Incremental Learning
Equilibrium selection in repeated games

HUANG Zhe, JIN Yuren, LIU Sibo

March 19, 2024

1 Introduction

In this report, we present an insightful article by **Sam Jindani**, titled **Equilibrium Selection in Repeated Games**[2]. This study stands at the intersection of economics and social sciences, leveraging game theory as a pivotal analytical tool for dissecting the intricate web of behaviors exhibited by rational individuals in scenarios necessitating interactive decision-making. Infinite repeated games, a core focus of this investigation, offer a robust theoretical scaffold to examine the perpetual influence individuals exert on one another's strategic choices across endless interactions.

In the realms of economics and social sciences, game theory is an essential tool for understanding the patterns of behavior among rational individuals in interactive decision-making. Specifically, infinite repeated games provide a theoretical framework for analyzing how individuals influence each other's choices through continuous interactions. These games theoretically admit multiple equilibria, each corresponding to different strategy combinations and potential outcomes. However, individuals often face the challenge of selecting the optimal or most appropriate strategy from these possible equilibria in practice. This selection issue has sparked extensive discussion in theory and has profound implications for economic policy formulation and social behavior norms in practice.

Although the literature on game theory is replete with studies on equilibrium selection, most theories still discuss this issue within a static framework, overlooking individuals' learning and adaptation behavior in actual decision-making processes. With the development of behavioral economics and experimental economics, researchers have begun to pay attention to how individuals adjust their strategy choices through the learning and information updating process in historical interactions. However, there is still a significant research gap in how individuals select among many possible equilibria through a dynamic learning process in infinite repeated games and how this learning mechanism affects the stability and convergence of equilibria.

This author's research aims to fill this gap by developing and analyzing a new learning rule that guides individuals to select specific equilibria in infinite repeated games. Particularly, this paper introduces a learning rule that combines theories from behavioral economics and dynamic game strategies, aiming to simulate how individuals adjust their expectations and behaviors through continuous interaction and feedback. Moreover, we specifically focus on how this learning rule leads to the selection of an equilibrium corresponding to the Kalai–Smorodinsky bargaining solution and analyze the changes in equilibrium stability and individual welfare during this process.

2 Theoretical Background

2.1 Kalai and Lehrer's[3] Interactive Learning Problem

In the study of repeated games, players do not make decisions in isolation but form expectations about their opponents' strategies through continuous interactions and respond accordingly. A key issue in this process is how each player adjusts their strategy based on their opponent's behavior, especially when they have limited information about their opponent's strategy. The work of Kalai and Lehrer[3] focuses on this so-called "interactive learning" problem, exploring how players learn from each other in interactions and ultimately reach consistent behavior patterns.

In their seminal analysis in 1993, Kalai and Lehrer[3] proposed a model to analyze the learning process of players in infinite repeated games. They assumed that each player holds certain "beliefs" (i.e., expectations) about their opponent's possible strategies at the beginning of the game and chooses their best strategy based on these beliefs. As the game progresses, players update their beliefs about their opponent's strategy according to the actual behavior observed.

Kalai and Lehrer's model emphasizes the importance of the "grain-of-truth" assumption, meaning that players' initial beliefs must be compatible with the equilibrium states the game may eventually reach. Although players may not fully understand their opponent's strategy initially, their beliefs contain enough information that, over time, through learning and strategy adjustments, players can accurately predict each other's behavior and reach a Nash equilibrium.

2.2 Foster and Young's[1] Near-Rational Learning

2.2.1 Core Theory and Method

Foster and Young, in their 2003 study, explored a learning model distinct from that of Kalai and Lehrer, focusing on how players, equipped with limited information and processing capabilities, approximate the equilibrium of games through near-rational behavior. This model particularly highlights the "near-rational" actions of players, where, in most instances, attempts are made to respond optimally, but non-optimal choices may occur due to reasons like misunderstandings, misjudgments, or constrained computational capacity.

In the model proposed by Foster and Young[1], players form hypotheses about their opponents' strategies based on limited-memory beliefs and provide ε-best responses to these hypotheses, allowing for a degree of deviation. This approach acknowledges the imperfections in individual decision-making and the realities of making choices with limited information and computational abilities.

2.2.2 Dynamics of Near-Rational Learning

- Hypothesis Formation and Testing: Players continuously develop hypotheses about their opponent's strategy based on past behavior and test the accuracy of these hypotheses through observation of actual behavior. Significant discrepancies between observed actions and expected hypotheses lead players to adjust or change their beliefs.
- ε-Best Response: When making decisions, players aim for the best strategy but, due to the nature of near-rationality, this choice encompasses a margin of error. This permits the exploration of potential strategies, reflecting the uncertainty in individual decision-making in reality.
- Convergence to Equilibrium: Foster and Young[1] demonstrated that under their learning rule, players eventually tend to converge to a subgame-perfect equilibrium of the repeated game. This indicates that, even in the presence of errors and near-rational behavior, continuous learning and strategy adjustment can stabilize the system's dynamics at an equilibrium state.

2.3 The Kalai–Smorodinsky[4] Bargaining Solution

The literature on learning for repeated-game strategies often omits the question of equilibrium selection. Addressing this, the presented model of near-rational learning in two-player, infinitely repeated games yields precise selection results. For any given stage game, the employed learning rule selects a subgame-perfect equilibrium corresponding to the repeated game in which the payoffs received by each player align closely with the Kalai–Smorodinsky[4] bargaining solution of the payoff space established in the seminal work of Kalai and Smorodinsky (1975) [4]. Kalai and Smorodinsky's solution is as follow:

The foundational assumptions of the model assume that players develop their beliefs grounded in empirical evidence, they exhibit rationality with significant probability—optimizing decisions based on these beliefs—and the learning mechanisms at play are uncoupled, implying no necessity for players to understand their opponent's payoff structure. Such assumptions are not only plausible but also central to the model's predictive power.

The dynamics of the model revolve around players forming and discarding bounded-memory beliefs based on observed behaviors in play. A belief is maintained if it is consistently validated by the opponent's actions and discarded upon frequent conflict, leading to a period of uncertainty and exploration for new strategic beliefs. This iterative belief-adjustment process—bounded by memory and defined by best responses, with allowances for human error—culminates in states that diverge from subgame-perfect equilibrium being inherently unstable.

This model underscores the parallels between traditional bargaining dilemmas and the strategic complexities of equilibrium selection in repeated games. Both contexts address multifaceted potential outcomes with varying implications for player payoffs, focusing on how players navigate and ultimately divide the surplus. The seamless translation of the bargaining solution into repeated games signifies a profound analytical symmetry, wherein the principles of equitable and efficient outcomes inherent in one domain are aptly applicable to another.

	C	D
C	$2, 2$	$-1, 3$
D	$5, -1$	$0, 0$

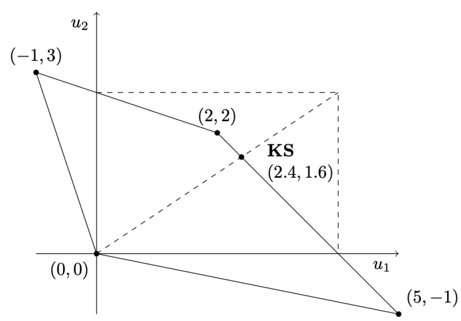

Figure 1: Game matrix and payoff space for the asymmetric prisoner's dilemma, demonstrating the Kalai–Smorodinsky solution.

For instance, the asymmetric Prisoner's Dilemma can be reinterpreted as a bargaining scenario where the set of feasible and individually rational payoffs can be visualized within the bargaining solution framework. The Kalai–Smorodinsky[4] solution, situated on the Pareto frontier, ensures an equitable allocation of payoffs, even in the context of asymmetry, granting a larger share to the player with a higher maximal payoff.

3 Problem Setting

3.1 Infinite Repeated Games

The strategic landscape of infinite repeated games is markedly distinct from that of their finite counterparts. In infinite repeated games, the choices made in each round of the game not only affect the outcome of the current round but also have potential implications for strategies and outcomes in all future rounds. This complexity requires players to consider long-term strategies over immediate gains, making the decision-making process considerably more intricate.

Consider the classic Prisoner's Dilemma game, where two prisoners are given the choice to either "Cooperate" with each other by remaining silent or "Defect" by betraying the other. The payoffs for a single round of the game are as follows: If both players cooperate, they both get a payoff of 3 (mutual cooperation). If one player defects while the other cooperates, the defector gets a payoff of 5 (temptation to defect), and the cooperator gets a payoff of 0 (sucker's payoff). If both players defect, they both get a payoff of 1 (punishment for mutual defection).

In the infinitely repeated version of this game, players play this stage game over and over again. The strategic consideration now includes not only the outcome of the current game but also the maintenance of cooperation or the threat of defection over an infinite horizon. The possibility of future retaliation can deter players from defecting in the present, potentially sustaining mutual cooperation as part of an equilibrium strategy, even though defection is the dominant strategy in the single-stage game.

3.2 Equilibrium Selection Problem

In game theory, an "equilibrium" is a state where no player has the incentive to deviate from their strategy, believing it to be optimal given the strategies of the others.

The Folk Theorem is a significant concept in game theory, especially in the context of infinite repeated games. Rather than a single theorem, it refers to a collection of theorems that demonstrate under certain conditions in infinite repeated games, almost any individually rational payoff configuration can be sustained as an equilibrium through appropriate strategy combinations.

The essence of the theorem is that as long as players are sufficiently patient (i.e., they place a high value on future payoffs) and capable of observing each other's past behaviors, they can support almost any payoff configuration above the "punishment baseline" (i.e., the minimum guarantee players can ensure by unilaterally deviating from cooperation) by threatening to punish those who deviate from cooperative behavior. Theoretically, this means that equilibrium selection becomes highly flexible and indeterminate.

With numerous possible equilibrium states in infinite repeated games, each corresponding to different strategy combinations and payoff configurations, a key question arises: if multiple equilibria are possible, how can we predict which one players will ultimately choose?

4 Learning Rules

The learning rule proposed in the current study takes inspiration from two notable precedents in the realm of game-theoretic learning. The initial reference point is the model by Foster and Young (2003)[1], which integrates the concept of belief formation based on the observed actions of opponents. In this framework, players exhibit near-rationality within the context of their beliefs, and it is these beliefs that they are prepared to abandon should they find them at odds with the empirical evidence of their opponent's behavior. A pivotal methodological decision made in the author's study is the exclusion of mixed strategies from consideration. This decision streamlines the analytical process by enabling a more direct approach where players can explicitly account for mistakes rather than having to deduce them through probabilistic means.

Furthermore, the learning dynamics in question also bear resemblance to the methodologies articulated by Young (2009) and Pradelski&Young (2012), particularly concerning actions within stage games. Common to these approaches is the delineation of two distinct operational modes for players: a stable mode characterized by adherence to a predetermined strategy with allowances for occasional departures, and a mode of exploratory behavior. This latter mode, denoted as a 'search mode', is typically instigated by suboptimal payoffs and embodies a trial-and-error spirit where players venture through various strategic alternatives. The trigger for this exploratory phase in the referenced models is generally associated with the receipt of low payoffs. However, in the model at hand, it is the discrepancies between a player's anticipations and the opponent's actual play that prompt the shift. This dual-mode concept echoes broader principles found in other domains, such as the adaptive mechanisms in computer science algorithms and the behavioral patterns observed in natural ecosystems.

4.1 Behavioral Modes

4.1.1 Payoff Structure in Repeated Games

1. Consider a two-player repeated game $\mathcal{G}$ with the following payoff structure. Let the action sets for the two players be A_1 and A_2, with typical elements $a_1 \in A_1$ and $a_2 \in A_2$.

2. The payoff functions for the players are given by u_1 and u_2:

$$U_i(v_i) = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} u_i^t, \tag{1}$$

   where δ is the discount factor and u_i^t is the payoff to player i at time t.

3. The learning rule R^ϵ, with ϵ denoting the error tolerance, dictates how players adapt their strategies in response to observed outcomes.

4. Define the maximum and minimum payoffs in $\mathcal{G}$ as:

$$\bar{v} := \max_{i\in\{1,2\},a\in A} u_i(a),$$
$$v := \min_{i\in\{1,2\},a\in A} u_i(a).$$

5. The convex hull of the set of feasible payoff pairs in $\mathcal{G}$ is denoted by V, and the minmax payoffs for each player, which exclude mixed actions, are given by:

$$w_i := \min_{a_j\in A_j} \max_{a_i\in A_i} u_i(a_i, a_j).$$

6. For normalization purposes, we adjust the payoffs such that $w_i = 0$ for all i. The set of feasible and strictly individually rational payoff pairs is:

$$V^* := \{v \in V : v \gg (0,0)\}.$$

7. Assuming the set V^* is non-empty, we define the maximum feasible payoff for each player as:

$$v_i^* := \sup_{v\in V^*} v_i.$$

8. We then introduce the Kalai–Smorodinsky bargaining solution concept to the payoff profile, which is given by:

$$x_i := \arg\max_{v\in V^*} \min_{i=1,2} \frac{v_i}{v_i^*}.$$

9. The Kalai–Smorodinsky payoff profile is derived from the following ratio, which represents the equalization of gains across players, ensuring fairness in the distribution of payoffs.:

$$\rho = \frac{x_i}{\bar{v}_i} \tag{2}$$

4.1.2 Certain/Uncertain Mode

In the context of the game we are considering, which is an infinitely repeated game, the situation is analogous to an endless chess match where each move depends on the history of the game. The challenge is that each player desires to win the game, or at the very least, not to lose. However, due to the game's infinite nature, there are multiple ways to "not lose." Hence, we require a set of guidelines—referred to as a learning rule—to aid players in deciding their subsequent moves.

This learning rule instructs players on how to make choices based on their memory of the game's history. It is similar to predicting your opponent's next move in a chess game by recalling their previous ones. The purpose of this rule is to assist players in finding a stable pattern—termed an "equilibrium"—that enables them to anticipate the opponent's actions and respond accordingly.

The game exhibits two fundamental mindsets:

- **Certain:** When a player is in this state, they have strong confidence in predicting their opponent's next move. This assurance allows them to react optimally according to the learning rule. If they follow the rule, they are unlikely to err, akin to being well-prepared for each move in chess.
- **Uncertain:** In this state, the player lacks confidence in forecasting their opponent's next move. They attempt to guess, but with considerable uncertainty, which increases the likelihood of random actions, similar to making an arbitrary move in chess due to an unknown opponent's strategy.

Players transition between these mindsets based on their understanding of the game. If a player perceives that their understanding aligns with the actual events, they remain certain. Otherwise, upon suspecting a discrepancy, they shift to an uncertain state, attempting to formulate a new understanding, much like pausing to reassess the opponent's strategy when it diverges from expectations in a game of chess.

4.2 Definition of Player's Behavior

In the study of repeated games, it is crucial to understand how players learn from their history of interactions to make strategic decisions. To investigate the learning rules and behavioral dynamics, we need to define a few key concepts that will establish the groundwork for our analysis.

- **Definition 1 (ℓ-history):** In the context of infinitely repeated games, players require a reference to a segment of history to inform their decisions. The ℓ-history represents the sequence of action profiles played in the most recent ℓ periods. This historical segment is fundamental to the decision-making process, as it allows players to anticipate opponents' next moves and respond accordingly.
- **Definition 2 (Memory k strategy):** Considering how players utilize historical information, we introduce the notion of a "memory k strategy." This implies a player's strategy is contingent on the most recent k rounds of action history. A strategy with memory k prescribes the same action for any two subgames that share the same k-history. This concept simplifies the strategic planning process by limiting the amount of historical data the player must consider.
- **Definition 3 (Consistency with strategies):** Beyond relying on history, players also evaluate the effectiveness of their strategies. Players begin to question their strategies when they notice a discrepancy with the observed m-history. If the number of deviations from the expected strategy exceeds a certain threshold τ, it signals a potential need for strategic adjustment. Conversely, consistency between history and strategy reinforces confidence in the chosen approach.

Through these definitions, we establish a clear analytical framework for player behavior. Players form memory k strategies based on their ℓ-history and adjust them over time through evaluations of consistency.

4.3 Strategy Transitions in the Learning Process

In repeated games, players undergo transitions between different modes of thinking based on their beliefs and the outcomes of recent interactions.

- **Certain Mode:** The certain mode, denoted as $c(s_i, s_j)$, represents a state of play where player i has a strong belief about player j's strategy s_j and vice versa. In this mode, players act on their convictions and respond with strategies they deem best against the anticipated actions of their opponents. Formally, we can express the condition for a player to remain in the certain mode as follows:

  As long as the most recent m-history is consistent with s_i and s_j, player i stays certain. Otherwise, player i transitions to the uncertain mode $u(s_i)$ with probability b in each period.

- **Uncertain Mode:** The uncertain mode, expressed as $u(s_i)$, occurs when a player doubts their current understanding of the game. This doubt leads to a state where the player is more likely to experiment with different strategies. The conditions for transition into and out of the uncertain mode can be summarized as:

  Each period, a transition occurs with probability b' from $(0, 1)$. If a transition occurs, the player picks a new belief s_j from the strategy set S_j based on the distribution $q(s_i, h)$, where h is the most recent m-history. The player then selects one of the best responses to s_j within the strategy set S_i at random with uniform probability and adopts it as the new strategy s_i. If the chosen strategy leads to consistent outcomes with the m-history, player i transitions back to the certain mode $c(s_i, s_j)$. Otherwise, they remain in the uncertain mode.

4.4 Stability and Effectiveness

In repeated games, a fundamental question is how to choose among many possible equilibria. Ideally, players learn and reach an equilibrium where their actions are coordinated, and each responds optimally to the actions of others. Therefore, we require a method to predict which equilibrium players are most likely to select eventually.

- **Stability:** Stability refers to the consistency of strategies and expected payoffs over the long run, meaning players do not change their strategies arbitrarily. A robust learning rule should assist players in identifying and adhering to successful strategies through a process of trial and error.

  The author views the stability of learning rules as a dynamic process, involving probabilities of transitioning from one strategic state, such as a certain mode, to another, like an uncertain mode. The duration a player remains on a strategy and the frequency with which they switch to new strategies depend on their confidence in the current strategy's effectiveness.

- **Effectiveness:** Regarding effectiveness, the learning rules presented in the original text focus on how players make strategic choices based on their historical experiences and the current state of equilibrium. Circumstances may arise where, despite reaching an equilibrium, players might deviate due to various reasons, such as random fluctuations or strategic errors. The learning rule needs to address these deviations and guide players back to an equilibrium state, ideally to the optimal one, such as the Kalai–Smorodinsky equilibrium.

The learning rule considers not only the immediate effectiveness of strategies but also their long-term stability and the trust players place in their strategies. The analysis suggests that when players employ this learning rule, they tend to favor a stable equilibrium state that offers higher expected payoffs. Such an equilibrium state should reflect the mutual best responses of players and withstand random perturbations over time.

5 Results

5.1 Main Result: Theorem 1

As we explore the world of repeated games, a key discovery that emerges is how we can predict how players will ultimately behave in repeated games. Imagine that you are participating in an endless strategy game, and each round you need to decide whether to cooperate or compete. As the game progresses, you begin to learn which strategies best bring about the results you want. Theorem 1 reveals the logic behind this process.

More formally, Theorem 1 addresses the conditions under which a learning rule in repeated games leads to a stochastically stable state. These states represent the strategies that players are most likely to adopt in the long run. The theorem asserts:

Given the learning parameters δ (the discount factor) and k (the memory length) are sufficiently large, the stochastically stable states of the learning process P^{ϵ} must satisfy the following conditions:

1. Players' beliefs about the opponents' strategies are accurate.
2. The strategies form a subgame perfect equilibrium. Meaning that given the strategies of other players, no one can achieve better results by changing their strategies.
3. The payoff vectors for the players are close to the Kalai-Smorodinsky bargaining solution.

Theorem 1's significance lies in that it demonstrates that under certain conditions, the process of learning and adaptation in repeated games converges to a state that is not only stable but also optimal in terms of fairness and efficiency, as indicated by its proximity to the Kalai-Smorodinsky bargaining solution.

5.2 The Role and Proof of the Lemmas

In exploring the dynamics of repeated games and how players converge to a stable state under specific learning rules, several key lemmas provide the foundation for understanding the proof of our main theorem. These lemmas dissect the process into manageable pieces, each addressing a unique aspect of the learning rule's impact on the game's equilibrium.

5.2.1 Lemma 1

Lemma 1 states that without uncertainty, the game cannot reach a stochastically stable state. Formally, if all players have complete certainty about the game's state, then those states are not part of the support of the stochastically stable distribution.

$$\text{If } \forall i,\ \Pr(s_i|h) = 1, \text{ then } s \notin \text{Support}(\mu^*). \tag{3}$$

This lemma highlights the importance of exploration and uncertainty in reaching an equilibrium that is robust and sustainable in the long run.

5.2.2 Lemma 2

Lemma 2 elaborates on the conditions necessary for a strategy profile to be part of a stochastically stable state. It specifically focuses on the role of player beliefs aligning with the strategies they observe.

$$\forall s \in S, \text{ if } b(s|h) > 0, \text{ then } s \text{ forms a Nash equilibrium.} \tag{4}$$

It underscores that for a state to be stochastically stable, the strategies employed must not only be a response to the current state but also be the best response given the beliefs about the opponents' strategies.

5.2.3 Lemma 3

The focus of *Lemma 3* is on the transition probabilities between states, asserting that transitions from any state to a stochastically stable state require a non-zero probability.

$$\forall s, s' \in S, \text{ if } s' \text{ is stochastically stable, then } \Pr(s \to s') > 0. \tag{5}$$

This lemma is crucial for proving that stochastically stable states are reachable from any starting point in the game, ensuring the robustness of these equilibrium states.

5.2.4 Lemma 4

Lemma 4 addresses the persistence of stochastically stable states, showing that once reached, the probability of remaining in such a state is high.

$$\text{If } s \text{ is stochastically stable, then } \Pr(s \to s) \approx 1. \tag{6}$$

This demonstrates the stability of these states, indicating that fluctuations away from these equilibria are infrequent and unlikely.

5.2.5 Lemma 5

Finally, *Lemma 5* deals with the efficiency of stochastically stable states, suggesting that these states are not only stable but also efficient.

$$\text{If } s \text{ is stochastically stable, then } u_i(s) \geq u_i(s') \, \forall s' \in S, \, \forall i. \tag{7}$$

This lemma ensures that the stochastically stable states are those where the utility (or payoff) for every player is at least as good as any other state they could transition to, highlighting the optimality of these states.

These lemmas together form the backbone of our understanding of how learning rules influence the emergence of stable and efficient equilibria in repeated games. At the same time, Lemma 1-5 established a theoretical framework to support the conclusion of Theorem 1. Through these mathematical explorations, we see a clearer picture of the conditions necessary for the stability and optimality of strategies in the long-term dynamics of the game.

5.3 Support of Auxiliary Theorems: Theorem A and B

Theorem A and B provide the background support and additional guarantees required by Theorem 1: Theorem A relates to the conditions for the existence of a subgame perfect equilibrium with a specific memory, which is a key part of steady state analysis. Theorem B provides a criterion for determining which states are stochastically stable, that is, those states that will occur with high frequency in the game in the long run.

5.3.1 Theorem A (Barlo, Carmona, and Sabourian 2016)

For all $v \in V^*$ and $\lambda > 0$, there exists a k-memory pure subgame-perfect equilibrium of $\mathcal{H}$ such that each player i receives continuation payoffs within λ of v_i at every subgame.

5.3.2 Theorem B (Ellison 2000)

Let ω, ω' be states within the process. If the resistance to leaving state ω exceeds the resistance to transitioning from ω to ω', denoted by $r(\omega) > r'(\omega', \omega)$, then state ω is not stochastically stable.

Theorems A and B not only assert the possibility of reaching a stable equilibrium but also ensure that the equilibrium is resilient to fluctuations inherent in the learning dynamics of the game. **These theorems are tightly coupled with the idealized Kalai-Smorodinsky bargaining solution.**

5.4 Implications and Prospects

The significance of these results lies in their profound implications for both theoretical and applied fields of economics and game theory. By demonstrating that players can achieve near-optimal outcomes through a process of learning and adaptation, we unlock a deeper understanding of the mechanisms that drive cooperation, competition, and coordination in various social and economic settings. This understanding challenges the traditional notion that equilibrium selection in repeated games is arbitrary and unpredictable, presenting instead a framework where stability and optimality are attainable even under constraints of bounded rationality.

5.4.1 Implications for Economic Theory and Practice

From a practical standpoint, these findings illuminate pathways for designing mechanisms and policies that foster desirable outcomes in markets, negotiations, and social dilemmas. For policymakers and strategists, the learning rule P^ϵ provides a tool for predicting and influencing the dynamics of strategic interactions, offering a blueprint for enhancing cooperation and social welfare.

5.4.2 Future Research Directions

The author's research discussed here not only deepens our understanding of repeated game dynamics but also opens up exciting paths for further study. Here are some key areas that could be explored next:

- Exploring the impact of different forms of learning rules on equilibrium selection and stability.
- Investigating the role of information asymmetry and how it influences the convergence to efficient equilibria.
- Extending the analysis to more complex game structures, including those with more than two players and continuous action spaces.
- Applying the insights from this research to real-world scenarios, such as market competition, environmental agreements, and public goods provision, to test the applicability and robustness of the theoretical findings.

6 Conclusion

Through the insightful framework provided by Sam Jindani's article, "Equilibrium Selection in Repeated Games," we have navigated the complexities of strategic interactions in infinite horizons, unveiling the critical role of learning rules in the stabilization and optimization of player outcomes.

At the heart of our exploration was the introduction of a novel learning rule, P^ϵ, which not only facilitates the convergence to equilibrium states corresponding to the Kalai–Smorodinsky bargaining solution but also illuminates the path towards achieving efficient and equitable outcomes. Theorem 1, with the foundational support from Theorem A (Barlo, Carmona, and Sabourian 2016) and Theorem B (Ellison 2000), serves as the linchpin in establishing the conditions under which these desirable states are attained. These conditions, notably, revolve around players' abilities to adjust their strategies based on accumulated experience and the resulting payoff outcomes. The Lemmas 1-5 provided the essential building blocks for the proof of Theorem 1, each dissecting different facets of the learning dynamics to elucidate how players transition from uncertainty to a state of equilibrium. From the initial uncertainty and exploration to the final convergence towards a stable and fair equilibrium, the learning rule P^ϵ is shown to guide players through a dynamic and adaptive process, rooted in realistic assumptions about human behavior and decision-making.

By exploring equilibrium choices in repeated games through the lens of learning rules, we can develop a nuanced and comprehensive understanding of the dynamics of strategic interactions. It provides a valuable framework for predicting and influencing outcomes in interactive decision-making scenarios. This research by Sam Jindani will also become the cornerstone of future research in the vast field of game theory.

References

[1] Dean P. Foster and H.Peyton Young. "Learning, hypothesis testing, and Nash equilibrium". In: *Games and Economic Behavior, Volume 45, Issue 1, October 2003, Pages 73-96*. 2003.

[2] Sam Jindani. "Equilibrium selection in repeated games". In: Department of Economics, University of Oxford. 2020.

[3] Ehud Kalai and Ehud Lehrer. "Rational Learning Leads to Nash Equilibrium". In: *Econometrica, Vol. 61, No. 5 (Sep., 1993), pp. 1019-1045*. 1993.

[4] Ehud Kalai and Meir Smorodinsky. "Other solutions to Nash's bargaining problem". In: *Econometrica, Vol. 43, No. 3 (May, 1975), pp. 513-518*. 1975.