidence to challenge the theoretical prediction. To address this issue, we modify the statistical procedure of Ioannou and Makris (2019) to assess the closeness (in a sense defined shortly) of the empirical distribution to the theoretical one. The modified approach thus allows us to determine the games where the theoretical model performs (relatively) better.

This statistical procedure assumes that pairs' behavior may be influenced by iid *behavioral errors* due to factors 'outside the theoretical model' (e.g. short attention spans, random distractions). To clarify, in the context of our games, we assume that the total number of non-compliant (to the equilibrium predictions) plays is a random variable $Z \sim \text{Bin}(N, \pi)$, where $\text{Bin}(N, \pi)$ is the binomial distribution with size parameter N (i.e. the number of pairs),³⁶ and 'success probability' π . The key assumption here is that $\pi \sim \text{Beta}(a, b)$ for some exogenously given scalars $a \geq 1, b > 1$, where $\text{Beta}(a, b)$ is the Beta distribution with parameters a, b . To investigate whether pairs' behavior is *close* to the theoretical prediction of zero non-compliant (to the equilibrium predictions) plays and get a sense of how close it is, we will examine whether pairs' behavior aligns with the H_0 that $Z \sim \text{Bin}(N, \pi)$ with $\pi \sim \text{Beta}(a, b)$ for predetermined values of a and b , paying particular attention to determine the set of values of (a, b) for which the H_0 is rejected.

To conduct our statistical procedure, we will use the total number of non-compliant plays Z as our statistic from Table 3. We therefore need to derive the probability distribution of Z under the above H_0 . It is shown in Ioannou and Makris (2019) that the probability distribution of statistic Z under the above H_0 is $\text{Bin}(N, \frac{a}{a+b})$. Denoting with $p(z, \frac{a}{a+b})$ the p -value that corresponds to the realization z of the statistic Z , it follows that the H_0 is rejected given a realization z of the statistic and a level of

³⁶The size parameter N thus includes both the convergent as well as the divergent sequences. We do so to obtain a higher bound on the threshold mean errors.

significance $\alpha \in (0, 1)$ if

$$p(z; \frac{a}{a+b}) := \sum_{j=z}^N \binom{N}{j} (\frac{a}{a+b})^j (\frac{b}{a+b})^{N-j} < \alpha. \quad (2)$$

We know that $p(0; \pi) = 1$ for all $\pi \in [0, 1]$. We thus have that the H_0 cannot be rejected following a realization $z = 0$ of the statistic for any mean error $\frac{a}{a+b}$. Turning to the case of a realization $z > 0$ of the statistic, the magnitude of the mean error is crucial in rejecting or not the H_0 . However, without additional information about pairs' behavior, the specific choice of the mean error is arbitrary. With this in mind, we now shift the focus to asking how large the range of mean error values must be for a given realization $z > 0$ to result in a rejection of the H_0 . We know that $p(z; \pi)$ is increasing in π for all $z \in \{1, \dots, N\}$, with $p(z; 1) = 1$ and $p(z; 0) = 0$. Therefore, there is a critical threshold for the mean error below which the H_0 is rejected, whereas for all (weakly) higher mean errors the H_0 is not rejected. This threshold mean, denoted hereafter by $\underline{\pi}(z)$, is in $(0, 1)$ and given implicitly by

$$\sum_{j=z}^N \binom{N}{j} (\underline{\pi}(z))^j (1 - \underline{\pi}(z))^{N-j} = \alpha. \quad (3)$$

Note now that $p(z, \pi)$ is decreasing in z for all $\pi \in (0, 1)$. Consequently, we also have that $\underline{\pi}(z)$ is increasing in z . It follows that the smaller the realization $z > 0$ of the statistic, the *closer* the observed behavior is to the theoretical prediction; that is, the lower the realization of the statistic, the 'less often' the above H_0 will be rejected.

The threshold mean errors are displayed on Table 4. We observe that the range of threshold mean errors is between 0.0260 and 0.3116. The lowest value is in SH2 where the predictions are well in line with the observed behavior. The highest value is in APD3 where three pairs played the cooperative action profile (instead of the predicted defecting one) and thirteen pairs did not converge to a pattern. Evidently, payoff asymmetry provides an additional layer of complexity in all game types. Overall, the theoretical model seems to perform the best in the SH games.

Table 4: THRESHOLD MEAN ERRORS

H_0 : $Z \sim \text{Bin}(N, \pi)$ with $\pi \sim \text{Beta}(a, b)$ for predetermined values of $a \geq 1, b > 1$	
<i>PD Games</i>	
PD1	0.2061
PD2	0.0797
APD3	0.3116
<i>SH Games</i>	
SH1	0.0878
SH2	0.0260
ASH3	0.1040
<i>BoS Games</i>	
BoS1	0.1933
ABoS2	0.2407

Notes: The level of statistical significance α is fixed at 5%.

6 Concluding Remarks

In the proof-of-concept application, we consider one-state and two-state machines, and use asymptotic theory in the machine game to derive long-term predictions in payoff variations of the Prisoner’s Dilemma, Stag Hunt and Battle of the Sexes. Furthermore, we conduct experiments with a high continuation probability and contrast the theoretical behavior with the empirical one. We find that the proposed theoretical framework does fairly well overall in delivering the experimental behavior. We discuss next some of the key findings of our study. Wherever possible, we also relate our findings to those in relevant studies.

We start off with a discussion about the machines (strategies) inferred in the PD games. As mentioned earlier, there exists a recent literature that applies strategy elicitation directly to infer subjects’ strategies in the PD games. Though the continuation probability is different in these studies from ours, the PD payoff matrices

used are identical.³⁷ Dal Bó and Fréchette (2019) find that subjects choose common strategies, such as ‘Tit-For-Tat,’ ‘Grim-Trigger’ and ‘Always Defect.’ Romero and Rosokha (2018) corroborate the findings of Dal Bó and Fréchette (2019). In Romero and Rosokha (2023), the authors find that even though many subjects do use mixed strategies in the beginning of the PD game, strategies tend to become less mixed over time, and move towards three focal pure-strategies: ‘Tit-For-Tat,’ ‘Grim-Trigger’ and ‘Always Defect.’ We reach similar conclusions. Figure 7 shows that predicted ‘Grim-Trigger’ and ‘Quick A, then B’ (i.e. ‘Always Defect’) explain a very high percentage of the frequent patterns mined in two of the three payoff variations in PD games and a lower percentage of frequent patterns in the third payoff variation.

We make two additional remarks. First, in the PD games, we *do* find machines with defecting absorbing states. This corroborates existing literature. Specifically, Dal Bó and Fréchette (2011) provide experimental evidence to suggest that even in treatments where cooperation can be supported in equilibrium (like our selected variations), the level of cooperation may remain at low levels even after significant experience is obtained. The authors conclude that “these results cast doubt on the common assumption that agents will make the most of the opportunity to cooperate whenever it is possible to do so in equilibrium” (p. 412). Second, Blonski, Ockenfels, and Spagnolo (2011) find that in PD games, the mean value of cooperation in other experimental studies (see their Table 6) is 0.43 if $\delta^* < \delta$, which is comparable to our mean value of cooperation in the three variations of 0.49.

In the SH games, a few of the predicted machines (e.g. ‘Tit-For-Tat,’ and ‘BA Triggers A’) explain more than 60% of the frequent patterns mined in the payoff variations. Importantly, we find that people do not necessarily coordinate on the payoff-dominant equilibrium, which corroborates the findings in Dal Bó, Fréchette, and Kim (2021). More specifically, in SH1, a large number of pairs end up playing the risk-dominant equilibrium. The fact that payoff dominance is not used by these pairs as an equilibrium selection criterion adds to the existing evidence suggesting that strategic uncertainty is important in coordination games. Though our payoff matrices are directly taken from Battalio, Samuelson, and Van Huyck (2001), the notable differences in their matching protocol and horizon with ours does not allow us to make any head-on comparisons between the two studies.³⁸ Yet we still find that, similar to

³⁷Dal Bó and Fréchette (2019) use the same symmetric PD payoff matrices we use here (i.e. PD1 and PD2) with $\delta = 0.5, 0.75, 0.9$. Romero and Rosokha (2018, 2023) also use PD1 with $\delta = 0.95$.

³⁸Battalio, Samuelson, and Van Huyck (2001) use a random matching protocol with 75 periods of

Battalio, Samuelson, and Van Huyck (2001), the risk-dominant equilibrium is more likely to emerge the larger is the optimization premium. Finally, we do not find any evidence of the predicted alternating sequence between the payoff-dominant and risk-dominant equilibria in the SH games. Clearly, this is an unwarranted prediction that is neither supported in the experimental data here nor, to the best of our knowledge, in the existing literature.³⁹

In the BoS games, predictability is compromised as almost all machines can be supported in equilibrium. Nevertheless, all predicted machines find support in the data. Specifically, as shown in Figure 7, all predicted machines can explain *at least* 30% of the frequent patterns mined in the two payoff variations of the BoS games. Furthermore, given that the BoS payoff matrices are taken from Arifovic and Ledyard (2018), despite their experimental games being finitely repeated, we can still compare (somewhat) our findings to theirs.⁴⁰ Arifovic and Ledyard (2018) find that in the symmetric game, 40% of the pairs alternate between the two pure-strategy Nash equilibria, 30% settle on one of the pure-strategy Nash equilibria, and the other 30% are confused or contentious. We find that in the symmetric game, 43% of the pairs alternated between the two pure-strategy Nash equilibria, 24% settled on one of the two pure-strategy Nash equilibria, and 33% did not converge to a pattern. Therefore, the percentages found in the BoS1 are quite similar to theirs. Finally, Arifovic and Ledyard (2018) find that the asymmetry in the payoffs led to an increase in the percentage of pairs ending at the pure-strategy Nash equilibrium, and a decrease in the percentage of those alternating between the two pure-strategy Nash equilibria. Both trends are observed in our data as well.

Concluding, we hope that in the near future, similar studies will be carried across other infinitely-repeated games to confirm the ability of the framework to capture well long-term behavior.

game play.

³⁹Mathevet (2018) in his axiomatization of plays in repeated games provides a rather weak axiom that eliminates the predicted alternating sequence in our SH games. Specifically, the *efficient simplicity* axiom states that a pair would only complexify their convention, if in doing so, at least one of the players earns more; clearly, in all versions of our SH games, the alternating sequences are strictly Pareto inferior to the simpler payoff-dominant equilibrium.

⁴⁰The length of game play in Arifovic and Ledyard (2018) is 40 periods and the matching protocol is fixed. All this information is common knowledge.

References

- Abreu, Dilip, and Ariel Rubinstein. “The Structure of Nash Equilibrium in Repeated Games with Finite Automata.” *Econometrica* 56: (1988) 1259–82.
- Arifovic, Jasmina, and John O. Ledyard. “Learning to Alternate.” *Experimental Economics* 21, 3: (2018) 692–721.
- Banks, Jeffrey S., and Rangarajan K. Sundaram. “Repeated Games, Finite Automata, and Complexity.” *Games and Economic Behavior* 2: (1990) 97–117.
- Battalio, Raymond, Larry Samuelson, and John Van Huyck. “Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games.” *Econometrica* 69, 3: (2001) 749–64.
- Bhaskar, V. “Egalitarianism and Efficiency in Repeated Symmetric Games.” *Games and Economic Behavior* 32, 2: (2000) 247–62.
- Binmore, Kenneth G., and Larry Samuelson. “Muddling Through: Noisy Equilibrium Selection.” *Journal of Economic Theory* 74, 2: (1997) 235–65.
- . “Evolutionary Drift and Equilibrium Selection.” *The Review of Economic Studies* 66, 2: (1999) 363–93.
- Blonski, Matthias, Peter Ockenfels, and Giancarlo Spagnolo. “Equilibrium Selection in the Repeated Prisoner’s Dilemma: Axiomatic Approach and Experimental Evidence.” *American Economic Journal: Microeconomics* 3, 3: (2011) 164–92.
- Brandt, Hannelore, and Karl Sigmund. “Indirect Reciprocity, Image Scoring, and Moral Hazard.” *Proceedings of the National Academy of Science* 102, 7: (2005) 2666–70.
- Brandts, Jordi, and Gary Charness. “The Strategy Versus the Direct-Response Method: A First Survey of Experimental Comparisons.” *Experimental Economics* 14, 3: (2011) 375–98.
- Carlsson, Hans, and Eric van Damme. “Global Games and Equilibrium Selection.” *Econometrica* 61, 5: (1993) 989–1031.

- Cason, Timothy N., Sau-Him Paul Lau, and Vai-Lam Mui. “Learning, Teaching, and Turn Taking in the Repeated Assignment Game.” *Economic Theory* 54: (2013) 335–57.
- Cooper, David J. “Supergames Played by Finite Automata with Finite Costs of Complexity in an Evolutionary Setting.” *Journal of Economic Theory* 68, 1: (1996) 266–75.
- Dal Bó, Pedro, and Guillaume R. Fréchette. “The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence.” *American Economic Review* 101, 1: (2011) 411–29.
- . “On the Determinants of Cooperation in Infinitely Repeated Games: A Survey.” *Journal of Economic Literature* 56, 1: (2018) 60–114.
- . “Strategy Choice in the Infinitely Repeated Prisoner’s Dilemma.” *American Economic Review* 109, 11: (2019) 3929–52.
- Dal Bó, Pedro, Guillaume R. Fréchette, and Jeongbin Kim. “The Determinants of Efficient Coordination: Experimental Evidence from Stag Hunt Games.” *Games and Economic Behavior* 130: (2021) 352–68.
- Ellison, Glenn. “Learning, Local Interaction, and Coordination.” *Econometrica* 61, 5: (1993) 1047–71.
- . “Learning from Personal Experience: One Rational Guy and the Justification of Myopia.” *Games and Economic Behavior* 19: (1997) 180–210.
- Embrey, Matthew, Guillaume R. Fréchette, and Sevki Yuksel. “Cooperation in the Finitely Repeated Prisoner’s Dilemma.” *Quarterly Journal of Economics* 133, 1: (2018) 509–51.
- Engle-Warnick, Jim, and Robert Slonim. “Inferring Repeated-Game Strategies from Actions: Evidence from Trust Game Experiments.” *Economic Theory* 28, 3: (2006) 603–32.
- Friedman, James E. “A Non-Cooperative Equilibrium for Supergames.” *The Review of Economic Studies* 38, 1: (1971) 1–12.

- Fudenberg, Drew, and Eric Maskin. “Evolution and Repeated Games.”, 1993. Mimeo.
- Fudenberg, Drew, David G. Rand, and Anna Dreber. “Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World.” *American Economic Review* 102, 2: (2012) 720–49.
- Hanaki, Nobuyuki, Rajiv Sethi, Ido Erev, and Alexander Peterhansl. “Learning Strategies.” *Journal of Economic Behavior and Organization* 56: (2005) 523–42.
- Harsanyi, John, and Reinhard Selten. *A General Theory of Equilibrium Selection in Games*. The MIT Press, 1988.
- Hauert, Christoph, and Gerhard Schuster. “Effects of Increasing the Number of Players and Memory Size in the Iterated Prisoner’s Dilemma: A Numerical Approach.” *Proceedings of the Royal Society B* 264, 1381: (1997) 513–9.
- Hilbe, Christian, Luis A. Martinez-Vequeiro, Krishnendu Chatterjee, and Martin A. Nowak. “Memory-n Strategies of Direct Reciprocity.” *Proceedings of the National Academy of Science* 114, 18: (2017) 4715–20.
- Hilbe, Christian, Bin Wu, Arne Traulsen, and Martin A. Nowak. “Cooperation and Control in Multiplayer Social Dilemmas.” *Proceedings of the National Academy of Science* 111, 46: (2014) 16,425–30.
- Imhof, Lorens A., Drew Fudenberg, and Martin A. Nowak. “Tit-For-Tat or Win-Stay, Lose-Shift?” *Journal of Theoretical Biology* 247, 3: (2007) 574–80.
- Ioannou, Christos A. “Asymptotic Behavior of Strategies in the Repeated Prisoner’s Dilemma Game in the Presence of Errors.” *Artificial Intelligence Research* 3, 4: (2014) 28–37.
- . “Coordination Games: Escaping the Straitjacket.” In *Handbook of Experimental Finance*, edited by Sascha Fullbrunn, and Ernan E. Haruvy. Edward Elgar Publishing, 2022, 150–8.
- Ioannou, Christos A., and Miltiadis Makris. “An Experimental Study of Uncertainty in Coordination Games.” *International Economic Review* 60, 2: (2019) 751–99.
- Ioannou, Christos A., Laurent Mathevet, Julian Romero, and Huanren Zhang. “Data Mining in Repeated Games.”, 2025. Mimeo.

- Ioannou, Christos A., and Julian Romero. “A Generalized Approach to Belief Learning in Repeated Games.” *Games and Economic Behavior* 87: (2014) 178–203.
- Jéhiel, Philippe. “Limited Foresight May Force Cooperation.” *The Review of Economic Studies* 68: (2001) 369–91.
- Kandori, Michihiro, George J. Mailath, and Rafael Rob. “Learning, Mutation, and Long Run Equilibria in Games.” *Econometrica* 61, 1: (1993) 29–56.
- Kraines, David P., and Vivian Y. Kraines. “Natural Selection of Memory-One Strategies for the Iterated Prisoner’s Dilemma Game.” *Journal of Theoretical Biology* 203: (2000) 335–55.
- Kuzmics, Christoph, Thomas R. Palfrey, and Brian W. Rogers. “Symmetric Play in Repeated Allocation Games.” *Journal of Economic Theory* 154: (2014) 25–67.
- Makris, Miltiadis. “Complementarities and Macroeconomics: Poisson Games.” *Games and Economic Behavior* 62: (2008) 180–9.
- Mannila, Heikki, Hannu Toivonen, and Inkeri Verkamo. “Discovery of Frequent Episodes in Event Sequences.” *Data Mining and Knowledge Discovery* 1: (1997) 259–89.
- Mathevet, Laurent. “An Axiomatization of Plays in Repeated Games.” *Games and Economic Behavior* 110: (2018) 19–31.
- McKelvey, Richard, and Thomas R. Palfrey. “Playing in the Dark: Information, Learning, and Coordination in Repeated Games.”, 2001. Mimeo.
- Moore, Edward F. “Gedanken Experiments on Sequential Machines.” *Annals of Mathematical Studies* 34: (1956) 129–53.
- Morris, Stephen, and Hyun Song Shin. “Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks.” *American Economic Review* 88, 3: (1998) 587–97.
- . “Global Games: Theory and Applications.” In *Proceedings of the 8th World Congress of the Econometric Society*, edited by M. Dewatripont, L. Hansen, and S. Turnovsky. Cambridge University Press, 2003.

- Nikiforakis, Nikos, Charles N. Noussair, and Tom Wilkening. “Normative Conflict and Feuds: The Limits of Self-Enforcement.” *Journal of Public Economics* 96, 9: (2012) 797–807.
- Nowak, Martin A., Karl Sigmund, and Esam El-Sedy. “Automata, Repeated Games, and Noise.” *Journal of Mathematical Biology* 33: (1995) 703–22.
- Osborne, Martin J., and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, 1994.
- Piccione, Michele. “Finite Automata Equilibrium with Discounting.” *Journal of Economic Theory* 56: (1992) 189–93.
- Rand, David G., Joshua D. Greene, and Martin A. Nowak. “Spontaneous Giving and Calculated Greed.” *Nature* 489, 7416: (2012) 427–30.
- Rand, David G., Alexander Peysakhovich, Gordon T. Kraft-Todd, George E. Newman, Owen Wurzbacher, Martin A. Nowak, and Joshua D. Greene. “Social Heuristics Shape Intuitive Cooperation.” *Nature Communications* 5, 3677: (2014) 1–14.
- Romero, Julian, and Yaroslav Rosokha. “Constructing Strategies in the Indefinitely Repeated Prisoner’s Dilemma Game.” *European Economic Review* 104: (2018) 185–219.
- . “The Evolution of Cooperation: The Role of Costly Strategy Adjustments.” *American Economic Journal: Microeconomics* 11, 1: (2019) 299–328.
- . “Mixed Strategies in the Indefinitely Repeated Prisoner’s Dilemma.” *Econometrica* 91, 6: (2023) 2295–331.
- Roth, Alvin E. “Bargaining Experiments.” In *Handbook of Experimental Economics*, edited by John Kagel, and Alvin E. Roth. Princeton: Princeton University Press, 1995, 253–348.
- Selten, Reinhard. “Die Strategiemethode zur Erforschung des Eingeschränkt Rationalen Verhaltens im Rahmen eines Oligopolexperiments.” In *Beiträge zur experimentellen Wirtschaftsforschung*, edited by Heinz Sauermann. Tübingen: J. C. B. Mohr, 1967, 136–68.

- . “A Re-Examination of the Perfectness Concept for Equilibrium Points in Extensive Games.” *International Journal of Game Theory* 4: (1975) 25–55.
- Simon, Herbert. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations*. The Free Press, 1947.
- Sonsino, Doron, and Julia Sirota. “Strategic Pattern Recognition - Experimental Evidence.” *Games and Economic Behavior* 44, 2: (2003) 390–411.
- Stahl, O. Dale. “Boundedly Rational Rule Learning in a Guessing Game.” *Games and Economic Behavior* 16: (1996) 303–30.
- . “Evidence Based Rules and Learning in Symmetric Normal-Form Games.” *International Journal of Game Theory* 28: (1999) 111–30.
- Stahl, O. Dale, and Ernan Haruvy. “Between-Game Rule Learning in Dissimilar Symmetric Normal-Form Games.” *Games and Economic Behavior* 74: (2012) 208–21.
- Stewart, Alexander J., Todd L. Parsons, and Joshua B. Plotkin. “Evolutionary Consequences of Behavioral Diversity.” *Proceedings of the National Academy of Science* 113, 45: (2016) E7003–9.
- Tirole, Jean. *The Theory of Industrial Organization*. MIT Press, 1988.
- Turocy, Theodore L. “A Dynamic Homotopy Interpretation of the Logistic Quantal Response Equilibrium Correspondence.” *Games and Economic Behavior* 51: (2005) 243–63.
- Vespa, Emanuel. “An Experimental Investigation of Strategies in the Dynamic Common Pool Game.” *International Economic Review* 61, 1: (2020) 417–40.
- Zhang, Boyu, and Josef Hofbauer. “Quantal Response Methods for Equilibrium Selection in 2×2 Coordination Games.” *Games and Economic Behavior* 97: (2016) 19–31.