Cooperation in the Evolutionary Iterated Prisoner's Dilemma Game with Risk Attitude Adaptation

Weijun Zeng^{1,2}, Minqiang Li¹, Fuzan Chen¹

¹College of Management and Economics, Tianjin University, Tianjin 300072, P.R. China

²School of Economics and Management, Hainan University, Haikou 570228, P.R. China

(wjzeng@tju.edu.cn, mqli@tju.edu.cn, fzchen@tju.edu.cn)

Abstract

The Iterated Prisoner's Dilemma (IPD) game has been commonly used to investigate the cooperation among competitors. However, most previous studies on the IPD focused solely on maximizing players' average payoffs without considering their risk preferences. By introducing the concept of income stream risk into the IPD game, this paper presents a novel evolutionary IPD model with agents seeking to balance between average payoffs and risks with respect to their own risk attitudes. We build a new IPD model of multiple agents, in which agents interact with one another in the pair-wise IPD game while adapting their risk attitudes according to their received payoffs. Agents become more risk averse after their payoffs exceed their aspirations, or become more risk seeking after their payoffs fall short of their aspirations. The aspiration levels of agents are determined based on their historical self-payoff information or the payoff information of the agent population. Simulations are conducted to investigate the emergence of cooperation under these two comparison methods. Results indicate that agents can sustain a highly cooperative equilibrium when they consider only their own historical payoffs as aspirations (called historical comparison) in adjusting their risk attitudes. This holds true even for the IPD with a short game encounter, for which cooperation was previously demonstrated difficult. However, when agents evaluate their payoffs in comparison with the population average payoff (called social comparison), those agents with payoffs below the population average tend to be dissatisfied with the game outcomes. This dissatisfaction will induce more risk-seeking behavior of agents in the IPD game, which will constitute a strong deterrent to the emergence of mutual cooperation in the population.

Key words: Evolutionary iterated prisoner's dilemma game, Risk attitude, Historical comparison, Social comparison, Agent-based simulation.

1. Introduction

Competition and cooperation are among the mostly noted paradoxical phenomena in the real world. For example, collaboration with competitors in business is common for modern firms [1], which is enforced by the rapid advance in information technology and the globalization of markets. Moreover, strategic alliances increase in various industries every year. Firms are hoping to improve their performances by collaboratively developing new technologies, obtaining critical resources, or attaining economies of scale [2]. Nonetheless, despite the benefit of mutual cooperation, firms still compete with one another for gaining sustained competitive advantages in the marketplace [3]. Then, a question arises as to how cooperation emerges among competitors when they still have strong incentives to compete. The iterated prisoner's dilemma game (IPD) is a prominent paradigm to examine the essence of this problem.

The Prisoner's Dilemma (PD) is a bilateral game in which each player can either cooperate (C) or defect (D) [4, 5]. A reward (R) will be given to both players if they cooperate with each other, whereas a punishment (P) will be enforced if they both defect. In the circumstance with one defecting but the other cooperating, the defector will obtain a tempting payoff (T) and the cooperator will receive a sucker's punishment (S). Given the conditions $T > R > P > S$ and $2R > T + S$, a dilemma emerges. Although defection is the dominant strategy for any player, mutual cooperation yields the highest total payoff. The situation will become more complex if the PD is iterated over many rounds (IPD). Thus, the IPD is often referred to as a mixed-motivation game in which opportunities for cooperation and defection (or competition) coexist.

The evolutionary IPD game has been widely employed by researchers to study the evolution of cooperative behavior among selfish individuals [4, 6-8]. The researchers usually start with a system of multiple agents (players) engaging in recurring IPD interactions while adjusting their own strategies by adaptive learning. As the adaptation proceeds, the aggregation of recurring interactions among agents can generate complex collective behavior of the system [9, 10], and the resulting cooperation (or defection) at equilibrium can be analyzed under different scenarios. The aim of this research is to verify what specific factors can influence the emergence of cooperation, and the observed results are used to explain real-world cooperative phenomena.

Evolutionary algorithms have been widely adopted as an appropriate analog for modeling the agents' adaptive learning in the evolutionary IPD game [11]. Agents are generally assumed to adapt their strategies by learning from the better-performing peers in the population. Such a learning mechanism echoes the social learning processes in human social interactions [12-14], in which individuals often draw upon others' experiences to refine own strategies and enhance own behavioral practices [15]. Recently, with the rapid development in behavioral disciplines [16], in which personal traits are commonly treated as independent variables to explain human behavior, it is strongly

suggested to design agents with explicit personal traits in light of behavioral studies through the agent-based modeling [17, 18].

Risk is a critical feature of individual activities, and risk attitude has been treated as an independent variable in explaining cooperation in the behavioral IPD experiments. Based on the principle that conditional cooperation supports a sub-game perfect equilibrium [19], Snijders and his colleagues [20, 21] argued that a player's strategy in the IPD game was whether or not to unilaterally exploit the opponent who cooperated conditionally. Their studies showed by behavioral experiments that risk-averse players favored cooperation in such a scenario, whereas risk-seeking players tended to unilaterally defect. However, their formulation of risk involved only individual preferences on gambles between deterministic and variable prospects. No study had explicitly defined what risk was in the IPD game.

Risk-averse players retain steady cooperation with their opponents because the cooperation could minimize the interaction uncertainty and yield stable payoffs for them in the IPD [20, 21]. This observation is in accordance with the proposition that firms in the iterated Cournot competition, a realistic IPD situation, would like to coordinate with each other on their production outputs to enhance their income stream stability [22]. Motivated by these studies, we first defined risk as the standard deviation of the one-move payoffs for players in the IPD game [23]. This definition on risk is in line with the concept of "income stream risk" in economics and finance [24], in which risk is conceived as a manifestation of uncertainty or uncontrollability of outcomes [25].

Based on this risk definition, we have re-examined the two-player IPD problem by considering both maximizing average payoff and minimizing risk for players [23]. Computational simulation did demonstrate that risk-averse players favored mutual cooperation in the IPD game [23]. Nonetheless, we focused only on the learning of the IPD for two agents of fixed risk attitudes, but did not consider the possibility of changes in their risk preferences over time. Therefore, we attempt to consider agents of adaptive risk attitudes in this paper by following the fact that risk preference is contextually dependent in human society [26, 27].

The evolutionary IPD game with pair-wise interactions of multiple agents is investigated, in which agents play the IPD game by adapting their risk attitudes according to relative performance feedback. Following the prospect theory [26] and empirical findings on human risk behavior [28, 29], agents respond to poor performances by taking greater risk or satisfactory performances by avoiding risk (see Section 3.2 for more details). In particular, the poor or satisfactory performances are distinguished by agents in comparison with their own past performances (historical comparison) or with the population's average performance (social comparison). Consequently, the evolution is characterized by the changes of risk preferences in the population, and the resulting outcome should be reexamined. The contributions of this study involve the following three respects:

(1) A novel evolutionary IPD model is formulated with agents of adaptive risk attitudes, which is aligned with the trend of integrating human traits into the agent-based modeling. To our best knowledge, this work is the first attempt to study the adaptive risk behavior in the evolutionary IPD game, in which the effects of historical comparison and social comparison on the evolution of cooperation are investigated particularly.

(2) Based on the evolutionary model, extensive simulation studies are conducted to examine the effect of dynamic risk consideration on the evolution of cooperation. Results indicate that high levels of cooperation can emerge in the IPD even with short game encounters when agents consider only their past payoffs as reference points (historical comparison) to assess their performances for adjusting risk attitudes. The whole population can achieve sustainable high returns at low risk in this case. The speed by which agents change their risk attitudes is shown to have a significant impact on the cooperation.

(3) When agents employ the population's average payoffs as reference points to appraise their performances (social comparison), we find that agents will become more dissatisfied on average. As a result, the evolution leads to a large proportion of highly risk-seeking agents, which deters mutual cooperation and brings low returns to the population. Thus, our findings present the evidence that uncooperative behavior can be perceived as a consequence of social comparison.

The rest of this paper is organized as follows. Section 2 reviews some related literature on the IPD game. Section 3 presents the definition of risk and introduces the theoretical basis for the risk attitude adaptation. Section 4 formulates the proposed evolutionary IPD model with agents of adaptive risk attitudes. The simulation studies are reported in Section 5. Section 6 concludes this study with remarks for future work.

2. Related Literature

Games are often employed to model socio-economic phenomena [30], and the evolutionary PD and IPD games have been extensively studied to understand the evolution of cooperation, which has been a considerable concern to evolutionary biologists [31], economists [32, 33], social scientists [34], cognitive scientists [35], and even computational scientists [36].

A classical evolutionary PD game involves a population of individuals who play the PD game against each other (called the well-mixed game setting). Because uncooperative individuals (defectors) always have higher average payoff (fitness) than cooperators, defection dominates cooperation in the population, driving cooperators to extinction under natural selection. This evolutionary outcome can be analytically derived in replicator dynamics, imitation dynamics, or selection-mutation dynamics [37]. Nonetheless, because cooperation is the beneficial choice to the whole population, a question arises as to what mechanisms can promote the cooperation in the

evolutionary PD game.

Ohtsuki et al. [31] considered an evolutionary PD game with individuals occupying vertices of a graph, in which the edges of the graph determined links between individuals in terms of game interaction. An individual only played the PD game with the connected ones in the population. Ohtsuki et al. [31] found that the graph-based interaction could allow cooperators to form clusters. They calculated the fixation probability that a single cooperator invaded a population of defectors under natural selection, and concluded that the fewer the connections among individuals, the easier the natural selection promoted cooperation. This mechanism was defined as network reciprocity [38]. Zhang et al. [39] extended the work of Ohtsuki et al. [31] by externally forcing some randomly chosen individuals to cooperate with a certain probability. Mathematical analyses proved that the cooperation could be enhanced with such an external constraint [39].

Izquierdo et al. [40] further provided individuals with an option to disconnect current links in response to the game outcomes, i.e., conditional dissociation. When a link was broken, both individuals had to find new links with other peers. Izquierdo et al. adopted computer simulations and analytical approximation to analyze the population dynamics with conditional dissociation under the standard evolutionary forces (selection and mutation). It was concluded that conditional dissociation could induce a significant level of cooperation in the evolutionary PD game. In other words, conditional dissociation could provide an evolutionary basis for the emergence of cooperation through generating endogenous positive assortment between cooperators.

Inspired by the notion of indirect reciprocity from the field of evolutionary biology [41], Stahl [42] developed an evolutionary PD model with individuals using Markov strategies, which mapped the reputation information of an individual and its opponent to the possible actions (i.e., C or D). Stahl [42] examined a class of reputation mechanisms, and summarized that, under the reputation mechanism entailing both punishment and penitence¹, a strategy that cooperated with good opponents (with good reputation) and defected against the bad ones (with bad reputation) was a global attractor of replicator dynamics. Therefore, cooperation was steadily sustained, highlighting the important role of indirect reciprocity in the evolution of cooperative behavior [41]. The positive effect of the reputation mechanism on the cooperation has also been supported by behavioral experiments in [43].

Real-world interactions always involve repeated encounters, such as the foraging competition between two species and the Cournot competition between two firms [44]. The IPD game indicates a more complex situation than the PD game. Fudenberg and Maskin [19] proved that cooperation could be self-enforcing by players if there were enough rounds in the IPD game². This kind of

¹ Cooperating with a good opponent when an individual was in the bad reputation state was penitence, which resulted in the restoration of the individual's reputation to good. Defecting against a bad opponent when an individual was in the good reputation state was punishment, which did not result in the loss of good reputation [42].

² The discount factor δ applied by Fudenberg and Maskin [19] could be considered as the probability of playing another PD round in the whole IPD game. Then, the average number of rounds could be given by $1/(1-\delta)$. Hence,

cooperation was termed as direct reciprocity [38]. In particular, their results depended on trigger strategies, which switched to relentless defection as long as opponents defected. It was this threat that forced cooperation in equilibrium in the IPD game, which was considered as conditional cooperation [20].

Experimental studies on human behavior have confirmed that people are actually conditional cooperators [45, 46]. Suzuki et al. [35] showed the neural and psychological bases of conditional cooperation and suggested that the cognitive inhibition of motivation to cooperate with defectors led to the conditional behavior. Axelrod [4] sponsored three IPD tournaments by inviting scholars from the fields of economics, mathematics, psychology, and social sciences to submit their strategies, and found that a conditional cooperative strategy, Tit for Tat, which cooperated in the first round, and then copied opponent's last choice, won in all of the tournaments. Inspired by his work, researchers have developed various efficient IPD strategies with the cognitive capacity to perform conditional cooperation, using artificial neural networks [44] or incremental learning technologies [36].

Theoretical studies on the evolutionary IPD game also supported conditional cooperative strategies [39, 47]. For example, Moreira et al. [48] devised an evolutionary IPD game model in which individuals can be unconditional defectors or memory-constrained conditional cooperators. The conditional cooperators exhibited different memory capacities, which helped them remember opponents' previous actions. It was theoretically verified that increasing the memory capacities could enable cooperation to emerge and sustain in populations of increasing sizes. While studies on reputation mechanism [42] were implicitly associated with individuals' cognitive capacities, the work of Moreira et al. [48] focused on cognition explicitly. With the aim to obtain analytical solutions, individuals were often restricted to a small set of pre-determined strategies in the theoretical IPD models [39, 47, 48]. As a consequence, the theoretical models often lack the capability to demonstrate innovative behavior of individuals in the IPD game [37].

The limitation of the theoretical models can be supplemented by multi-agent models [37], in which agents (individuals) are allowed to adopt novel strategies to play the IPD game through adaptive learning. Studies of multi-agent models are usually performed by simulations, which involves a variety of agents [17]. Multi-agent IPD models are more flexible to address the cooperation problem. For example, Fogel [6] studied the evolutionary IPD game with agents represented by finite-state machines under the well-mixed game setting. He found that cooperation (i.e., conditional cooperation) hardly emerged in the case with very short game encounters. Note that this result was in line with the argument in [19], which stated that cooperation could only be sustained in the IPD game with enough rounds.

After the seminal work of Fogel [6], researchers have tended to examine the mechanisms promoting the evolution of cooperation in the short-encounter IPD games with the multi-agent

“there were enough rounds” means “a sufficiently large discount factor δ ” in [19].

models. Alkemade et al. [49] introduced the tag-mediated mechanism into the IPD game, following the theory of kin reciprocity from evolutionary biology [50]. Based on this mechanism, each agent carried an additional tag, and only those with similar ones were matched to play the IPD game. As similar to conditional dissociation, the tag-mediated mechanism allowed the positive assortment between cooperators, and thus ensured the benefits of mutual cooperation to outweigh the losses against defectors. In this case, cooperative behavior could be effectively promoted in the IPD with short game encounters [49].

Ishibuchi and his colleagues [7, 51] examined the network reciprocity in the evolutionary IPD game by assuming agents to be distributed across square lattices. They concluded that fewer connections among agents (i.e., small interaction neighborhoods) could facilitate the evolution of cooperation. Furthermore, they also considered the spatial mechanism in the random pairing IPD game, in which each agent played against a randomly chosen neighbor in each round and thus the interaction length against the same opponent was very short. It was verified that cooperation could be even evolved in such a difficult case when the neighborhood structure was specified appropriately [51].

Chong and Yao [52] developed agents with an additional ability to learn about reputations of opponents in the evolutionary IPD game. Agents selected strategies according to the opponents' reputations, and high levels of cooperation could be steadily evolved in this case. This outcome was verified in all circumstances whatever the length of game was specified [52], demonstrating the successfulness of indirect reciprocity [41, 42] in promoting the evolution of cooperative behavior in the IPD game.

Agents were supposed to maximize their average payoffs in the IPD game in the above studies. In our previous research on the two-agent IPD game [23], we have demonstrated that risk aversion of agents could also enhance mutual cooperation in the IPD with short game encounter, and this result was also in accordance with the outcomes of behavioral experiments in [20, 21], which stated that risk-averse individuals were more cooperative than the risk-seeking ones in the IPD game. In particular, the risk in [23] was defined as the standard deviation of the one-move payoffs for agents in the IPD game by following the concept of "income stream risk" in the field of economics and finance [24].

However, we only considered games with agents of fixed risk attitudes in [23]. Since an individual's risk preference strongly depends on the context in which it is situated [53, 54], we attempt to extend our investigation into the IPD game with agents of adaptive risk attitudes. A multi-agent IPD model is considered, in which agents engage in pair-wise IPD interactions while adapting their risk attitudes based on prospect theory [26]. Our main purpose is to examine the novel effect of the dynamic risk consideration on the evolution of cooperation in the multi-agent IPD model, especially with a short-encounter game setting. Using prospect theory to study the behavior of agents was in aligned with the trend of constructing cognitively sophisticated agents with designs

informed by psychological theories [17].

3. Theoretical Preliminaries

3.1 Risk Definition in the IPD

Suppose that the one-move payoffs for a player in the IPD is $p = (p_1, p_2, \dots, p_l)$, where l is the length of the encounter and p_k represents the payoff on the k -th move ($k = 1, 2, \dots, l$). Then, the income stream risk for the player can be measured as:

$$r = \sqrt{\frac{1}{l} \cdot \sum_{k=1}^l (p_k - \bar{p})^2} \quad (1)$$

and the player's utility can be expressed as follows:

$$u = \frac{\bar{p}}{1 + \alpha \cdot r} \quad (2)$$

where $\bar{p} = \sum_{k=1}^l p_k / l$, and α represents the player's risk attitude. The value of α is set as 0 if the player is completely risk-seeking. Otherwise, $\alpha > 0$. The more risk-averse a player is, the larger α becomes [23].

The utility function (2) involves the two most important factors for individual performance measures, namely, the average payoff and the risk [27], which thus creates a decision-making methodology for players to balance between average payoff and risk based on their own risk preferences in the IPD game. Risk-seeking players (with small values of α) tend to ignore the risk of interaction. They will try to maximize their potential payoff in each move. Such an attempt may entice them to exploit their cooperative opponents, even though the exploitation will expose them to being retaliated, which often yields lower average payoffs for both players. On the contrary, risk-averse players (with large values of α) are willing to maintain a steady cooperation with their opponents for reducing risk in the IPD game.

3.2 Risk Attitude Adaptation in the IPD game

According to prospect theory [26], an individual's risk preference strongly depends on the decision-making context in which it is situated. Prospect theory assumes that individuals normally perceive outcomes (in the decision-making context) as gains or losses relative to some reference points, and they tend to be risk averse if the perceived outcome lies in the domain of gains, otherwise, they become risk seeking when the perceived outcome lies in the domain of losses. In the past decades, researchers have provided a large amount of evidences for the reversal change of risk preference in a variety of decision-making domains, including politics [55], financial markets [56], and organization management [29].

We attempt to incorporate the above reversal change of risk preference into the evolutionary IPD game by allowing agents to adjust risk attitudes according to their game payoffs. Following previous research [28, 53, 54, 57], the reference point used by an agent for evaluating the outcomes (i.e., payoffs) are defined as its aspiration level (see Section 4.3.1), which is further decided by its past payoff (historical comparison) and by the population average payoff (social comparison). Then, high-performance agents whose payoffs are above their target aspiration levels will increase their risk-aversion levels (α) in subsequent game iterations, whereas low-performance agents with payoffs below aspirations are more likely to decrease their risk-aversion levels. As a result, the dynamic risk attitude adaptation and its effect on the emergence of cooperation can be investigated in the evolutionary IPD game.

The above risk attitude adaptation mechanism also has special meaning in the IPD game. First, our study defines risk as the standard deviation of one-move payoffs for agents in the IPD interactions, and thus an agent's risk aversion level actually reflects how the agent evaluates the stability of its payoffs (profits) [24]. There exists a large amount of empirical evidence that individuals with performances below aspirations (in the loss situations) would like to initiate exploration of new practices [54]. Such an exploration, however, may directly give rise to a large fluctuation in the individuals' performances (payoffs). In contrast, individuals with performances above aspirations (in the gain situations) will become more concerned with their steady development, i.e., their income stream stabilities [29, 58]. Therefore, the risk attitude adaptation mechanism represents these concerns of individuals in the gain or loss situations.

Second, as discussed in Section 3.1, individuals with higher risk aversion levels are more willing to maintain a steady cooperation with their opponents in the IPD game, and this proposition has been well verified by laboratory experiments [20, 21], and also by computational experiments in our previous study [23]. Thus, the risk attitude adaptation mechanism embodies the fact that satisfied (or dissatisfied) individuals, with payoffs higher (or lower) than aspirations, will become more cooperative (or uncooperative), echoing well with previous research findings [59, 60] that individuals experiencing positive (or negative) affect will become more cooperative (or uncooperative). The reason is that individuals tend to adopt a more mutual-rewarding (or self-serving) interpretation when they are making decisions in the domain of gains (or losses) [59, 61]. A successful long-term cooperative relationship will surely rely on high levels of satisfaction of players in the IPD game.

More detailed description of the evolutionary IPD model with multiple agents of adaptive risk attitudes is presented in the following section.

4. Evolutionary Model

Multi-agent models, with agents interacting and co-adapting in a population, have been used to

address various real-world problems, such as manufacturing process organization [62] and natural resource allocation [63]. This section describes the proposed evolutionary IPD model with multiple agents of adaptive risk attitudes, which helps provide insights into the relationship of risk adaptation to cooperation. Fig. 1(a) shows the main process of the model. The evolution starts with a random initialization of a population of N agents, including the initialization of agents' risk attitudes and strategies. Afterwards, in each generation, agents play the IPD game, and perform risk attitude adaptation and strategy modification subsequently. This evolution process iterates until the maximum generation is reached.

Fig. 1(a). Evolutionary IPD Model

Fig. 1(b) depicts the pseudocode of the evolutionary process. The risk attitude adaptation is implemented on the basis of the prospect theory [26, 64], and is designed to be payoff-oriented. On the other hand, the strategy modification is performed with regard to agents' utilities, integrating agents' payoffs and risk attitudes. Evolutionary learning is applied for the strategy modification.

Evolutionary Process

```
1: Initiate a population of  $N$  agents. Each agent  $i$  ( $i = 1, 2, \dots, N$ ) is specified by a risk attitude  $\alpha_{i0}$  and a strategy  $s_{i0}$ 
2: Set:  $t = 0$ 
3: While  $t \leq G$  do
4:   // Interaction
5:   For  $i = 1$  to  $N$  do
6:     Randomly select  $g_n$  opponents  $\{Opp_{ijt}, j = 1, 2, \dots, g_n\}$  from the population for Agent  $i$ 
7:     Agent  $i$  plays the IPD game with each opponent
8:     Calculate the average payoff  $P_{it}$  for Agent  $i$  // Equation (3)
9:   End For
10:  // Payoff-oriented Adaptation of Risk Attitude according to Prospect Theory
11:  For  $i = 1$  to  $N$  do
12:    Determine the aspiration goal  $G_{it}$  of Agent  $i$  // Equation (4)
13:    If  $P_{it} \geq (1 + \gamma) \times G_{it}$  then
14:       $\alpha_{i,t+1} \leftarrow \alpha_{it} \times (1 + r_{up}) + N(0, \alpha_{it} \times r_{up} \times v)$ 
15:    Else If  $P_{it} \leq (1 - \gamma) \times G_{it}$  then
16:       $\alpha_{i,t+1} \leftarrow \alpha_{it} \times (1 - r_{down}) + N(0, \alpha_{it} \times r_{down} \times v)$ 
17:    Else  $\alpha_{i,t+1} \leftarrow \alpha_{it}$ 
18:    End If
19:  End If
20: End For
21: // Utility-oriented Modification of Strategy based on Evolutionary Learning
22: Select the best  $N/2$  agents with the highest average payoffs in Interaction, and their strategies are denoted as
    $\{s_{1,best}, s_{2,best}, \dots, s_{N/2,best}\}$ 
23: For  $i = 1$  to  $N$  do
24:   Randomly select two strategies from  $\{s_{1,best}, s_{2,best}, \dots, s_{N/2,best}\}$  with probabilities proportional to the average
   payoffs of the corresponding agents
25:   A candidate strategy  $C_{it}$  is generated from the two chosen strategies by combination and mutation
26:   Calculate the utilities  $U(C_{it})$  and  $U(s_{it})$  based on the risk attitude  $\alpha_{i,t+1}$  // Equations (6) & (7)
27:   If  $U(C_{it}) \geq U(s_{it})$ 
28:     then  $s_{i,t+1} \leftarrow C_{it}$ 
29:   Else  $s_{i,t+1} \leftarrow s_{it}$ 
30:   End If
31: End For
32: Set:  $t = t + 1$ 
33: End While
```

Fig. 1(b). Pseudocode of Evolutionary Process

The agent definition, interaction structure, and learning dynamics, which are fundamental to the behavior of a multi-agent model, are elaborated respectively as follows.

4.1 Agent Definition

With the initialization, each agent, indexed as i ($i = 1, 2, \dots, N$), is specified by a risk attitude α_{i0} and an IPD strategy s_{i0} (the “0” in the subscript denotes the initial generation). The value of α_{i0} determines the payoff-risk tradeoff of the agent, and it is limited to the range of $[0, 1]^3$. The strategy s_{i0} specifies the agent’s actions in the IPD game, which is conditioned on the historical information of prior three moves [4, 65]. The binary string representation of strategy is exploited [66]. Because each move has four possible outcomes (“D-D”, “D-C”, “C-D”, and “C-C”), a total of 4^3 possible histories exist. Thus, a string has 64 bits to store an action for each possible history, in which “1” represents “C” and “0” represents “D”. For example, bit position 15, with its binary code as “001110” (the decoded value starts at zero), specifies the action for the history with an agent playing “D-C-C” and the opponent playing “D-C-D” in the previous three moves. If its value is “0”, then the agent will defect; otherwise, the agent will cooperate.

A history of previous three moves is randomly assigned for agents at the beginning of an IPD game [66]. This can help agents find strategies more flexible in different situations. No information is assumed about the length of a game in the binary string representation, which means that agents cannot figure out when the IPD interaction will come to an end. Thus, no “shadow of the future” exists [67].

4.2 Interaction Structure

To filter out the interaction factors that might contribute to the maintenance of cooperation, such as kin-biased interaction [49] or spatial (or network) interaction [8, 68], each agent is assumed to play the IPD game against g_n opponents randomly chosen from the population. Here, we apply a scheme for selecting the g_n opponents for an agent to reduce the effect of randomness. In each generation t ($t = 0, 1, 2, \dots, G$), when agent i selects its opponents, the remaining agents are sorted into a list of an ascending order with respect to their risk attitudes α_{jt} ($j = 1, \dots, i-1, i+1, \dots, N$). Then, those agents are sectioned into g_n equal groups, and an opponent is randomly selected from each group.

The average payoff for agent i in the IPD game is computed as

$$P_{it} = \frac{1}{g_n} \sum_{j=1}^{g_n} \bar{p}(s_{it}, Opp_{ijt}) \quad (i = 1, 2, \dots, N) \quad (3)$$

where $\bar{p}(s_{it}, Opp_{ijt})$ represents the average payoff per move for agent i to play against its opponent agent Opp_{ijt} , $j = 1, 2, \dots, g_n$.

4.3 Learning Dynamics

³ The risk attitudes of agents were picked from $\{0, 0.1, \dots, 1.8\}$ in our previous work [23]. But agents with $\alpha > 0.8$ were observed to perform similarly well and were identified as highly risk-averse in the IPD game. Thus, we restrict the range of the risk attitudes of agents to be $[0, 1]$ in this paper.

4.3.1 Payoff-oriented Adaptation of Risk Attitude

After the pair-wise interactions, all agents undergo risk attitude adaptation, which follows the tenets of prospect theory [26]. As described in Section 3.2, prospect theory suggests that an agent usually compares its payoff with a reference point, and becomes risk averse (or risk seeking) when the former is higher (or lower) than the latter. Hence, the key factor to apply prospect theory here is to determine the reference points used by agents to evaluate their game payoffs.

(1) Setting Aspirations as Reference Points: Historical Comparison and Social Comparison

Aspiration levels or goals are usually used by individuals as reference points to appraise own performances [26, 28, 53]. For example, firms normally evaluate their management achievements against their pre-determined goals. Thus, agents are set to appraise their payoffs P_{it} ($i = 1, 2, \dots, N$) against pre-specified aspiration goals G_{it} ($i = 1, 2, \dots, N$) for changing their risk attitudes in the evolutionary IPD game. In particular, the agents' aspiration goals are determined by their past payoffs (historical comparison) and the payoffs of peers in the population (social comparison).

Individuals' comparison with past performance (historical comparison) or comparison to relevant referents (social comparison) has been demonstrated to strongly influence the aspiration goals of individuals. On one hand, historical comparison can help individuals to evaluate their current states relative to the past, and thus to assess any progress or regress [54, 57]. On the other hand, social peers stand as an external benchmark for individuals to make comparison [69, 70]. Thus, social comparison is also an important source of aspiration. For example, people would like to pursue a relatively higher social rank for contentment and confidence [71]. Likewise, organizations often set goals to outperform competitors in the market competition [28, 29].

Major extant studies, including empirical studies [29, 53, 72] and agent-based simulation studies [73], use the average performance of peers as the standard for social comparison. In particular, the aspiration of an individual is usually defined as a linear combination of its own past performance (historical comparison) and the average performance of the reference peers (social comparison) [54, 57]. Following these studies, we define an agent's aspiration goal as the weighted sum of its own past payoff and the population average payoff in the evolutionary IPD game:

$$G_{it} = (1 - \beta) \times P_{i,t-1} + \beta \times \bar{P}_t \quad (i = 1, 2, \dots, N) \quad (4)$$

where β ($0 \leq \beta \leq 1$) denotes the proportion of the social comparison in the overall aspiration level, and $\bar{P}_t$ represents the population average payoff in the current generation. Here, $\bar{P}_t$ is defined as the population median payoff rather than the mean payoff to avoid the effect of extreme cases [58].

(2) Mechanism of Risk Attitude Adaptation

The adjustment of risk attitude of each agent i is formulated as follows⁴:

$$\alpha_{i,t+1} = \begin{cases} \alpha_{it} \times (1 + r_{up}) + N(0, \alpha_{it} \times r_{up} \times v) & P_{it} \geq (1 + \gamma) \times G_{it} \\ \alpha_{i,t} & (1 - \gamma) \times G_{it} < P_{it} < (1 + \gamma) \times G_{it} \\ \alpha_{it} \times (1 - r_{down}) + N(0, \alpha_{it} \times r_{down} \times v) & P_{it} \leq (1 - \gamma) \times G_{it} \end{cases} \quad (5)$$

First, if P_{it} is higher than G_{it} by a ratio larger than γ ($\gamma \geq 0$), then agent i is considered to be satisfied with its performance, and it will increase its risk aversion level by a ratio of r_{up} . Second, if P_{it} is lower than G_{it} by a ratio greater than γ , then agent i is regarded to be dissatisfied, and it will decrease its risk aversion level by a ratio of r_{down} . Both of these two kinds of adjustments are perturbed by a Gaussian disturbance. If the Gaussian disturbance drives the risk attitude value outside the range of $[0, 1]$, the attitude will be adjusted back to the closer boundary value. The parameter v is assumed to be 0.2, which specifies the scale of the Gaussian disturbance. Moreover, since individuals are always loss averse, the impact of a loss on the risk attitude is always larger than that of an equivalent gain. This has been known as the rule of “losses loom larger than gains” in prospect theory [26]. Thus, the inequality $r_{up} \leq r_{down}$ is further required. Third, if the deviation of P_{it} from G_{it} is smaller than the ratio of γ , then agent i is considered to be indifferent about the game outcome, and its risk attitude will remain unchanged.

Based on Equation (5), agents’ aspiration goals can exert a significant impact on the changes of their risk attitudes in the evolutionary IPD game. The incorporation of historical comparison or social comparison into agents’ aspirations is in line with previous literature, and it is demonstrated that historical comparison [53, 54] or social comparison [29, 71] do affect an individual’s risk behavior. Moreover, a direct examination can be conducted about the effects of the two comparison approaches on the evolution of cooperation in the evolutionary IPD game.

4.3.2 Utility-oriented Modification of Strategy

Agents update their strategies based on a social learning approach [5, 50, 68] after the adjustment of risk attitude. The underlying mechanism is that agents tend to learn from those who are successful. Such a learning mechanism, as mentioned in Section 1, echoes the social learning processes in human societies [12-14], in which individuals are usually reliant on the successful peers’ experiences to refine their strategies [15].

Here, successful agents are assumed to be the best $N/2$ agents with the highest average payoffs in the current population, whose strategies are denoted by $\{s_{1,best}, s_{2,best}, \dots, s_{N/2,best}\}$. When agent i ($i = 1, 2, \dots, N$) modifies its strategy, two strategies are randomly selected from $\{s_{1,best}, s_{2,best}, \dots, s_{N/2,best}\}$ with probabilities proportional to the average payoffs of corresponding successful agents. Then, a candidate strategy (C_{it}) is generated with each of its bit value randomly picked from either of the two selected strategies (combination). Afterwards, each bit of the candidate

⁴ Risk attitude adaptation is absent in the initial generation if $\beta < 1$ because no historical payoffs exist.

strategy is subject to a bit-flip mutation with a probability of 0.05⁵. Interpreted in the economic context, the above selection scheme reveals the principle that the better one performs, the more its strategy will be known and learned by others. The combination operation can be interpreted as the integration of different efficient strategies, and mutation means the attempt of innovation.

The utility of the candidate strategy C_{it} is calculated as:

$$U(C_{it}) = \frac{1}{g_n} \sum_{j=1}^{g_n} \frac{\bar{p}(C_{it}, Opp_{ijt})}{1 + \alpha_{i,t+1} \cdot r(C_{it}, Opp_{ijt})} \quad (6)$$

where $r(C_{it}, Opp_{ijt})$ represents the risk for C_{it} to play against Opp_{ijt} , $j = 1, 2, \dots, g_n$. The strategy s_{it} of agent i will be renewed as C_{it} if the following condition is satisfied:

$$U(C_{it}) \geq U(s_{it}) = \frac{1}{g_n} \sum_{j=1}^{g_n} \frac{\bar{p}(s_{it}, Opp_{ijt})}{1 + \alpha_{i,t+1} \cdot r(s_{it}, Opp_{ijt})} \quad (7)$$

Otherwise, s_{it} remains unchanged and is directly copied as $s_{i,t+1}$ for agent i in the next generation.

Based on Equation (7), an agent will change its strategy as long as its utility can be increased. This strategy adaptation mechanism relaxes the rather strict cognitive assumption in classical game theory, which states that individuals are looking for the optimal response. The adaptation by social learning has been demonstrated to be more cost-effective and closer to real cognitive processes [75]. The strategies $\{s_{1,best}, s_{2,best}, \dots, s_{N/2,best}\}$ are unchanged in each generation, and thus the strategy updating is synchronous for agents. After the strategy modification process is accomplished for every agent, the strategy modification rate of the population is calculated as the ratio of the number of agents with successful strategy renewal to the total number of agents (N).

The routines of game interaction, attitude adaptation, and strategy modification are repeated in each generation until the maximum generation G is reached. According to Equations (6) and (7), the risk attitudes of agents exert a significant impact on their strategic choices in terms of the payoff-risk tradeoff in the IPD game. Note that agents try to update their strategic behavior in each generation. Therefore, given $\gamma \geq 0$, the evolutionary model manifests the intuition that the preferences of agents may adapt slower than their strategies. The parameter γ can be regarded as an indication of the sensitivity of agents' risk attitudes with respect to their performances in the evolutionary IPD game. A smaller value of γ means that agents change their risk attitudes in a quicker response to the game outcomes, whereas a larger value of γ means that agents change their risk attitudes in a slower

⁵ The mutation operator is applied for the exploration of innovative strategies for agents in the learning of the IPD game, and thus the mutation rate cannot be too low. Certainly, the mutation rate cannot be too high either, since high mutation will degenerate the evolutionary leaning into a random search [67]. The mutation rate of 0.05 used in this study is chosen based on the previous work in [36, 74]. Experiments demonstrate that the agents' cooperative or defective behaviors are robust in a wide range of mutation probabilities.

manner. These settings reflect the fact that individuals (or organizations) may respond to what happens swiftly or slowly.

4.4 Experimental Setting

Table 1
A Brief Summary of the Experiments

Experiment	Initial Attitudes of Agents	Attitude Adaptation	β	γ	r_{up}	r_{down} / r_{up}
<i>E0</i>	0	No	/	/	/	/
<i>E1</i>	[0, 1]	No	/	/	/	/
<i>E2</i>	[0, 1]	Yes	0	{0,0.05,...,0.5}	{0.2,0.4,...,1}	1
<i>E3</i>	[0, 1]	Yes	{0.05,0.1,...,0.3,0.5,1}	{0,0.05,...,0.5}	{0.2,0.4,...,1}	1
<i>E4</i>	[0, 1]	Yes	{0, 1}	{0,0.05,...,0.5}	0.4	{1,1.2,...,2}

We are interested in the emergence of cooperation in the evolutionary IPD model and the way by which it is affected by parameters β , γ , r_{up} , and r_{down} . This model is analytically infeasible, and therefore we rely on computational simulations. The standard payoff values $T=5$, $R=3$, $P=1$, and $S=0$ are consistent with the parameterization in previous studies [4, 51, 65]. The encounter length l is fixed as 10. Thus, mutual cooperation is difficult to achieve in the IPD without risk consideration [6]. The population size is specified as $N=200$, and every agent is set to play against $g_n = 20$ opponents in each generation t ($t = 0, 1, 2, \dots, G$). A maximum of $G = 20,000$ generations are performed.

A brief summary of parameter settings for the computational experiments is depicted in Table 1. We begin with the experiment of the evolutionary IPD game with no risk consideration (*E0*), in which agents' risk attitudes are fixed as 0 in the evolution. Then, the experiment *E1* examines the IPD game with risk consideration but without risk attitude adaptation. Risk attitudes of agents are independently and randomly drawn from the uniform interval of [0, 1] in the initial generation and then remain unchanged throughout the evolutionary process. The *E0* and *E1* serve as two baselines for subsequent experiments.

The experiment *E2* examines the dynamics of risk adaptation, in which risk attitudes of agents are independently and randomly initialized and then are adjusted in the evolutionary process. In particular, agents consider only their own historical payoffs as reference points in adjusting their risk attitudes. The value of β is fixed to 0. The rates of changes of risk attitude are fixed as $r_{up} = r_{down} \in \{0.2, 0.4, 0.6, 0.8, 1\}$. The parameter γ is altered from 0 to 0.5 in a step of 0.05, resulting in a total of 55 different scenarios. The experiment *E3* further investigates the circumstances in which agents also benchmark their payoffs against the population average payoffs in adapting their risk attitudes. All parameters are the same as in *E2* except for β , which is varied from 0.05 to 1. In the experiment *E4*, we fix the value of r_{up} as 0.4 and change the ratio of

r_{down}/r_{up} from 1 to 2 in an increment of 0.2 to examine any changes in the evolutionary outcomes, following the principle of “losses loom larger than gains” in the prospect theory [26]. The parameter β is specified as 0 or 1 for the considerations of historical or social comparisons respectively.

5. Simulation Results

The payoff value of 2.6 is taken as a surrogate for the emergence of mutual cooperation in the IPD game [6], which is higher than the average payoff of 2.5 that agents can obtain from the move of “CD” or “DC”. Each experiment with a specific parameter setting is repeated for 20 independent runs, and the average results are reported⁶.

5.1 Two baselines (E0 and E1)

Fig. 2. Population Average Payoff by generation for the evolutionary IPD game without risk consideration (E0)

Fig. 3. Population Average Payoff by generation for the evolutionary IPD game with risk consideration but without risk attitude adaptation (E1)

⁶ Experiments show that the results averaged over 20 runs is not statistically different from (with 95% confidence level) those averaged over 30, 40, or 50 runs.

Fig. 2 illustrates the population average payoff by generation for *E0*. We find an initial drop in the population average payoff from 2.25 to about 1.7, indicating that the population tends to mutually defect in early generations. The reason is that the strategies of agents are initialized randomly and the best strategy is to always defect in such an environment. However, this condition last for only a few generations. From about the 30th generation, some strategies begin to reciprocate cooperation, and become proliferated in the population. This makes the population average payoff steadily rise to about 2.5 at about the 350th generation (Fig. 2), which, on the other hand, slowly declines to and converges to a value of about 2.25.

The latter decline in the population average payoff can be attributed to the spite effect in the IPD game. A spiteful action is one that worsens the payoffs of others at no benefit or even at a cost to the actor [76, 77]. The purpose of the action is to reduce the possibility of being outcompeted. Thus, although the population can achieve a certain level of cooperation after the initial drop in *E0*, the spite emerges afterwards due to the frequent adoption of defection by some agents, which can yield them higher, at least not lower, payoffs than their opponents in the IPD game. However, the defection bring the agents into lose-lose situations⁷.

Fig. 4. Comparison of *E0* and *E1* in the percentage of moves of “D-D”, “D-C” and “C-D”, and “C-C”: (a) *E0*, (b) *E1*

⁷ According to Vriend [78] and Bergin and Bernhardt [79], social learning from the better-performing peers may (but not necessarily) induce spiteful behavior among agents. Thus, we also conduct additional experiments in which agents select strategies of the other peers with equal probabilities in modifying their IPD strategies. The results are similar to those reported here. This also holds true for the following experiments with strong social comparison (Section 5.3), in which the evolution of high levels of cooperation also fails. Thus, the resulting spiteful (defective) behavior is not due to the social learning mechanism in the model.

Fig. 3 presents the population average payoff by generation in the experiment *E1*. It is observed that the population average payoff stabilizes at about 2.5 at the end of evolution, verifying that the cooperation level in the IPD is higher with risk consideration than without [23]. Fig. 4 compares the percentages of moves “D-D”, “D-C” and “C-D”, and “C-C” performed by agents in *E0* and *E1*. Since performance stabilizes much earlier, only the result in the first 10,000 generations is shown. Compared with *E0*, *E1* is able to elicit the “C-C” interaction right from the start of the evolution, which is essential for the emergence of mutual cooperation in the IPD with short game durations [52]. This observation is fairly consistent in the subsequent experiments with dynamic risk consideration. On the other hand, although the percentage of “D-C” and “C-D” is higher than that of “D-D” at the beginning of *E1*, the latter exceeds the former from about the 1,000th generation onwards. Although the payoff by alterations of cooperation and defection is greater than that by mutual defection, agents tend to defect for preventing themselves from being outperformed by others. Thus, the population average payoff cannot be maintained higher than 2.6 in *E1* either.

Further inspection on the percentages of moves played by agents with different risk attitudes indicates that agents perform similarly well in *E1*. Moreover, no exploitation exists among agents in their mutual interactions. That is, no agents can unilaterally defect (“D-C”) against their opponents without being punished (“C-D”). The reason is that social learning mechanism facilitates information exchange among agents [36], which helps deter the emergence of unilateral exploitation in the population. This observation also holds true for the following experiments with agents of adaptive risk attitudes.

5.2 Dynamic Risk Consideration: Aspirations Consisting of only Historical Comparison (*E2*)

In this subsection, we examine the effect of dynamic risk consideration on the evolution of cooperation in the IPD game, in which agents’ aspirations are restricted to their past performances (*E2*).

5.2.1 Moderate Speed of Attitude Adaptation Leads to the Evolution of Cooperation

Fig. 5(a) illustrates the population average payoff for the 55 scenarios of *E2*. Results are averaged over 20,000 generations. Fig. 5(a) indicates that the parameter γ has an inverted U-shaped influence on the population average payoff. That is, moderate values of γ maximize the population average payoff in *E2*. This observation is fairly consistent for $r_{up} = r_{down} \in \{0.2, 0.4, 0.6, 0.8, 1.0\}$. In particular, the larger are the values of r_{up} and r_{down} , the higher is the value of γ , at which the population average payoff attains its peak value. Fig. 5(b) depicts the corresponding population average risk attitude, and it indicates also an inverted U-shaped relationship between γ and the population average risk attitude, which suggests a strongly positive correlation between risk aversion and cooperation in the IPD game (Fig. 5(c)).

Fig. 5. Results for the evolutionary IPD game with agents' aspirations determined as their past performances (*E2*): (a) population average payoff, (b) population average risk attitude, and (c) population average risk. Results are averaged over 20,000 generations

Fig. 6. Population average payoff by generation for *E2* under different speeds of adjustment of risk attitudes ($r_{up} = r_{down} = 0.4$)

A closer inspection of the population average payoff by generation reveals that *E2* with small value of γ exhibits a similar evolutionary behavior to that of *E0*. Fig. 6 illustrates such a phenomenon with $r_{up} = r_{down} = 0.4$ and $\gamma = 0$, in which the population average payoff is observed to increase to exceed 2.5 after the initial drop, but it then decreases and stabilizes at about 2.25. Moreover, agents tend to be highly risk-seeking in this case (Fig. 7(a)). Second, *E2* with γ close to 0.5 leads to a convergence of the population average payoff towards about 2.5, which is a similar game dynamic with *E1*. This is because few agents change their risk attitudes when γ is large enough (Fig. 7(c)). Third, if γ is chosen with the population average payoff attaining its peak value, agents will become highly risk-averse in the IPD game (Fig. 7(b)). In this case, the population average payoff will steadily increase with the generation counter after the initial drop, and finally stabilize above 2.6 at the end of evolution, indicating the successful formation of mutual cooperation (see the example with $\gamma = 0.15$ in Fig. 6).

Fig. 7. Proportions of different agents for $E2$ with $r_{up} = r_{down} = 0.4$ and (a) $\gamma = 0$, (b) $\gamma = 0.15$ and (c) $\gamma = 0.5$ (Data are averaged over every 400 generations)

Table 2
Strategy modification rate of the population in $E2$.

r_{up} (r_{down})	γ										
	0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
0.2	0.32±0.00	0.38±0.01	0.40±0.00	0.39±0.00	0.40±0.00	0.40±0.00	0.39±0.00	0.38±0.00	0.38±0.00	0.37±0.00	0.38±0.01
0.4	0.32±0.00	0.34±0.00	0.37±0.00	0.40±0.00	0.40±0.00	0.40±0.00	0.40±0.00	0.39±0.00	0.38±0.00	0.37±0.00	0.38±0.00
0.6	0.32±0.00	0.33±0.00	0.41±0.02	0.36±0.01	0.40±0.00	0.40±0.00	0.40±0.00	0.39±0.00	0.38±0.00	0.37±0.00	0.38±0.00
0.8	0.32±0.01	0.33±0.00	0.41±0.01	0.36±0.02	0.37±0.01	0.39±0.01	0.40±0.00	0.39±0.00	0.38±0.00	0.38±0.00	0.37±0.00
1	0.33±0.01	0.35±0.01	0.41±0.02	0.36±0.02	0.35±0.01	0.38±0.01	0.39±0.01	0.39±0.00	0.38±0.00	0.38±0.00	0.37±0.00

* Averages ± 95% confidence intervals: the averages are obtained from 20 independent runs of 20,000 generations and the 95% confidence intervals are computed from 20 independent runs. This also applies for Tables 3, 4, 5, and 6.

The above observations verify the fact that historical comparison promotes the evolution of cooperation among agents⁸. Nonetheless, an appropriate setting of γ is crucial for the learning of cooperative behavior. Since strategies are selected by agents regarding their risk preferences in the IPD game (Section 4.3.2), the parameter γ can have a direct impact on the effectiveness of the agents' strategy modification. A too fast ($\gamma \leq 0.05$) or a too slow ($\gamma \geq 0.4$) adjustment of risk attitude are not anticipated, because they may make the strategy modification too fast or too slow to closely adapt to the changes of the game environment across generations. As summarized in Table 2, the rates of strategy modification of the population are generally smaller at the extremes of the range of γ but larger in the middle. Thus, a too fast or a too slow adjustment of risk attitude will degrade

⁸ Tanabe and Masuda [80] also used historical comparison to define aspirations for individuals in the IPD game. However, different from our work, Tanabe and Masuda used the payoff of previous PD rounds to define the aspiration payoff of the current PD for an individual in the IPD game. They designed an on-line reinforcement learning mechanism in the IPD game, in which the propensity of playing the move C or D would be increased (or decreased) as long as the move brought an above-aspiration (or below-aspiration) payoff to an individual. Such a mechanism enabled the learning of cooperative behavior in the evolutionary IPD game. Tanabe and Masuda concluded their result as an affirmative example of the Baldwin effect, which states that learning accelerates the evolution to optimality.

the effectiveness of strategy learning of agents, which will certainly decrease their satisfactions with own performances and further decrease their willingness to reduce risk in the IPD game. As a result, high levels of cooperation become impossible in these scenarios.

5.2.2 Risk Consideration Counteracts the Spite Effect

As detailed in Section 4.1, agents play the IPD game based on histories of previous three moves. To further explain the different outcomes in the evolutionary games with different speeds of attitude adaptation of agents, we turn to examine the average cooperation rates of agents under different histories of previous three moves. Table 3 depicts the result for $E2$ with $r_{up} = r_{down} = 0.4$. Similar results can be obtained for other values of r_{up} and r_{down} as well. It is seen that agents are more cooperative under the moderate speeds of attitude adaptation than under the speeds that are too fast or too slow. In particular, in the scenarios with γ ranging from 0.15 to 0.25, in which the population tend to be highly risk-averse in the evolution (Fig. 5(b)), agents generally cooperate with a rate greater than 0.6 after the moves of “D-D” or “D-C”, except for the case of three “D-C”, in which opponents are so foolish to be exploited thrice in the previous three moves. Thus, the contrition for defection is prevalent among agents when they are highly risk-averse in the IPD. This contrition successfully counteracts the spite effect, and ultimately enhances the evolution of high levels of cooperation.

Table 3

Cooperation rates of agents under different histories of previous three moves in $E2$ with $r_{up} = r_{down} = 0.4$.

		γ										
		0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
Number of “D-D” in Previous Three Moves	0	0.22±0.01	0.34±0.01	0.37±0.01	0.69±0.01	0.69±0.01	0.63±0.01	0.55±0.01	0.44±0.01	0.40±0.02	0.37±0.01	0.38±0.02
	1	0.17±0.01	0.28±0.01	0.34±0.03	0.74±0.01	0.74±0.01	0.67±0.01	0.57±0.01	0.41±0.02	0.37±0.03	0.32±0.01	0.34±0.02
	2	0.14±0.01	0.19±0.01	0.49±0.00	0.75±0.00	0.75±0.00	0.67±0.00	0.60±0.00	0.53±0.01	0.51±0.01	0.50±0.00	0.50±0.01
	3	0.99±0.00	0.99±0.00	0.95±0.01	0.98±0.00	0.98±0.00	0.98±0.00	0.97±0.00	0.96±0.00	0.95±0.00	0.95±0.01	0.95±0.01
Number of “D-C” in Previous Three Moves	0	0.30±0.01	0.41±0.01	0.50±0.01	0.73±0.00	0.73±0.00	0.68±0.00	0.63±0.00	0.55±0.01	0.53±0.01	0.51±0.00	0.52±0.01
	1	0.14±0.01	0.26±0.01	0.34±0.02	0.76±0.01	0.76±0.01	0.68±0.01	0.58±0.01	0.42±0.02	0.37±0.03	0.33±0.01	0.35±0.02
	2	0.08±0.02	0.16±0.02	0.18±0.03	0.66±0.01	0.66±0.01	0.58±0.02	0.45±0.01	0.26±0.02	0.21±0.04	0.16±0.01	0.18±0.03
	3	0.01±0.01	0.02±0.01	0.02±0.01	0.12±0.00	0.11±0.01	0.10±0.01	0.08±0.01	0.03±0.01	0.02±0.01	0.01±0.00	0.01±0.01
Number of “C-D” in Previous Three Moves	0	0.31±0.01	0.44±0.01	0.53±0.01	0.82±0.00	0.82±0.00	0.77±0.01	0.70±0.01	0.60±0.01	0.57±0.02	0.54±0.01	0.55±0.01
	1	0.14±0.01	0.24±0.01	0.32±0.02	0.73±0.01	0.73±0.01	0.65±0.01	0.54±0.01	0.39±0.02	0.34±0.03	0.31±0.01	0.32±0.02
	2	0.07±0.01	0.13±0.01	0.17±0.02	0.52±0.01	0.52±0.01	0.44±0.01	0.35±0.01	0.22±0.01	0.19±0.03	0.16±0.01	0.17±0.02
	3	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
Number of “C-C” in Previous Three Moves	0	0.06±0.01	0.09±0.01	0.18±0.02	0.58±0.01	0.57±0.01	0.49±0.01	0.38±0.01	0.23±0.02	0.19±0.03	0.16±0.01	0.17±0.02
	1	0.19±0.02	0.36±0.01	0.44±0.02	0.81±0.01	0.81±0.01	0.74±0.01	0.66±0.01	0.53±0.01	0.48±0.03	0.44±0.01	0.46±0.02
	2	0.54±0.02	0.72±0.01	0.75±0.00	0.91±0.00	0.91±0.00	0.88±0.00	0.85±0.00	0.81±0.00	0.79±0.01	0.77±0.00	0.77±0.01
	3	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00

Table 4

Percentage of moves for agents with risk attitude α_1 to play against agents with risk attitude α_2 in $E2$ with $r_{up} = r_{down} = 0.4$ and $\gamma = 0.15$.

α_1	α_2	“D-D”	“D-C”	“C-D”	“C-C”
0~0.2	0~0.2	0.11±0.00	0.12±0.00	0.12±0.00	0.65±0.00
	0.2~0.4	0.09±0.00	0.12±0.00	0.11±0.00	0.68±0.00
	0.4~0.6	0.09±0.00	0.12±0.00	0.11±0.00	0.68±0.00
	0.6~0.8	0.09±0.00	0.12±0.00	0.11±0.00	0.68±0.00
	0.8~1.0	0.09±0.00	0.12±0.00	0.11±0.00	0.68±0.00
0.2~0.4	0.2~0.4	0.09±0.00	0.11±0.00	0.11±0.00	0.69±0.00
	0.4~0.6	0.09±0.00	0.11±0.00	0.11±0.00	0.69±0.00
	0.6~0.8	0.09±0.00	0.11±0.00	0.11±0.00	0.69±0.00
	0.8~1.0	0.08±0.00	0.11±0.00	0.11±0.00	0.70±0.00
0.4~0.6	0.4~0.6	0.08±0.00	0.11±0.00	0.11±0.00	0.70±0.00
	0.6~0.8	0.08±0.00	0.11±0.00	0.11±0.00	0.70±0.00
	0.8~1.0	0.08±0.00	0.11±0.00	0.10±0.00	0.71±0.00
0.6~0.8	0.6~0.8	0.08±0.00	0.11±0.00	0.11±0.00	0.70±0.00
	0.8~1.0	0.08±0.00	0.11±0.00	0.10±0.00	0.71±0.00
0.8~1.0	0.8~1.0	0.08±0.00	0.10±0.00	0.10±0.00	0.72±0.00

The above observation can be further confirmed by comparing the percentages of the moves “D-D”, “D-C” and “C-D”, and “C-C” performed by different agents. Table 4 illustrates the relevant results in $E2$ with $r_{up} = r_{down} = 0.4$ and $\gamma = 0.15$. Data are averaged over 20,000 generations⁹. It is clearly shown that the percentage of the move “C-C” is consistently higher than that of “D-C” and “C-D”, which is consistently higher than that of “D-D” in the interactions between all agents, implying the successful inhibition of spiteful behavior in this circumstance. Nonetheless, this finding does not hold when γ is too small or too large. Moreover, compared with the result in $E1$, more “C-C” are performed between agents with $0.2 \leq \alpha \leq 1$ in this scenario. Thus, mutual cooperation become more valued when highly risk-seeking agents ($0 \leq \alpha < 0.2$) contribute only a small part of the population in the evolutionary game (Fig. 7(b)).

5.3 Experiment 3: Aspirations Involving Social Comparison ($E3$)

Social comparison is a fundamental characteristic of human interactions, in which individuals often assess their own performances by comparing with others to ascertain the validity of their behavior [69]. Therefore, in following experiments, social comparison, in addition to historical comparison, is also employed as a component for agents to determine their satisfactions with their payoffs in the evolutionary IPD game ($E3$). The parameter β is varied from 0.05 to 1 to investigate the cooperation levels that evolve in the population.

⁹ Generations with the absence of agents with particular risk attitudes have been excluded from the averaging process for the corresponding agents. This operation also applies for Table 5.

5.3.1. Social Comparison Harms the Evolution of Cooperation

Fig. 8. Results for the evolutionary IPD game with agents' aspirations determined by both their past performances and the population average performance ($\beta = 0.5$): (a) population average payoff, (b) population average risk attitude, and (c) population average risk. Results are averaged over 20,000 generations

The experiment $E3$ with $\beta \leq 0.1$ exhibits similar result to that of $E2$, but larger values of β lead to lower levels of cooperation of agents in $E3$. Figs. 8(a), (b), and (c) illustrate the population average payoff, population average risk attitude, and population average risk, respectively, for $E3$ with $\beta = 0.5$. Results are averaged over 20,000 generations. First, mutual cooperation is impossible in the evolutionary game when agents depend too much on social comparison in adapting their risk attitudes. This phenomenon verifies the fact that social comparison of one's performance to others rewards defective behavior¹⁰.

Fig. 9. Population average payoff by generation for $E3$ with different speeds of risk attitude adaptation for agents ($\beta = 0.5, r_{up} = r_{down} = 0.4$)

¹⁰ In [75], Zschache also used social comparison to measure agents' satisfaction in a network-based public good game (i.e., N-person PD game). The satisfaction (or dissatisfaction) of agents was related to strengthening (or weakening) their current behavior and existing relationships. He showed that cooperation could prevail under the social comparison in this network-based public good game, which differed from the result of our experiment $E3$. Note that the endogenous relationship adjustments played an important role in the evolution of cooperation in [75] and its effect on the evolutionary IPD game should be further investigated.

Second, different from the case with $\beta = 0$, in which the value of γ has an inverted U-shaped influence on the population average payoff, a U-shaped relationship is found between γ and the population average payoff (also the population average risk attitude) when $\beta = 0.5$. That is, a too fast or a too slow adjustment of risk attitude would have a negative impact on the spread of the defective behavior in the population when social comparison exists among agents. This outcome follows the same reason as the survival of cooperative behavior in *E2*, which is associated with the validity of strategy learning behavior of agents [81]. As shown by the example with $r_{up} = r_{down} = 0.4$ in Fig. 9, an appropriate speed of attitude adaptation ($\gamma = 0.1$) that assures the accuracy of strategy modification of agents will accelerate the evolution of defective behavior in the social comparison process. On the contrary, a too high ($\gamma = 0$) or too low speed ($\gamma = 0.5$) of attitude adaptation will prevent the defection. Thus, a certain level of cooperation can emerge in the population in these two cases, but none with population average payoff stabilizing higher than 2.6.

5.3.2. Why Social Comparison Destroys Mutual Cooperation

Fig. 10(a) depicts the population average risk attitude (averaged over 20,000 generations) for *E3* with different settings of β . The parameters r_{up} and r_{down} are fixed as 0.4. Similar results can be obtained for other values of r_{up} and r_{down} . To illustrate the effect of attitude adaptation on the learning of the cooperative or the defective behavior, we compute the average results of $\gamma = 0.1$, 0.15, and 0.2 that indicate different speeds of attitude adaptation.

Fig. 10. Change of (a) the population average risk attitude, and (b) the population average payoff with respect to β in the evolutionary IPD game with $r_{up} = r_{down} = 0.4$. Data are averaged over 20,000 generations and further averaged with $\gamma = 0.1, 0.15$, and 0.2 , at which the cooperative (or defective) behavior can be adequately learned by agents.

It is observed that larger values of β give rise to lower population average risk attitudes in the evolutionary IPD game. Thus, agents tend to be more risk seeking when they base their risk attitude adaptation more on social comparison. Although agents performing better than their aspirations (containing the population averages) tend to increase their risk aversion levels, their opponents

performing worse than aspirations (containing the population averages) become more risk seeking. Thus, mutual cooperation may fail in the population with strong social comparison, because it can only be established when both players are risk averse in the IPD game. With the increasing failure of cooperation, agents (including those previously performing better than their aspirations) will become more dissatisfied with the game outcomes, and the whole population will tend to be more risk seeking as the evolution proceeds. Therefore, the evolution of mutual cooperation becomes difficult in the evolutionary IPD game with larger β (Fig. 10(b)).

Table 5 further reports the percentages of moves performed by agents in $E3$ with $\beta = 0.5$, $r_{up} = r_{down} = 0.4$, and $\gamma = 0.1$. The percentage of the move “C-C” is smaller than 0.6 in the interactions between all agents, which is similar to the result in $E0$ (Fig. 4(a)). Moreover, the percentage of the move “D-D” is greater than that of “D-C” and “C-D” in the interactions between agents with $0 \leq \alpha < 0.2$, who make up most of the population in the evolution. Thus, social comparison makes agents tend to be aggressive, which fuels the spite effect in the IPD game. This outcome is fairly consistent for β ranging from 0.2 to 1. In particular, the contrition for defection is lacking among agents. For example, in $E3$ with $\beta = 1$ and $r_{up} = r_{down} = 0.4$ and $\gamma = 0.15$, agents generally cooperate with a rate lower than 0.3 after the move “D-D” or “D-C”, except for the case in which they cooperate with a rate of about 0.9 following three moves of “D-D” to escape from endless mutual defections (Table 6). Note that the rate of 0.3 is only half of that (0.6) in $E2$ with $r_{up} = r_{down} = 0.4$ and $\gamma = 0.15$ as well (Table 3). These observations present an explicit explanation for the failure of high levels of cooperation under social comparison.

Table 5

Percentage of moves for agents with risk attitude α_1 to play against agents with risk attitude α_2 in $E3$ with

$$r_{up} = r_{down} = 0.4, \beta = 0.5 \text{ and } \gamma = 0.1.$$

α_1	α_2	“D-D”	“D-C”	“C-D”	“C-C”
0~0.2	0~0.2	0.33±0.00	0.05±0.00	0.05±0.00	0.57±0.00
	0.2~0.4	0.24±0.02	0.11±0.01	0.10±0.01	0.55±0.01
	0.4~0.6	0.21±0.02	0.13±0.01	0.12±0.01	0.54±0.01
	0.6~0.8	0.19±0.02	0.14±0.01	0.13±0.01	0.54±0.01
	0.8~1.0	0.19±0.02	0.14±0.01	0.13±0.01	0.54±0.01
0.2~0.4	0.2~0.4	0.21±0.02	0.12±0.01	0.12±0.01	0.55±0.01
	0.4~0.6	0.20±0.02	0.12±0.01	0.12±0.01	0.56±0.01
	0.6~0.8	0.18±0.02	0.13±0.01	0.13±0.01	0.56±0.01
	0.8~1.0	0.18±0.02	0.13±0.01	0.13±0.01	0.56±0.02
0.4~0.6	0.4~0.6	0.18±0.02	0.13±0.01	0.13±0.01	0.56±0.01
	0.6~0.8	0.18±0.02	0.13±0.01	0.13±0.01	0.56±0.01
	0.8~1.0	0.18±0.02	0.13±0.01	0.13±0.01	0.56±0.02
0.6~0.8	0.6~0.8	0.17±0.02	0.13±0.01	0.13±0.01	0.57±0.02
	0.8~1.0	0.17±0.02	0.13±0.01	0.13±0.01	0.57±0.02
0.8~1.0	0.8~1.0	0.17±0.02	0.13±0.00	0.13±0.00	0.57±0.02

Table 6

Cooperation rates of agents under different histories of previous three moves in $E3$ with $\beta = 1$ and

$$r_{up} = r_{down} = 0.4.$$

		γ										
		0	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
Number of “D-D” in Previous Three Moves	0	0.48±0.06	0.15±0.03	0.12±0.01	0.13±0.02	0.16±0.01	0.27±0.06	0.34±0.03	0.38±0.03	0.37±0.02	0.37±0.01	0.38±0.03
	1	0.46±0.07	0.12±0.02	0.10±0.01	0.12±0.02	0.15±0.02	0.24±0.07	0.29±0.04	0.34±0.04	0.33±0.03	0.32±0.02	0.35±0.04
	2	0.35±0.02	0.13±0.01	0.12±0.01	0.17±0.07	0.46±0.00	0.47±0.02	0.48±0.01	0.50±0.01	0.50±0.01	0.50±0.01	0.50±0.01
	3	0.96±0.02	0.99±0.00	1.00±0.00	0.97±0.06	0.74±0.02	0.87±0.02	0.94±0.01	0.95±0.00	0.95±0.01	0.94±0.01	0.95±0.01
Number of “D-C” in Previous Three Moves	0	0.50±0.03	0.24±0.02	0.22±0.01	0.24±0.01	0.32±0.01	0.41±0.04	0.48±0.02	0.51±0.02	0.51±0.01	0.51±0.01	0.51±0.02
	1	0.45±0.07	0.09±0.03	0.07±0.02	0.09±0.02	0.15±0.02	0.24±0.07	0.30±0.04	0.35±0.04	0.34±0.03	0.33±0.01	0.35±0.04
	2	0.41±0.08	0.05±0.03	0.04±0.02	0.05±0.02	0.07±0.01	0.13±0.08	0.14±0.04	0.19±0.05	0.18±0.04	0.16±0.02	0.19±0.04
	3	0.14±0.04	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.00	0.02±0.02	0.01±0.01	0.02±0.01	0.02±0.01	0.01±0.01	0.02±0.01
Number of “C-D” in Previous Three Moves	0	0.59±0.05	0.25±0.03	0.22±0.01	0.25±0.02	0.33±0.01	0.43±0.06	0.50±0.03	0.54±0.03	0.54±0.02	0.53±0.01	0.54±0.02
	1	0.41±0.06	0.09±0.02	0.07±0.01	0.08±0.02	0.14±0.01	0.23±0.06	0.28±0.04	0.32±0.04	0.32±0.03	0.30±0.01	0.33±0.03
	2	0.26±0.05	0.04±0.02	0.03±0.01	0.04±0.02	0.06±0.01	0.12±0.05	0.14±0.03	0.18±0.03	0.17±0.02	0.16±0.01	0.18±0.03
	3	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
Number of “C-C” in Previous Three Moves	0	0.27±0.04	0.06±0.02	0.05±0.01	0.07±0.02	0.14±0.01	0.16±0.05	0.15±0.03	0.18±0.03	0.17±0.03	0.16±0.02	0.18±0.03
	1	0.53±0.07	0.13±0.03	0.10±0.02	0.12±0.02	0.19±0.02	0.30±0.07	0.39±0.04	0.45±0.03	0.45±0.02	0.44±0.01	0.46±0.03
	2	0.76±0.05	0.40±0.03	0.32±0.02	0.34±0.03	0.38±0.02	0.59±0.05	0.73±0.02	0.76±0.01	0.77±0.01	0.77±0.00	0.77±0.01
	3	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00

5.4 Experiment 4: Concern with the Rule of “losses loom larger than gains” (E4)

Thus far, we have assumed that agents adjust their risk attitudes upwards and downwards by the same amount while experiencing gains and losses ($r_{up} = r_{down}$). According to prospect theory [26], individuals are always loss averse, and thus they dislike a loss more than they like a gain of equivalent amount. Thus, in the following experiment, we fix r_{up} as 0.4 and change the value of r_{down}/r_{up} from 1 to 2 in steps of 0.2 to examine any change of the cooperative outcome in the evolutionary IPD game (E4). The population average payoffs (averaged over 20,000 generations) are shown in Fig. 11.

Fig. 11. Population average payoff for $E4$ with $r_{up} = 0.4$ and (a) $\beta = 0$, (b) $\beta = 1$. Results are averaged over 20,000 generations

Comparable results are obtained for the six specifications of the ratio r_{down}/r_{up} . That is, an increasing γ has an inverted U-shaped influence on the evolution of cooperation in the evolutionary IPD game when $\beta = 0$, but a U-shaped influence when $\beta = 1$. In particular, a larger value of r_{down}/r_{up} leads to a larger value of γ at which the population average payoff attains its peak or valley values. Moreover, when $\beta = 0$, the influence of γ on the population average payoff becomes relatively modest with larger values of r_{down}/r_{up} , which interferes with the evolution of cooperation. This observation suggests that the positive impact of risk consideration on mutual cooperation would be weakened when agents adjust their risk attitudes at a higher rate in the downward direction than in the upward one. When r_{down} is much larger than r_{up} , risk-averse agents constitute only a small proportion in the population, meaning that agents will be unable to maximize joint gains in their mutual interactions. Thus, the evolution of cooperation becomes difficult.

5.5 Discussion

Risky decision making has been a central topic in the study of individual behavior. In this study, we carried out a group of simulation tests to determine the impact of dynamic risk consideration on the emergence of cooperation in the evolutionary IPD game, which has been widely used to model the interaction among competitors [22, 44]. Based on the empirical findings on human risky behavior [26, 28], agents were designed to become more risk-averse after obtaining a satisfying performance, but get more risk-seeking after experiencing a dissatisfying one. Detailed analyses of the induced population dynamics were provided under different settings of parameters in the evolutionary model. The learning of cooperative behavior is attributed partially to the way in which agents evaluate satisfaction and dissatisfaction (in terms of historical comparison or social comparison), and partially to the speed that agents adjust their risk attitudes.

Previous studies have shown that cooperation is difficult to achieve in the IPD with short game encounters [6, 51] because reciprocal altruism cannot be fostered in a short interaction. Our studies, however, suggested that mutual cooperation is still possible in the short-encounter IPD game when agents take a consideration of the income stream risk in their interactions. In the historical comparison, agents evaluate their current payoffs against their past payoffs, and adapt to be more (or less) concerned with the risk factor when their current payoffs are higher (or lower) than their past payoffs. Such a dynamic risk consideration can inhibit the emergence of spiteful behavior in the IPD game and induce agents to cultivate and maintain stable cooperation with each other. Thus, the common pursuit of improvement in self-payoff motivates the high-level cooperation. This finding is highly in accordance with the results of behavioral experiments in [32], which concluded that subjects cooperating in the IPD game were primarily motivated by own-payoff maximization.

Nonetheless, the high level of cooperation (under historical comparison) occurs only when the speed by which agents adjust risk attitudes is neither too fast nor too slow. The reason is that an inappropriate adjustment speed would prevent agents from adapting to the game environment. A too fast speed would make risk-averse agents switch to risk-seeking incentives before risk-seeking agents are converted to be more risk-averse, which is difficult for cooperation to be rewarded and maintained. On the other hand, when the adjustment speed is too slow, individuals are unwilling to change their risk attitudes and the cooperation just stays at an intermediate level. These outcomes highlight the significance of effective learning in the evolution of cooperation in the IPD game, which is in accordance with the simulation results in [44][82]. Moreover, these outcomes manifest the fact that sometimes people attempt to pursue their interests but fail because of the way they do it.

Individuals always compare their performances with others to assess their own status in the real world. For instance, managers are likely to compare their firms' profitability to the industry average to make plans for future improvements. This process has been construed as social comparison [69]. Therefore, we further conducted experiments to examine the impact of social comparison on the evolution of cooperation, in which agents adjust their risk attitudes based on the comparison of their own payoffs to the population averages. Results illustrate that high levels of cooperation fail with a strong social comparison. The reason is that agents with payoffs below the population averages tend to be dissatisfied with the game outcomes under social comparison and thus become more spiteful, which deters mutual cooperation in the IPD game. This outcome echoes the previous psychological findings that social comparison fuels competition [83] and harms cooperation [84]. On the other hand, the level of cooperation would increase when agents adapt their risk attitudes at a too fast or a too slow speed, but it still stays below the threshold 2.6 that indicates mutual cooperation.

In addition to the social learning mechanism used in the strategy modification of agents, which has been widely accepted by evolutionary game theorists as a move towards greater realism [12, 14], our model further integrated the goal-oriented risk attitude adaptation of agents in the evolutionary IPD game, which allows a direct investigation of the effects of historical comparison and social

comparison on the evolution of cooperation. Moreover, the results from the evolutionary experiments also verify the significance of effective leaning in the emergence of cooperation [81]. An ineffective leaning of agents may decrease (or increase) the level of cooperation in evolutionary IPD game with historical (or social) comparison, in which mutual cooperation is (or not) likely to emerge.

Our work demonstrates that the multi-agent modeling can be a fertile approach to theory development for scientists interested in human social phenomena, and the results from our evolutionary IPD model also have some interesting implications. We suggest that those concerned with income stream stability will be more likely to engage in mutual cooperation via historical comparison. On the other hand, social comparison can mitigate the concern of stability by causing unnecessary discontentment, and thus induce uncooperative behavior among individuals. Thus, an economic policy maker should care about the steady development of individuals to advance cooperation, and policies oriented to long-term sustainable development of individuals should be rewarded and supported.

6. Conclusion

This study introduced the concept of income stream risk into the evolutionary IPD model. Agents play pair-wise IPD games and adapt their risk attitudes in response to relative performance feedback. A detailed simulation study has been carried out to explore the relationship between the dynamic risk consideration and the evolution of cooperation. Results indicate that mutual cooperation is possible when agents focus on historical comparison to adapt their risk attitudes, but fails with the strong presence of social comparison. The reason is that comparing one's own performance with others reduces the satisfaction of agents whose payoffs are lower than the average of the population, and thus discourages their incentives to reduce income stream risk, which is determinative to the emergence of high-level cooperation in the IPD game.

Our work can be extended in several directions. First, we assumed that agents play the IPD game with randomly chosen opponents in the population. In fact, individuals may have some degree of autonomy with whom they are linked to. The mechanism of kin-biased or spatial-biased interaction can be incorporated. Second, the social comparison process in our model can be modified to allow agents to evaluate themselves with similar ones, not the whole population. This assumption seems to echo well with the real-world situations, in which individuals often determine how they should behave in the context of reference groups.

Acknowledgments

This work was supported by the National Science Fund for Distinguished Young Scholars of China (Grant No. 70925005), the General Program of the National Science Foundation of China (No.71371135), and the Scientific Research Foundation of Hainan University, China (No. kyqd1529).

It was also supported by the Program for Changjiang Scholars and Innovative Research Teams in Universities of China (PCSIRT). The authors would like to thank The High Performance Computing Centre (HPCC) of Tianjin University for providing computing support.

Reference

- [1] J. Zhang, G.V. Frazier, Strategic alliance via co-opetition: Supply chain partnership with a competitor, *Decision Support Systems* 51 (4) (2011) 853-863.
- [2] P.R. Varadarajan, M.H. Cunningham, Strategic Alliances: A Synthesis of Conceptual Foundations, *Journal of the Academy of Marketing Science* 23 (4) (1995) 282-296.
- [3] A.M. Brandenburger, B.J. Nalebuff, *Co-opetition: a revolution mindset that combines competition and cooperation*, Harvard Business Press, Cambridge, MA, 1996.
- [4] R. Axelrod, *The Evolution of Cooperation*, Basic Books, New York, 1984.
- [5] C. Fosco, F. Mengel, Cooperation through imitation and exclusion in networks, *Journal of Economic Dynamics and Control* 35 (5) (2011) 641-658.
- [6] D.B. Fogel, On the relationship between the duration of an encounter and the evolution of cooperation in the iterated prisoner's dilemma, *Evolutionary Computation* 3 (3) (1995) 349-363.
- [7] H. Ishibuchi, H. Ohyanagi, Y. Nojima, Evolution of Strategies With Different Representation Schemes in a Spatial Iterated Prisoner's Dilemma Game, *IEEE Transactions on Computational Intelligence and AI in Games* 3 (1) (2011) 67-82.
- [8] R. Chiong, M. Kirley, Effects of iterated interactions in multiplayer spatial evolutionary games, *IEEE Transactions on Evolutionary Computation* 16 (4) (2012) 537-555.
- [9] Z. Qin, T. Wan, Y. Dong, Y. Du, Evolutionary collective behavior decomposition model for time series data mining, *Applied Soft Computing* (in press).
- [10] L. Tesfatsion, Agent-based computational economics: modeling economies as complex adaptive systems, *Information Sciences* 149 (4) (2003) 262-268.
- [11] X. Thibert-Plante, P. Charbonneau, Crossover and evolutionary stability in the Prisoner's Dilemma, *Evolutionary Computation* 15 (3) (2007) 321-344.
- [12] J.W. Weibull, *Evolutionary game theory*, MIT press, Cambridge, MA, 1997.
- [13] H. Dawid, *Adaptive learning by genetic algorithms: analytical results and applications to economic models*, Springer-Verlag, New York, 1996.
- [14] G.M. Hodgson, K. Huang, Evolutionary game theory and evolutionary economics: are they different species?, *Journal of Evolutionary Economics* 22 (2) (2012) 345-366.
- [15] M. Tomasello, *The Cultural Origins of Human Cognition*, Harvard University Press, Cambridge, MA, 2000.
- [16] C.F. Camerer, *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press, Princeton, 2003.
- [17] S.-H. Chen, Varieties of agents in agent-based computational economics: A historical and an interdisciplinary perspective, *Journal of Economic Dynamics and Control* 36 (1) (2012) 1-25.
- [18] S.-H. Chen, Software-Agent Designs in Economics: An Interdisciplinary, *IEEE Computational Intelligence Magazine* 3 (4) (2008) 18-22.

- [19] D. Fudenberg, E. Maskin, The folk theorem in repeated games with discounting or with incomplete information, *Econometrica* 54 (3) (1986) 533-554.
- [20] C. Snijders, W. Raub, Revolution and risk: Paradoxical consequences of risk aversion in interdependent situations, *Rationality and Society* 10 (4) (1998) 405-425.
- [21] M. van Assen, C. Snijders, The effect of nonlinear utility on behaviour in repeated prisoner's dilemmas, *Rationality and Society* 22 (3) (2010) 301-332.
- [22] H.D. Dixon, Keeping up with the Joneses: competition and the evolution of collusion, *Journal of Economic Behavior & Organization* 43 (2) (2000) 223-238.
- [23] W. Zeng, M. Li, F. Chen, G. Nan, Risk consideration and cooperation in the iterated prisoner's dilemma, *Soft Computing* (in press).
- [24] K.D. Miller, P. Bromiley, Strategic risk and corporate performance: An analysis of alternative risk measures, *Academy of Management Journal* 33 (4) (1990) 756-779.
- [25] G.A. Holton, Defining risk, *Financial Analysts Journal* 60 (6) (2004) 19-25.
- [26] D. Kahneman, A. Tversky, Prospect theory: An analysis of decision under risk, *Econometrica: Journal of the Econometric Society* (1979) 263-291.
- [27] A. Fiegenbaum, H. Thomas, Strategic risk and competitive advantage: an integrative perspective, *European Management Review* 1 (1) (2004) 84-95.
- [28] R.M. Cyert, J.G. March, A behavioral theory of the firm, Prentice Hall, Englewood Cliffs, NJ, 1963.
- [29] A. Fiegenbaum, Prospect theory and the risk-return association: An empirical examination in 85 industries, *Journal of Economic Behavior & Organization* 14 (2) (1990) 187-203.
- [30] H. Oliveira Jr, A. Petraglia, Establishing Nash equilibria of strategic games: a multistart Fuzzy Adaptive Simulated Annealing approach, *Applied Soft Computing* 19 (2014) 188-197.
- [31] H. Ohtsuki, C. Hauert, E. Lieberman, M.A. Nowak, A simple rule for the evolution of cooperation on graphs and social networks, *Nature* 441 (7092) (2006) 502-505.
- [32] A. Dreber, D. Fudenberg, D.G. Rand, Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics, *Journal of Economic Behavior & Organization* 98 (2014) 41-55.
- [33] C. Ioannou, Coevolution of finite automata with errors, *Journal of Evolutionary Economics* 24 (3) (2014) 541-571.
- [34] N. Hanaki, A. Peterhansl, P.S. Dodds, D.J. Watts, Cooperation in evolving social networks, *Management Science* 53 (7) (2007) 1036-1050.
- [35] S. Suzuki, K. Niki, S. Fujisaki, E. Akiyama, Neural basis of conditional cooperation, *Social Cognitive and Affective Neuroscience* 6 (3) (2011) 338-347.
- [36] H.-Y. Quek, K.C. Tan, C.-K. Goh, H.A. Abbass, Evolution and incremental learning in the iterated prisoner's dilemma, *IEEE Transactions on Evolutionary Computation* 13 (2) (2009) 303-320.
- [37] K. Safarzyńska, J.C. van den Bergh, Evolutionary models in economics: a survey of methods and building blocks, *Journal of Evolutionary Economics* 20 (3) (2010) 329-373.
- [38] M.A. Nowak, Five rules for the evolution of cooperation, *science* 314 (5805) (2006) 1560-1563.
- [39] H. Zhang, M. Gao, W. Wang, Z. Liu, Evolutionary prisoner's dilemma game on graphs and social networks with external constraint, *Journal of Theoretical Biology* 358 (0) (2014) 122-131.
- [40] L.R. Izquierdo, S.S. Izquierdo, F. Vega-Redondo, Leave and let leave: A sufficient condition to explain the evolutionary emergence of cooperation, *Journal of Economic Dynamics and Control* 46 (0) (2014) 91-113.
- [41] M.A. Nowak, K. Sigmund, Evolution of indirect reciprocity by image scoring, *Nature* 393 (6685) (1998) 573-577.
- [42] D. Stahl, Cooperation in the sporadically repeated prisoners' dilemma via reputation mechanisms, *Journal of Evolutionary Economics* 21 (4) (2011) 687-702.
- [43] D.O. Stahl, An experimental test of the efficacy of a simple reputation mechanism to solve social dilemmas, *Journal*

of Economic Behavior & Organization 94 (0) (2013) 116-124.

[44] J. Barr, F. Saraceno, Organization, learning and cooperation, *Journal of Economic Behavior & Organization* 70 (1) (2009) 39-53.

[45] U. Fischbacher, S. Gächter, E. Fehr, Are people conditionally cooperative? Evidence from a public goods experiment, *Economics Letters* 71 (3) (2001) 397-404.

[46] E. Fehr, U. Fischbacher, The nature of human altruism, *Nature* 425 (6960) (2003) 785-791.

[47] M.A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Emergence of cooperation and evolutionary stability in finite populations, *Nature* 428 (6983) (2004) 646-650.

[48] J. Moreira, J. Vukov, C. Sousa, F.C. Santos, A.F. d'Almeida, M.D. Santos, J.M. Pacheco, Individual memory and the emergence of cooperation, *Animal Behaviour* 85 (1) (2013) 233-239.

[49] F. Alkemade, D. Van Bragt, J.A. La Poutré, Stabilization of tag-mediated interaction by sexual reproduction in an evolutionary agent system, *Information Sciences* 170 (1) (2005) 101-119.

[50] R.L. Riolo, M.D. Cohen, R. Axelrod, Evolution of cooperation without reciprocity, *Nature* 414 (6862) (2001) 441-443.

[51] H. Ishibuchi, N. Namikawa, Evolution of iterated prisoner's dilemma game strategies in structured demes under random pairing in game playing, *IEEE Transactions on Evolutionary Computation* 9 (6) (2005) 552-561.

[52] S.Y. Chong, X. Yao, Multiple choices and reputation in multiagent interactions, *IEEE Transactions on Evolutionary Computation* 11 (6) (2007) 689-711.

[53] P.G. Audia, H.R. Greve, Less likely to fail: Low performance, firm size, and factory expansion in the shipbuilding industry, *Management Science* 52 (1) (2006) 83-94.

[54] K. Wennberg, C. Holmquist, Problemistic search and international entrepreneurship, *European Management Journal* 26 (6) (2008) 441-454.

[55] B. Vis, K. Van Kersbergen, Why and how do political actors pursue risky reforms? *Journal of Theoretical Politics* 19 (2) (2007) 153-172.

[56] R.H. Thaler, *Advances in behavioral finance*, Princeton University Press, Princeton, 2005.

[57] H.R. Greve, *Organizational learning from performance feedback: A behavioral perspective on innovation and change*, Cambridge University Press, Cambridge, 2003.

[58] R.Z. Gooding, S. Goel, R.M. Wiseman, Fixed versus variable reference points in the risk-return relationship, *Journal of Economic Behavior & Organization* 29 (2) (1996) 331-350.

[59] J.P. Forgas, On feeling good and getting your way: mood effects on negotiator cognition and bargaining strategies, *Journal of personality and social psychology* 74 (3) (1998) 565.

[60] S. Williams, W.T. Shiaw, Mood and organizational citizenship behavior: The effects of positive affect on employee organizational citizenship behavior intentions, *The Journal of Psychology* 133 (6) (1999) 656-668.

[61] Y. Feldman, A. Schurr, D. Teichman, Reference Points and Contractual Choices: An Experimental Examination, *Journal of Empirical Legal Studies* 10 (3) (2013) 512-541.

[62] L.M. Hercog, Better manufacturing process organization using multi-agent self-organization and co-evolutionary classifier systems: The multibar problem, *Applied Soft Computing* 13 (3) (2013) 1407-1418.

[63] N. Barbalios, P. Tzionas, A robust approach for multi-agent natural resource allocation based on stochastic optimization algorithms, *Applied Soft Computing* 18 (2014) 12-24.

[64] A. Tversky, D. Kahneman, Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and uncertainty* 5 (4) (1992) 297-323.

[65] S. Mittal, K. Deb, Optimal strategies of the iterated prisoner's dilemma problem for multiple conflicting objectives, *IEEE Transactions on Evolutionary Computation* 13 (3) (2009) 554-565.

[66] N. Franken, A.P. Engelbrecht, Particle swarm optimization approaches to coevolve strategies for the iterated

prisoner's dilemma, *IEEE Transactions on Evolutionary Computation* 9 (6) (2005) 562-579.

[67] S.Y. Chong, X. Yao, Behavioral diversity, choices and noise in the iterated prisoner's dilemma, *IEEE Transactions on Evolutionary Computation* 9 (6) (2005) 540-551.

[68] S.T. Shuttters, Punishment leads to cooperative behavior in structured societies, *Evolutionary computation* 20 (2) (2012) 301-319.

[69] L. Festinger, A theory of social comparison processes, *Human relations* 7 (2) (1954) 117-140.

[70] T. Mussweiler, Comparison processes in social judgment: mechanisms and consequences, *Psychological review* 110 (3) (2003) 472.

[71] S.E. Hill, D.M. Buss, Risk and relative social rank: positional concerns and risky shifts in probabilistic decision-making, *Evolution and Human Behavior* 31 (3) (2010) 219-226.

[72] M. Washburn, P. Bromiley, Comparing aspiration models: the role of selective attention, *Journal of Management Studies* 49 (5) (2012) 896-917.

[73] J.A. Baggio, E. Papyrakis, Agent-Based Simulations of Subjective Well-Being, *Social Indicators Research* 115 (2) (2014) 623-635.

[74] S.Y. Chong, P. Tiño, D.C. Ku, X. Yao, Improving generalization performance in co-evolutionary learning, *IEEE Transactions on Evolutionary Computation* 16 (1) (2012) 70-85.

[75] J. Zschache, Producing public goods in networks: Some effects of social comparison and endogenous network change, *Social Networks* 34 (4) (2012) 539-548.

[76] W.D. Hamilton, Selfish and spiteful behaviour in an evolutionary model, *Nature* 228 (1970) 1218-1220.

[77] K.R. Foster, T. Wenseleers, F.L. Ratnieks, Spite: Hamilton's unproven theory, *Annales Zoologici Fennici* 38 (3-4) (2001) 229-238.

[78] N.J. Vriend, An illustration of the essential difference between individual and social learning, and its consequences for computational analyses, *Journal of economic dynamics and control* 24 (1) (2000) 1-19.

[79] J. Bergin, D. Bernhardt, Comparative Learning Dynamics, *International Economic Review* 45 (2) (2004) 431-465.

[80] S. Tanabe, N. Masuda, Evolution of cooperation facilitated by reinforcement learning with adaptive aspiration levels, *Journal of Theoretical Biology* 293 (0) (2012) 151-160.

[81] Y. Bereby-Meyer, A.E. Roth, The speed of learning in noisy games: partial reinforcement and the sustainability of cooperation, *The American Economic Review* 96 (4) (2006) 1029-1042.

[82] T. Jun, R. Sethi, Reciprocity in evolving social networks, *Journal of Evolutionary Economics* 19 (3) (2009) 379-396.

[83] S.M. Garcia, A. Tor, R. Gonzalez, Ranks and rivals: A theory of competition, *Personality and Social Psychology Bulletin* 32 (7) (2006) 970-982.

[84] J. Dunn, N.E. Ruedy, M.E. Schweitzer, It hurts both ways: How social comparisons harm affective and cognitive trust, *Organizational Behavior and Human Decision Processes* 117 (1) (2012) 2-14.

Process of the Evolutionary IPD Model

```

1: Initiate a population of  $N$  agents. Each agent  $i$  ( $i = 1, 2, \dots, N$ ) is specified by a risk attitude  $\alpha_{i0}$  and a strategy  $s_{i0}$ 
2: Set:  $t = 0$ 
3: While  $t \leq G$  do
4:   // Interaction
5:   For  $i = 1$  to  $N$  do
6:     Randomly select  $g_n$  opponents  $\{Opp_{ijt}, j = 1, 2, \dots, g_n\}$  from the population for Agent  $i$ 
7:     Agent  $i$  plays the IPD game with each opponent
8:     Calculate the average payoff  $P_{it}$  for Agent  $i$  // Equation (3)
9:   End For
10:  // Payoff-oriented Adaptation of Risk Attitude according to Prospect Theory
11:  For  $i = 1$  to  $N$  do
12:    Determine the aspiration goal  $G_{it}$  of Agent  $i$  // Equation (4)
13:    If  $P_{it} \geq (1 + \gamma) \times G_{it}$  then
14:       $\alpha_{i,t+1} \leftarrow \alpha_{it} \times (1 + r_{up}) + N(0, \alpha_{it} \times r_{up} \times v)$ 
15:    Else If  $P_{it} \leq (1 - \gamma) \times G_{it}$  then
16:       $\alpha_{i,t+1} \leftarrow \alpha_{it} \times (1 - r_{down}) + N(0, \alpha_{it} \times r_{down} \times v)$ 
17:    Else  $\alpha_{i,t+1} \leftarrow \alpha_{it}$ 
18:    End If
19:  End If
20: End For
21: // Utility-oriented Modification of Strategy based on Evolutionary Learning
22: Select the best  $N/2$  agents with the highest average payoffs in Interaction, and their strategies are denoted as
     $\{s_{1,best}, s_{2,best}, \dots, s_{N/2,best}\}$ 
23: For  $i = 1$  to  $N$  do
24:   Randomly select two strategies from  $\{s_{1,best}, s_{2,best}, \dots, s_{N/2,best}\}$  with probabilities proportional to the average
    payoffs of the corresponding agents
25:   A candidate strategy  $C_{it}$  is generated from the two chosen strategies by combination and mutation
26:   Calculate the utilities  $U(C_{it})$  and  $U(s_{it})$  based on the risk attitude  $\alpha_{i,t+1}$  // Equations (6) & (7)
27:   If  $U(C_{it}) \geq U(s_{it})$ 
28:     then  $s_{i,t+1} \leftarrow C_{it}$ 
29:   Else  $s_{i,t+1} \leftarrow s_{it}$ 
30:   End If
31: End For
32: Set:  $t = t + 1$ 
33: End While

```