Kindness Matters: A Theory of Reciprocity*

Yi Shi†

February 20, 2024

Abstract

A large body of evidence suggests that people are willing to sacrifice their own material payoffs to reward those who are kind to them or to hurt those who are unkind to them. In this paper, we present a theory of reciprocity, which consists of a two-sided underlying intentional kindness of other people (incorporating views about what payoffs other people can receive and what they should receive), and a one-sided consequential kindness of other people (in which the decision maker dislikes inequitable outcomes for themselves but does not care about inequitable outcomes for others). We also introduce a new definition of efficient strategy that successfully solves paradoxes in existing behavioural models. We further show that our model reflects the findings of a host of experiments on games such as the ultimatum game and the sequential prisoner's dilemma, which neither the standard theory nor other existing reciprocity models can explain. Finally, our model explains why the decision maker's positive or negative reciprocity falls when they have an opportunity to punish others.

Keywords: Reciprocity, Intention, Consequence, Psychological game theory

JEL classification: C79, D63, D64

*This paper was presented at Essex RSS seminar, and 18th Doctorissimes (2023). We are grateful to the participants at the seminar and conference for their helpful comments.

†Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom. e-mail: yi.shi@essex.ac.uk

1 Introduction

There is considerable laboratory and real-world evidence suggesting that people care about the behavioural intentions and consequences of others. They are willing to bear a personal cost to reward kind behaviour (known as positive reciprocity) and, conversely, to punish unkind behaviour (known as negative reciprocity). For example, tourists tip the tour guide after a pleasant journey even though they are unlikely to encounter the tour guide again, players often like to use the tit-for-tat strategy in the repeated prisoner's dilemma, and charities can increase the propensity of potential donors to donate when solicitation letters are accompanied by gifts (Falk, 2007). In the growing body of work on experimental games over the last three decades, reciprocity considerations have cast doubt on the classic assumptions of rationality and material self-interest[1]. All these observations raise two fundamental questions: what promotes positive reciprocal behaviour, and what promotes negative reciprocal behaviour?

The main purpose of this paper is to provide a comprehensive analysis to understand the nature of reciprocity. Despite a large body of theoretical literature on reciprocity, also known as other-regarding preferences, we are aware of no models that efficiently explain why people behave differently when encountering the same decision problem (see the examples in Figures 1 and 2). In our analysis, we revisit the central issue of when one's behaviour can be perceived as kind, especially when costly forms of punishment are available, and we develop the concept of expected reciprocity equilibrium (ERE). In our model, positive reciprocity is triggered by the player's intentions, and negative reciprocity is induced by the player's intentions as well as the consequences of their actions.

Taking into account the intentions of an action mean that the decision maker's beliefs about what another player might do (first-order belief) and about what another player believe they might do (second-order belief) are able to influence their perception of kindness. In psychology, and more recently in economics, researchers have contemplated preferences that depend on beliefs about other players' intentions. This suggests that these beliefs will enter directly into the utility, and further influence the preferences of the decision maker. It provides an alternative line of thinking that a player might use to explain why another player makes a particular choice, taking into account whether they only want to avoid a worse outcome for themselves, whether they want to help others, or both. But how are intentions captured? Two forms of intentions are included in our model:

1. What people can get: this depends on their own possible outcomes and partially determines others' kindness.

[1]See, for example, work on the dictator game (Andreoni and Bernheim, 2009), the ultimatum game (Güth, Schmittberger, and Schwarze, 1982; Thaler, 1988; Falk, Fehr, and Fischbacher, 2003; Falk and Fischbacher, 2006), the prisoner's dilemma (Clark and Sefton, 2001; Ahn et al., 2007; Dhaene and Bouckaert, 2010; Klempt, 2012; Charness, Rigotti, and Rustichini, 2016; Engel and Zhurakhovska, 2016; Orhun, 2018; Gächter, Lee, and Sefton, 2022), and the gift-exchange game (Fehr and Schmidt, 1999; Berg, Dickhaut, and McCabe, 1995).

2. What people should get: this depends on possible outcomes for other people and considers to what extent other people sacrifice their own payoffs to help or hurt them.

One contribution of this paper is that we take a novel approach to incorporating the status of other people (i.e. what they should get) to evaluate the underlying intentions. The empirically supported implications of this proposal are twofold. First, the possible material payoffs of both players resulting from all possible efficient strategies can influence the decision maker's perception of kindness and further influence their reciprocal behaviour. For example, in the ultimatum game, the responder may respond differently for the same offer proposed by the proposer when the proposer has different alternative offers available. Experiments show that the responder is more likely to accept the same offer when the alternative offer brings the proposer less material payoffs (Falk, Fehr, and Fischbacher, 2003). The second implication of our model is related to the power of punishment. Specifically, faced with kind behaviour, the decision maker is more willing to behave selfishly when they have the power to punish the unkind behaviour of other players at a low cost to themselves. For example, in the sequential prisoner's dilemma, the second mover is less likely to positively reciprocate (by choosing cooperation) the first mover's cooperation when the second mover has the chance to punish the first mover's defection compared with the general prisoner's dilemma, in which there is not a costly punishment option (Orhun, 2018). The two findings imply that the extent to which others help the decision maker will influence their preference.

Apart from the underlying intentions of an action, the consequences of an action[2] are another central part of our model. This refers to the final distributions resulting from all players' actions and explains why people respond negatively even if there does not exist any intention of other people: people care about what they may have obtained compared with others. They do not want to be treated unequally, and they experience disutility for gaining less than others.

The second main contribution of our model is that we take a new approach to the question of when people experience disutility from inequity. Our model suggests that the decision maker not only pays attention to other players' underlying intentions but also to their relative material payoff compared with others players. The decision maker prefers to choose an action that decreases the negative differences[3] between them and other players but will not experience disutility when they face a positive difference[4].

Two pioneering intention-based reciprocity models: Rabin (1993) [hereafter Rabin] and Dufwenberg and Kirchsteiger (2004) [hereafter D&K]) instead argue that the decision maker assesses the kindness of other people only on the basis of "what people can

[2]In many studies, some authors also prefer to call this "inequity-aversion". See Fehr and Schmidt (1999) or Bolton and Ockenfels (2000).

[3]The negative difference here means that the decision maker gains less than other players in material payoffs.

[4]The positive difference here means that the decision maker gains more than other players in material payoffs.

get"[5]. Therefore, contrary to the above two implications of our model, their models imply that the decision maker's perception of other players' kindness will not be influenced when the alternative brings the other player less or more material payoffs. Furthermore, their models also imply that the possibility of punishment increases the likelihood of positive reciprocity of the decision maker, which is refuted by experimental studies (Orhun, 2018). Theoretically, their models pay attention to "positive reciprocity" (one player's material payoff decreases and another player's material payoff increases) but ignore "negative reciprocity" (both players' material payoffs decrease). That is why the predictions of these models have been contradicted by a growing number of experimental studies (Falk, Fehr, and Fischbacher, 2003; Falk, Fehr, and Fischbacher, 2008).

In our model, we incorporate the status of other players (i.e. reflecting what people can get and what people should get) into our analysis by implementing a new reference standard, which is used for the evaluation of other players' intentional kindness[6]. Instead of simply applying half-half division, our analysis takes into consideration whether the other player intends to be kind or only to avoid a worse outcome for themselves. Our model yields testable predictions that can be compared with the results of several prior experiments (Falk, Fehr, and Fischbacher, 2003; Ahn et al., 2007; Orhun, 2018; Gächter, Lee, and Sefton, 2022).

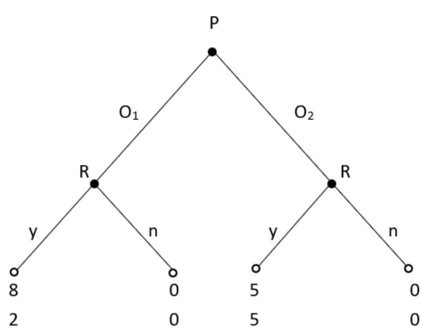

Figure 1: ultimatum game-(5,5)

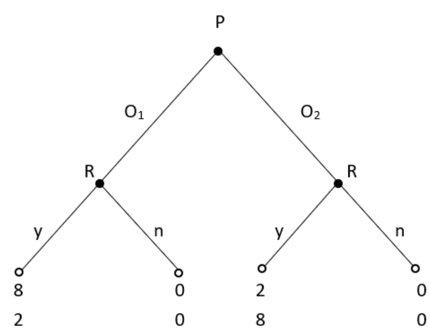

Figure 2: ultimatum game-(2,8)

Experiments with punishment are especially useful devices to measure subjects' psychological propensities. This can be best illustrated by ultimatum games, shown in Figures 1 and 2 (Falk, Fehr, and Fischbacher, 2003). Is "O_1" in ultimatum game-(2,8) more kind than "O_1" in ultimatum game-(5,5)? Experimental results indicate

[5]The rule is simply defined as a comparison between their actual material payoff and the equal division of the highest and lowest material payoffs they might obtain in the game. In our study, we further argue that the status of other players, i.e. "what people should get", will also influence the perception of the kindness of others from the perspective of the player.

[6]The reference point is one of the most important criteria in our model to determine whether a person is kind or not. If the actual material payoff of the decision maker is greater than this reference point, the other player is kind. Conversely, if the actual material payoff of the decision maker is smaller than this reference point, the other player is unkind.

that almost all responders (referred to as player R) accept (i.e. they choose "y") the (5,5) and (2,8) offer in Figure 1 and Figure 2, respectively. However, 44.4% of them reject (i.e. they choose "n") the (8,2) offer in Figure 1 and 26.7% of them reject the (8,2) offer in Figure 2. Intuitively, in Figure 1, the proposer (referred to as P) has the option to allocate a fair enough offer (5,5) or an offer (8,2) that will hurt the responder significantly; in Figure 2, the proposer still has an unfair offer (8,2) but this seems to be more acceptable than the same offer in Figure 1. This is because the proposer has an excuse for not choosing the (2,8) offer: the (2,8) offer will hurt the proposer significantly. The proposer might argue "I want to be a kind person, but helping you will hurt me too much, so I am forced to select the (8,2) offer". The experimental results agree with this intuition. This suggests that the statuses and outcomes of other players should also be taken into consideration. Our model serves this purpose. Instead of using the half-half weighting rule, our new reference standard puts a higher weight on "O_1" in ultimatum game-(2,8) than in ultimatum game-(5,5). Hence, even though "O_2" brings the responder more material payoffs in ultimatum game-(2,8) than in ultimatum game-(5,5), the reference point in Figure 2 is smaller than the reference point in Figure 1 in our model, which suggests the proposer is more kind by choosing "O_1" in Figure 2 than choosing "O_1" in Figure 1 and that more responders will reject the (8,2) offer in Figure 1 than in Figure 2. The prediction of our model is in line with the experimental results.

What if we apply existing intention-based models (Rabin and $D\&K$)? Unfortunately, the models yield a counterfactual answer, which suggests that the rejection rate of the (8,2) offer in Figure 2 is greater than in Figure 1. In addition to the intention-based reciprocity models, there are some other models that purely depend on the consequences of an action (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002). Since the utility of the decision maker depends only on the final payoff distributions, these models are also unable to explain the experimental results in Falk, Fehr, and Fischbacher (2003) as "O_1" would provide the same options for the responder in both Figure 1 and Figure 2. Thus, they predict the same rejection rate for the (8,2) offer in Figure 1 and Figure 2. This is still a contradiction. In addition to this experiment, in Section 4 we discuss more experimental games that existing models fail to explain, but in which our model shows good agreement with experimental results.

What makes a behavioural strategy convincing? Our third contribution is that we propose a method to eliminate the influence of inefficient strategies. We wish to capture to what extent another player is kind to the decision maker. Consider now a different aspect with regard to underlying intentions, namely wasteful strategies, which are strategies that no one is motivated to choose as they always lead to worse outcomes for both players. These strategies should not be seen as the preferred option of another player from the perspective of the decision maker. This causes another major challenge in the evaluation of intentions: how do we eliminate the effect of wasteful strategies and keep so-called efficient strategies? In the existing literature, there are three approaches to solving this problem (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Dufwenberg and Kirchsteiger, 2019), but they all contradict some experimental results (Dufwenberg and Kirchsteiger, 2019; Isoni and Sugden, 2019), which we discuss in detail in Section 5. In our model, we propose the "potential worst outcome" as a sequential rationality refinement to successfully find all efficient strategies. We confirm that our definition of efficient strategy can avoid the contradictions in previous literature and address problems in experimental studies.

To summarize, the objective of this paper is to study intention- and consequence-based reciprocity by investigating: (I) intentions: (1) many experimental studies contradict previous intention-based models, especially in experimental games where costly punishment is an option. How do we correctly capture players' intentions and successfully predict their behaviour? and (2) how to rule out wasteful strategies since no one is motivated to play in kindness consideration? (II) consequence: if there does not exist any intention, why do people still would like to punish others?

The rest of the paper is organized as follows: Section 2 reviews the literature; in Section 3, we present our model that incorporates consequential kindness and intentional kindness into the decision maker's preferences. We also develop the concept of expected reciprocity equilibrium (ERE) and prove its existence. In Section 4, we apply our model to some experimental games: the ultimatum game, the sequential prisoner's dilemma with punishment, and the prisoner's dilemma with asymmetric payoffs. The results show that our model is in line with the experimental results. Section 5 discusses the contribution of our model compared with the Rabin, D&K, and F&F models, such as the efficient strategy and the influence on equilibria. Section 6 concludes.

2 Literature Review

With respect to how reciprocity is defined, two main classes of models can be distinguished: consequence-based models and intention-based models. The consequence-based models (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002) claim that reciprocity or fairness refers to the final distribution of material payoffs. One decision maker's behaviour is driven not only by their own material payoff, but also by how other players gain compared with their payoff. For example, Fehr and Schmidt (1999) assume that an agent feels envy if other players' material payoffs exceed their own material payoff and feels unequal when other players' material payoffs are lower than their own material payoff. Furthermore, the agent experiences more inequity if another player gains more than them than they experience if they gain more than another player; the model of Bolton and Ockenfels (2000) assumes that the decision maker cares about their own relative status, and their utility depends on their own material payoff relative to the average overall payoff. This class of models capture the foremost determinants of other-regarding behaviour.

A similar model is proposed by Falk and Fischbacher (2006). Although their model includes the beliefs and intentions, the perception of kindness is determined by the expected material payoffs: "player i believes that player j aims to let them get more out of the exchange than player j wants for themselves", then player i perceives player

j as kind. They also consider the intention factor but use a set of given values to influence how one player wants to achieve the kind or unkind behaviour.

Another class of reciprocity models, which are built on the framework of psychological games (Geanakoplos, Pearce, and Stacchetti, 1989; Battigalli and Dufwenberg, 2009), focus on the underlying intentions to influence one's preference (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006). In a normal-form construct, Rabin (1993) uses the decision maker's beliefs about other players' actions, as well as the decision maker's beliefs about other player's beliefs about their actions, to assess the kindness of other players. Researchers observe surprising effects of the move order in many empirical studies(Cooper et al., 1993; Camerer, Knez, and Weber, 1996; Camerer, 1997; Dhaene and Bouckaert, 2010) where the results we observe are completely different from the simultaneous move games. This sheds light on the importance of sequential moves when we study reciprocity. Dufwenberg and Kirchsteiger (2004) address it in their sequential reciprocity model, which extends Rabin's model to sequential games. They make three changes compared with Rabin (1993), and hence propose a different influence on the decision maker's preference. One controversial change is their definition of efficient strategy, which is independent of the players' beliefs, while Rabin's definition is belief-dependent. The limitation of Rabin's definition has been discussed in Dufwenberg and Kirchsteiger (2004). Regarding to their belief-independent definition, Isoni and Sugden (2019) claim that their definition leads to a paradox in trust games. Although Dufwenberg and Kirchsteiger (2019) response critique by a new definition, it still remains a paradox. Another controversial change is related to the belief updating rules, in which one player's evaluation of their opponents' kindness is decided by the most updated beliefs at their decision nodes. Jiang and Wu (2019) argue that the D&K belief updating rule is not appropriate for games with more than two players; they propose another belief updating rule by categorizing players' beliefs via whether perceived kindness is calculated using their most updated forms.

Experimental literature provides a great deal of evidence about the effect of reciprocity (Dawes and Thaler, 1988; Falk, 2007). One typical experiment is the ultimatum game where the proposer makes an offer and the responder decides to accept or reject it (Güth, Schmittberger, and Schwarze, 1982; Thaler, 1988; Blount, 1995). Contrary to pure selfishness, experimental results suggest that the proposer may make a generous offer and the responder is likely to reject a positive but unequal offer. The responder would say "I would rather give up some material payoffs than accept an unfair offer". Falk, Fehr, and Fischbacher (2003) explore the nature of fairness by proposing four mini-ultimatum games. They find that people evaluate fairness not only by the consequences of an action but also by the intention of the player. A similar result is also observed in the moonlighting game Falk (2007), where player A first chooses a number a, denoting the payoff. A gives B a if $a \geq 0$, and A takes a away from B if $a < 0$. In the case of $a \geq 0$, the experimenter triples a so that B receives $3a$. After player B observes a, they can choose a reward or a sanction. Another famous experimental game is the prisoner's dilemma. For example, Clark and Sefton (2001) and Dhaene and Bouckaert (2010) show that reciprocity can influence an individual's behaviour; Ahn et al. (2007)

find that players with an advantage are less likely to cooperate to reward their opponent; Gächter, Lee, and Sefton (2022) find that the way cooperative behaviour varies with payoffs is consistent with the predictions of consequence-based reciprocity models; and Orhun (2018) finds that existing reciprocity models cannot explain subjects' behaviour when the second mover has the chance to punish defective behaviour.

When we mention reciprocity, the main question in our mind is related to the construct of kindness. In addition to the set of models we have mentioned. Çelen, Schotter, and Blanco (2017) propose a new definition of kindness with the notion of blame: suppose I were another player and wonder if I would choose an action that is more or less kind than this player's choice. I would blame the opponent if the latter is true, otherwise I would blame myself. Dufwenberg, Smith, and Van Essen (2013) consider vengeance and develop the concept of vengeance equilibrium to reflect negative reciprocity; Sohn and Wu (2022) extend the D&K model with the perspective of uncertainty, explore the threshold of cooperation, and propose extended sequential reciprocity equilibrium; and Battigalli and Dufwenberg (2022) summarize the existing models with belief-dependent motivations and psychological game theory.

3 The Model

3.1 Baseline framework

Our discussion in this paper will be confined to two-player, two-stage, extensive-form games with finitely many (usually two) actions at each stage and complete and perfect information. Therefore, the player's choice appears sequentially and fully observed. The purpose of our model is to provide a rigorous discussion of how we capture players' intentions and how to evaluate their perception of kindness.

Formally, let $N = \{1, 2\}$ be the set of players, H_i be the set of nodes of player $i \in N$ (or histories), A_i be the set of behavioural strategies of player i, and each strategy assigns at each node/history $h \in H$ a probability distribution over the set of possible actions of player i at h. With $a_i \in A_i$, $h \in H$, $a_{i,h}$ denotes the strategy that prescribes the same choice as a_i, except for the choice that decides history h that is made with probability 1. The material payoff of player i is given by $\pi_i : A \to \mathbb{R}$, for example in Figure 1, if the proposer chose "O_1", then $\pi_R(a_R, O_1) = \pi_R(y, O_1)Pr(y|O_1) + \pi_R(n, O_1)Pr(n|O_1)$. The material payoff means cash or some other measurable quantity (e.g. the number of vouchers), which denotes the selfish payoffs that we generally use in the game theory.

Moreover, our discussion in this paper includes the utility from intentions which reflect a player's psychological consideration. Like other intention-based reciprocity models (Rabin; D&K), therefore, we apply the framework of psychological game theory (Geanakoplos, Pearce, and Stacchetti, 1989; Battigalli and Dufwenberg, 2009). That is, when playing the game, player i's beliefs about others' strategies (first-order belief) and about others' beliefs about their own strategy (second-order belief) are important as they can influence the player's inference about others' intentions. Furthermore, in sequential games, the choice that a player has made can be fully observed by their opponents at each stage. Hence, the updating rule is also necessary. For the purposes of the following analysis, we define two types of beliefs (first-order belief and second-order belief) and updating rules[7].

Definition 1 (Beliefs and updating rules). *Let $B_{ij} \in A_j$ be player i's first-order belief and $B_{ij}(h)$ be the updated first-order belief that describes player j's actual behavioural strategy that leads to history h. $C_{iji} \in A_i$ denotes player i's second-order belief and $C_{iji}(h)$ denotes the updated second-order belief that describes player j's actual behavioural strategy that leads to history h.*

Example. Ultimatum game-(5,5) in Figure 1. Definition 1 describes that after observing the actions that their opponents have made and what they have done, players will update their first-order belief and second-order belief. In Figure 1, assume that the responder initially forms the first-order belief B_{RP} with $B_{RP}(O_1) = 0.3$ and $B_{RP}(O_2) = 0.7$. Once the proposer has made their choice: assume "O_1", then the responder updates their first-order belief with $B_{RP}(O_1) = 1$. In the same fashion, suppose that the proposer initially forms the first-order belief B_{PR} with $B_{PR}(y, O_1) = 0.4$, $B_{PR}(n, O_1) = 0.6$, $B_{PR}(y, O_2) = 0.8$, and $B_{PR}(n, O_2) = 0.2$. Once the responder has made their choice, assume "n" after "O_1" and "y" after "O_2", then the proposer updates their first-order belief with $B_{PR}(n, O_1) = 1$ and $B_{PR}(y, O_2) = 1$.

Suppose the responder forms the initial second-order belief C_{RPR} with $C_{RPR}(y, O_1) = 0.4$ and $C_{RPR}(n, O_1) = 0.6$. Once the responder has made their choice, assume "y" after "O_1", then they update their second-order belief with $C_{PRP}(y, O_1) = 1$. Similarly, suppose the proposer initially forms the second-order belief C_{PRP} with $C_{PRP}(O_1) = 0.4$ and $C_{PRP}(O_2) = 0.6$. Once the proposer has made their choice, assume "O_1", then they update they second-order belief with $C_{PRP}(O_1) = 1$.

3.2 Efficient strategy

Generalized reciprocity suggests we would like to reward those who give us more and punish those who give us less. This seems to be the best way to reflect their intentional kindness and unkindness. However, what if the only reason that other people give us more is that they can also benefit from this action (i.e. a mutually beneficial action)? How do we identify the purpose of their action and define their kindness? One solution is that before evaluating players' kindness, we first rule out all wasteful strategies that no one is motivated to play (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004). An appropriate definition of efficient strategy is able to rule out all such strategies, and addressing this concern. As discussed in Section 1, existing definitions of efficient strat-

[7]Some scholars, such as Lianjie Jiang et al. (2018), argue that the D&K definition of updating rules causes a series of problems, but these drawbacks are of no relevance to the two-stage games studied in this work.

egy cause a range of problems (this is further discussed in Section 5). Especially for so-called "punishment without cost":

Punishment without cost. *Assume the material payoff of player i is $\pi_i(a_{i,h}, a_{j,h})$ and of player j is $\pi_j(a_{i,h}, a_{j,h})$. Then if $\pi_i(a_{i,h}, a_{j,h}) \geq \pi_i(a'_{i,h}, a_{j,h})$ and $\pi_j(a_{i,h}, a_{j,h}) < \pi_j(a'_{i,h}, a_{j,h})$, then compared with $a'_{i,h}$, the strategy $a_{i,h}$ is called the behavioural strategy of punishment without cost.*

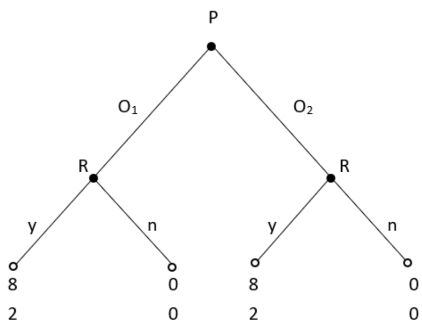

Figure 3: ultimatum game-(8,2)

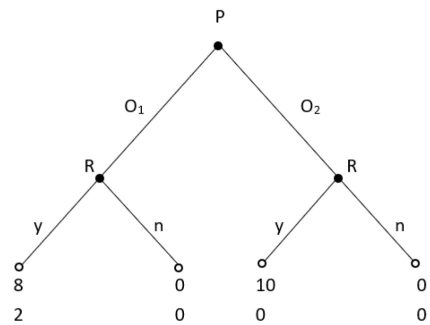

Figure 4: ultimatum game-(10,0)

This can be best illustrated by another two mini-ultimatum games studied by Falk et al. (2003), shown in Figures 3 and 4. Is "O_1" in ultimatum game-(10,0) the same as "O_1" in ultimatum game-(8,2) from the perspective of the responder? If yes, what is the implication? In Figure 3, "O_1" and "O_2" the are same for the proposer and the responder as (8,2) is the only offer. Clearly, we are not able to read any intention from the proposer's action. In Figure 4, although the proposer has two options, the (8,2) offer and the (10,0) offer, the experimental results obtained by Falk et al. (2003) indicate that there is no difference between ultimatum game-(8,2) and ultimatum game-(10,0) for the responder's response of "O_1". This implies that the (10,0) offer does not influence the responder's inference about the proposer's intentions. It seems that having the alternative offer (10,0) cannot make the (8,2) offer more generous from the perspective of the responder. Why does the responder behave this way?

To understand this phenomenon, let us try to infer the responder's consideration in Figure 4. On the one hand, the responder will receive zero material payoff regardless of whether they accept or reject the offer if the proposer chooses the (10,0) offer ("O_2"). On the other hand, the responder has the chance to receive a material payoff of 2 if they accept the (8,2) offer. Therefore, the (10,0) offer can never be a kind behaviour. Now, given the chance to punish this unkind behaviour without cost (the material payoffs of playing "y" and playing "n" after "O_2" are both zero), the responder should choose "n" since no one is motivated to reward an unkind behaviour. Correspondingly, still in Figure 4, if the proposer believes that the responder never rewards (by choosing "y") the unkind (10,0) offer, and the responder also believes that the proposer believes the responder never rewards an unkind behaviour (means "O_2" must lead "n"), then the proposer should never select "O_2" from the perspective of the responder since this would lead both players to the worst material payoff pair (0,0). Hence, "O_1" in ultimatum game-(10,0) will be the only choice for the proposer from the perspective of the responder. We can answer now that "O_1" in ultimatum game-(10,0) is the same as "O_1" in ultimatum game-(8,2) from the responder's perspective. In other words, "O_2" should be considered a "wasteful strategy" that should not be taken into consideration in Figure 4 when the responder considers the underlying intentions from the proposer.

It is common to find similar situations in other games and settings. It is necessary to rule out such wasteful strategies before the evaluation of intentional kindness. This is also what the Rabin and D&K models aim to achieve. Unfortunately, these approaches do not successfully rule out all wasteful strategies. Rabin's method leads to contradictions in sequential games (see Section 5 in D&K for details). Although the D&K method of ruling out efficient strategies works for some games, it still fails to explain many games (see Section 5 for further discussion), for example in Figure 4, their model suggests that both "O_1" and "O_2" are efficient[8]. Therefore, we propose the idea of "potential worst outcome" (PWO) to address such issues.

PWO as a sequential rationality refinement. To rule out all wasteful strategies, the idea of sequential rationality seems to be feasible. We can define a special second-order belief that assigns the probability 1 to a strategy that satisfies sequential rationality. Does it really work? For some situations, it seems to provide us with a convincing explanation. But we notice that we might meet the "punishment without cost" case like we have described in Figure 4. Under such situations, we are not able to assign both actions with probability 1. Based on our inference and discussion, we propose a refinement of sequential rationality by adding the idea of "potential worst outcome" to define a special second-order belief to rule out all wasteful strategies. The word "potential" refers to the sequential rationality and "worst outcome" means that when one has the chance to freely punish others, they would like to apply this chance since no one is motivated to reward an unkind person. At this point, we adopt three steps to meet our purpose: in step (i), we find the most advantageous strategy that should be unique for player i (by sequential rationality); in step (ii), if a_i does not satisfy (i) because a_i might be not unique, then player i may have the chance to punish their opponents without cost. Here we suppose that the player i would punish their opponent to satisfy our definition of "worst outcome"; and in step (iii), if a_i does not satisfy (ii) because a_i is still not unique, then there exist multiple strategies that bring both players the same material payoffs. The payoffs of players are indifferent between these strategies, thus i chooses randomly.

Definition 2 (Special second-order belief). *For any $a_i \in A_i$, $a_j \in A_j$, and $h \in H$, let*

[8] If both "O_1" and "O_2" are efficient in ultimatum game (10,0), then "O_1" must be more kind in Figure 4 than in Figure 3. This is a contradiction.

$\hat{A}_{i,h}(a_{j,h}) \equiv \arg\max_{a_{i,h} \in A_{i,h}} \pi_i(a_{i,h}, a_{j,h})$ *and define* $C^{pwo}_{iji}(a_{j,h}) \subseteq \hat{A}_{i,h}(a_{j,h})$ *as follows:*

$$a_{i,h} \in C^{pwo}_{iji}(a_{j,h}) \begin{cases} \textit{if either (i) } \pi_i(a_{i,h}, a_{j,h}) > \pi_i(a'_{i,h}, a_{j,h}) \; \forall \; a'_{i,h} \in A_{i,h}(a_{j,h}) - \{a_{i,h}\} \\ \textit{or if (ii) } \pi_j(a_{i,h}, a_{j,h}) < \pi_j(a'_{i,h}, a_{j,h}) \; \forall a'_{i,h} \in \hat{A}_{i,h}(a_{j,h}) - \{a_{i,h}\} \\ \textit{or if (iii) } \pi_k(a_{i,h}, a_{j,h}) = \pi_k(a'_{i,h}, a_{j,h}) \; \forall \; k \in \{i,j\} \; \forall a'_{i,h} \in \hat{A}_{i,h}(a_{j,h}). \end{cases}$$

Example. Ultimatum game-(8,2) and ultimatum game-(10,0) in Figure 3 and Figure 4, respectively. Assume player i is the responder and player j is the proposer. In Figure 3, to find $C^{pwo}_{RPR}(a_{P,h})$, $a_{P,h}$ might be "O_1" or "O_2": when $a_{P,h}$ is "O_1", because $\pi_R(y, O_1) > \pi_R(n, O_1)$, then action "y" satisfies definition 2. Therefore, $C^{pwo}_{RPR}(y, O_1)$= 1. Similarly, when $a_{P,h}$ is "O_2", because $\pi_R(y, O_2) > \pi_R(n, O_2)$, then action "y" satisfies Definition 2. Therefore, $C^{pwo}_{RPR}(y, O_2)$= 1.

Now we turn to Figure 4 and see if anything is different. In this game, similarly, $a_{P,h}$ still might be "O_1" or "O_2": when $a_{P,h}$ is "O_1", because $\pi_R(y, O_1) > \pi_R(n, O_1)$, then action "y" satisfies definition 2. Therefore, $C^{pwo}_{RPR}(y, O_1)$= 1. However, if $a_{P,h}$ is "O_2", we notice that $\pi_R(y, O_2) = \pi_R(n, O_2)$. Then we need to consider Definition 2(ii)[9]. Since $\pi_P(n, O_2) < \pi_P(y, O_2)$, then action "n" satisfies Definition 2(ii). Therefore, we get $C^{pwo}_{PRP}(n, O_2)$=1.

So far, we have defined a special second-order belief by the "potential worst outcome" to explain the player's intention appropriately. Here the "potential worst outcome" is not the "absolute worst outcome" for any history and any strategy. It describes when player i has the chance to punish player j without cost, player i would like to do so. But why not apply the absolute worst outcome to define the second-order belief in which player i would directly choose the action that will cause player j to receive the worst outcome? On the one hand, although player i can give player j the worst outcome, player i might lose some part of their own material payoff (so-called costly punishment). We cannot argue that they should definitely cede their own material payoff to punish player j[10]. On the other hand, the objective of defining an efficient strategy set is to have a reference point in the following studies such that the kindness/intention will not be underestimated or overestimated. Indeed, player i might choose the costly punishment to punish the unkind behaviour, but they may avoid punishing others when the cost is too high to bear. Therefore, the idea of the potential worst outcome is the appropriate one.

Definition 2 provides us with a way to rule out wasteful strategies. Furthermore, we apply this special second-order belief $C^{pwo}_{iji}(a_{j,h})$ for player i to define the efficient strategy for player j.

[9]This is the difference compared with the sequential rationality. If we apply the idea of the sequential rationality, then the responder has no preference between "y" and "n" when the proposer plays "O_2". However, we want to evaluate the player's intention, so our definition 2(ii) serves the purpose.

[10]Of course, costly punishment is possible, and it is also an important part of our research. We will solve this problem in Definition 5. Otherwise, the model is for those who are anti-social.

Definition 3 (Efficient strategy). *Define an efficient strategy set for $j \in \{1,2\}$ as follows:*

$$E_j^{pwo} = \left\{ a_j \in A_j \;\middle|\; \begin{array}{l} \textit{if } \nexists\, a'_j \in A_j \textit{ such that:} \\ (i)\ \pi_k(a'_j, C_{iji}^{pwo}(a_{j,h})) \geq \pi_k(a_j, C_{iji}^{pwo}(a_{j,h}))\ \forall k \in \{i,j\} \textit{ and} \\ (ii)\ \pi_k(a'_j, C_{iji}^{pwo}(a_{j,h})) > \pi_k(a_j, C_{iji}^{pwo}(a_{j,h}))\ \textit{ for some } k \in \{i,j\} \end{array} \right\}.$$

In other words, Definition 3 excludes a strategy a_j if and only if there exists at least one a'_j which describes the choice that leads to Pareto-superior outcomes given the special second-order belief $C_{iji}^{pwo}(a_{j,h})$.

Example. We apply this definition to Ultimatum game-(8,2) and Ultimatum game-(10,0), illustrated in Figures 3 and 4, respectively. In Figure 3, we obtain $C_{RPR}^{pwo}(y, O_1)$= 1 and $C_{PRP}^{pwo}(y, O_2)$=1. From the perspective of the responder, according to Definition 3, both "O_1" and "O_2" are efficient strategies since $(8,2) = (8,2)$. However, in Figure 4, we obtain $C_{RPR}^{pwo}(y, O_1)$=1 and $C_{PRP}^{pwo}(n, O_2)$=1 by Definition 2. From the perspective of the responder, according to the Definition 3, only "O_1" is efficient now since $(0,0) < (8,2)$. This result is consistent with our inference and the experimental results.

In addition to the games discussed above, our definition successfully addresses many other potential problems encountered with existing definitions. We discuss this in detail in Section 5.

3.3 Kindness

Kindness consideration is the most important element in our model. Summing up the last two sections, the perceived kindness should include both (i) the player's intentions and (ii) relative consequences. Our study tries to incorporate both (i) and (ii). According to our model, the utility consists of three parts: the material payoff, **intentional kindness** (**two sides**) $\mathbf{\Psi_i}$ and **consequential kindness** $\mathbf{\Phi_i}$. In the next two subsections, we will introduce the last of those two elements in detail.

3.3.1 The intentional kindness

Having defined the efficient strategy, we now move to one important part of kindness - the intention of an action. The intention has become a popular approach to explaining reciprocal behaviour in the last three decades (Rabin; D&K; F&F). Their idea of underlying intentions (one side) seems to successfully explain some non-selfish behaviour, but retains many potential drawbacks, as discussed in Section 1. So we interpret the intentional kindness again and propose our new idea of the underlying intentions.

What is the intentional kindness? (our view) *I never expect that you sacrifice your own material payoffs in helping (hurting) me, but once you do it, I might perceive you as a(n) kind (unkind) person, and I would like to help (punish) you as well if this* *will not hurt me significantly.*

Our view of intentional kindness consists of two sides. To understand it, take the ultimatum games in Figure 1 and Figure 2 as an example. Assume that the proposer chose "O_1" in two games. Then the responder might reject the (8,2) offer if they think they are treated unkindly. This raises a basic question as to what elements influence their inference of being treated kindly or unkindly. On the one hand, apart from the possibility that they can receive 2 or 0 (the proposer chose "O_1"), the responder might get 5 in Figure 1 and 8 in Figure 2 (if the proposer chose "O_2"). Such consideration can partially influence the responder's inference of the proposer's kindness. However, it should not be the only influence. Otherwise, the responder in Figure 2 will feel worse than in Figure 1, as one might get 8 but have not achieved and one can get 5 but have not achieved.[11] On the other hand, the responder also considers why the proposer decides "O_1" instead of "O_2". From the perspective of the responder, the proposer might receive 8 if they play "O_1", but 5 if they play "O_2" in Figure 1 and only 2 if they play "O_2" in Figure 2. At this point, the proposer choosing "O_1" in Figure 2 might be more acceptable than in Figure 1. This aspect can also partially determine the responder's perception of the proposer's kindness. Therefore, when modelling intentional kindness, we should consider both two sides and then make a trade-off.

Definition 4 (Reference point). *Let the reference point of player i with the consideration of player j's intention function $\vartheta(a_{j,h})$ be*

$$\pi_i^r = \textstyle\sum_{a_{j,h}\in E_j^{pwo}} \vartheta(a_{j,h}) \cdot \pi_i(C_{iji}, a_{j,h}).$$

To interpret this expression, as we argued above, the evaluation of kindness depends on two aspects: what people can get (determined by player i's own material payoffs) and what people should get (determined by the material payoffs of player j). Our definition of reference point suggests such considerations where $\pi_i(C_{iji}, a_{j,h})$ measures what player i can get and $\vartheta(a_{j,h})$ measures what player i should get and to what extent player j sacrifices their own material payoff in helping player i. It is convincing that player j is more likely to choose the action that can bring themselves a higher material payoff from the perspective of player i. This is in line with empirical studies and real-world observations (Charness, Rigotti, and Rustichini, 2016; Orhun, 2018).

Intention function $\vartheta(\mathbf{a_{j,h}})$. For our purpose, the logistic quantal response function (McKelvey et al., 1995) is the appropriate one: $\vartheta(a_{j,h}) = \frac{\exp[\lambda\cdot\pi_j(C_{iji},a_{j,h})]}{\sum_{\hat{a}_{j,h}\in E_j^{pwo}}\exp[\lambda\cdot\pi_j(C_{iji},\hat{a}_{j,h})]}$ [12]. Generally, $\vartheta(a_{j,h})$ has the following four properties: (i) $\vartheta(a_{j,h})$ is non-decreasing in

[11] This is also one reason why all existing intention-based reciprocity models fail to predict many experimental results (Falk et al., 2003 & 2008; Orhun, 2018).

[12] This specific function will be used in the remaining examples with $\lambda = 1$.

$\pi_j(a_{j,h})$, $\frac{\partial\vartheta(a_{j,h})}{\partial\pi_j(a_{j,h})} \geq 0$ where $a_{j,h} \in E_j^{pwo}$, and non-increasing in $\pi_j(\tilde{a}_{j,h})$, $\frac{\partial\vartheta(a_{j,h})}{\partial\pi_j(\tilde{a}_{j,h})} \leq 0$ where $\tilde{a}_{j,h} \in E_j^{pwo} - \{a_{j,h}\}$; (ii) if we have $\pi_j(a_{j,h})$ ¿ $\pi_j(\tilde{a}_{j,h})$, then we must have $\vartheta(a_{j,h}) \geq \vartheta(\tilde{a}_{j,h})$; (iii) we must have $\vartheta(a_{j,h}) > 0$ $\forall a_{j,h} \in E_j^{pwo}$; and (iv) we have $\sum_{a_{j,h} \in E_j^{pwo}} \vartheta(a_{j,h}) = 1$. The first two properties guarantee that others would like to choose the action that brings them a higher material payoff; the last two properties ensure that the reference point is located between the lowest and highest material payoff that player i might receive.

To explain why we introduce the intention function, we make the following two points. (1) The drawback of the existing definition of the reference point is that using "equitable payoff" as a simple principle to determine kindness means that people do not consider their decisions deeply and analytically (Messick, 1995)). They simply take a quick read on a situation and then make their decision. However, In most experimental games and real-world observations, this heuristic processing produces incoherent results, especially when there is an option for costly punishment (e.g. prisoner's dilemma with punishment, ultimatum games). Moreover, if we have more than two actions, "equitable payoff" might be pointless as it is very easy to reject kind behaviour when the highest material payoff is too large and easy to accept behaviour that is not quite kind when the lowest material payoff is very small. (2) One benefit of our definition of reference point is its universality: the existing method of "equitable payoff" can be part of our definition when players only care about their own material payoffs and have a quick thought on the kindness. If not, our definition can still explain players' behaviour very well. Another benefit is its reasonableness: our reference point takes into consideration a player's material payoff as well as that of their opponent. It is convincing in explaining the player's psychological concerns. Besides, applying our definition in Section 4, the efficiency is approved by the experimental results where the "equitable payoff" fails to predict such empirical results. Having defined the reference point, we now turn to the intentional kindness term.

Definition 5 (Intention part). *Player i's belief about how intentionally kind player $j \neq i$ is to i at history $h \in H$ is given by the following function:*

$$\delta_{ji} = \pi_i(a_i, B_{ij}(h), C_{iji}(h)) - \pi_i^r(h),$$

where $\pi_i(a_i, B_{ij}(h), C_{iji}(h))$ is the material payoff that player i received according to player j's behaviour $B_{ij}(h)$ and player i's response a_i.

Having $\pi_i^r(h)$ as a condition for player i's the "expected payoff", $\delta_{ji} > 0$ suggests that player i perceives player j as intentionally kind, $\delta_{ji} < 0$ suggests that player i perceives player j as intentionally unkind, and $\delta_{ji} = 0$ suggests that player i perceives player j as neither intentionally kind nor intentionally unkind.

Having introduced Definitions 4 and 5, we can specify player i's utility from the intentional kindness intention term. We use Ψ_i to denote the intentional kindness.

Definition 6 (Intentional kindness). *Player i's utility from the intentional kindness at history $h \in H$ is defined by:*

$$\Psi_i = \beta_i \cdot \delta_{ji} \cdot \pi_j(a_i, B_{ij}(h), C_{iji}(h)),$$

where β_i is an exogenously given non-negative number, which measures how sensitive player i is to the intention concerns with respect to player j.

A value of β_i=0 indicates that player i does not care about whether player j intends to help or hurt them, while β_i>0 means that player i cares about the underlying intentions of player j and prefers to give player j more material payoff for kind behaviour and less material payoff for unkind behaviour. The larger β_i is, the more player i cares about the underlying intentions.

3.3.2 The consequential kindness

Having the intentional kindness term Ψ_i, we find that for most games (e.g. the sequential prisoner's dilemma with/without punishment and the ultimatum games with different alternatives), it provides us with a solid explanation of the player's behaviour. However, there still exist some shortcomings. For example, the experimental results in Figure 3 (Falk, Fehr, and Fischbacher, 2003) suggest that if we only consider the intention term, then the responder should always accept the offer since there does not exist any intention for the proposer as the proposer is forced to propose the (8,2) offer. In reality, this is not the case: some responders still reject the offer in the experiment. Therefore, the intentional kindness term Ψ alone fails to explain players' reciprocal behaviour.

A growing body of experimental and theoretical evidence demonstrates the importance of relative gains between players (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Falk, Fehr, and Fischbacher, 2003), especially for those who are inequity-averse. Based on this evidence, some economists propose utility models that incorporate other-regarding preferences. The most well-known is the model suggested by (Fehr and Schmidt, 1999) who argue that individual behaviour is driven not only by selfishness but also by concerns for others' well-being. In their model, the utility of one player comprises three parts: their own material payoff, the disutility for gaining less than others, and the disutility for gaining more than others. For the experiment of (Falk, Fehr, and Fischbacher, 2003), it is easy to notice that there does exist a "disutility for gaining less than others", and hence some responders choose a "fair" enough result "n" even though (8,2) is the only offer for the proposer. But what about "the disutility for gaining more than others" - does the decision maker truly want to refuse the only offer by which they can earn more than others? For example, if the (2,8) offer is the only offer for the proposer, will the responder choose "n"? In other words, we want to know whether the decision maker cares about the inequity only for themselves or inequity for others too.

The best experimental game to capture whether the decision maker also cares about inequity of others should be the dictator game, which was first proposed by Kahneman, Knetsch, and Thaler (1986)[13]. Many economists also run the similar dictator games and find that people are not selfish and would like to help others. But their experimental results face many challenges such as the influence of reputation and social image because of anonymity. As a result, in recent decades, for example, Andreoni and Bernheim (2009) has suggested that people only like to be perceived as fair, instead of sincerely caring about others' well-being. Without social image concerns, subjects are more likely to select the action that benefits themselves, and the 50-50 norm will not hold anymore[14]. Another similar experiment is proposed by Franzen and Pointner (2012) who apply the randomized response technique that was originally designed by Warner (1965) to achieve anonymity. They find that some dictators choose to assign some points to the recipient in the general dictator game (but few dictators assign more than half ot the points to others)[15], but most dictators choose keeping all amounts in the randomized response technique dictators. Other studies, such as Bardsley (2008) and List (2007) find that dictators even withdraw resources from recipients when they have the option to do so in the "taking game"[16]. Further evidence is provided by Dana, Weber, and Kuang (2007), whose experiment indicates that sharing is vastly reduced if the recipient's payoff is not only determined by the dictator but also by chance. These results suggest that dictators behave non-selfishly only because they want to be seen as a fair person in the eyes of the recipients or even the experimenters. The implication from these experiments is that people care about inequity only for themselves and not for others.

Therefore, in our model, we have the following empirically supported assumptions. First, in addition to the purely selfish players, there exist players who dislike unequal outcomes and are inequity-averse (even if this is not their opponents' purpose or even if their opponents make the best choice for them). They think they are experiencing consequential unkindness if they are worse off in material payoffs than their opponents. This assumption is in line with the description of "the disutility for gaining less than

[13]In the experiment, subjects are given a choice of either an equal split of 20 dollars (10 dollars each) with another student or an unequal split (18,2). Three-quarters of the participants selected the equal split.

[14]In their experiment on the dictator game, they set a probability $1-p$ that the dictator chooses the transfer, and probability p that the nature sets the transfer equal to some fixed value. By increasing the value of p, in the experiment, more participants choose the nature transfer instead of the half-half offer. But a drawback of their experiment is that it has no option that is bad for the dictator: the two nature selected offers are both beneficial to the dictator (the dictator can gain more than half of the payoffs).

[15]This suggests that the decision maker does not want to be treated unequally, and does not want others' material payoffs over their own material payoffs even with the consideration of social image and reputation.

[16]In this experiment, both players are allocated 5 dollars. Participants who are dictators are allocated an additional 5 dollars. In a one-shot allocation game, the dictator can allocate from 0 to 5 of this additional endowment to their randomly assigned partner. The "taking game" allows the dictator to allocate from -5 to 5, which suggests the dictator can take some money from their paired partner.

others" in Fehr and Schmidt (1999). Moreover, in their model setting, the decision maker also experiences "the disutility for gaining more than others". However, according to empirical evidence, without social image or reputation concerns, the decision maker will not experience this. Therefore, in our model, we further assume that the decision maker experiences disutility if and only if their material payoffs are less than those of other players. This means that the decision maker's only concern is that they do not want to be treated unkindly; they are otherwise unconcerned about the payoffs received by other players.

To illustrate our idea more directly, we introduce the following notations. First, we set $L_{ij}(a_i, B_{ij}(h)) = \min\{\pi_i(a_i, B_{ij}(h)) - \pi_j(a_i, B_{ij}(h)), 0\}$, which describes that player i might suffer disutility if and only if they gain less than player j (a negative difference). Second, we set $\hat{L}_{ij}(B_{ij}(h)) = \max_{a'_i \in A_i} L_{ij}(a'_i, B_{ij}(h))$, which indicates the smallest negative difference in all available alternatives. Finally, we use $\{L_{ij}(a_i, B_{ij}(h)) - \hat{L}_{ij}(B_{ij}(h))\}^- = \min\{L_{ij}(a_i, B_{ij}(h)) - \hat{L}_{ij}(B_{ij}(h)), 0\}$ to denote that player i might not feel inequity even if they gain less than their opponents only if all alternatives will cause a greater negative difference. In other words, player i will experience disutility if and only if they gain less than their opponents (a negative difference) and the negative difference is also smaller than the smallest negative difference in all available alternatives given the history h. We employ the term Φ_i to denote the consequential kindness.

Definition 7 (Consequential kindness). *Player i's utility from the consequential kindness at history $h \in H$ is defined by*

$$\Phi_i = \alpha_i \cdot \{L_{ij}(a_i, B_{ij}(h)) - \hat{L}_{ij}(B_{ij}(h))\}^-,$$

where α_i is an exogenously given non-negative number, which measures how sensitive player i is to the consequence concerns with respect to player j. Different individuals and different opponents might have different values of α_i.

A value of α_i=0 indicates that player i does not care about how much player j might receive compared to what they receive, while α_i >0 indicates that player i suffers disutility when obtaining less than others.

3.4 The utility function and the equilibrium

Having defined two important kindness elements, the consequence part Φ_i and the intention part Ψ_i, we can move to the utility of player i in sequential games.

Definition 8 (The utility function). *Let player i and player j be two players of the game. The utility of player i at history $h \in H$ in the sequential move game is defined as*

$$U_i(a_i, B_{ij}(h), C_{iji}(h)) = \pi_i(a_i, B_{ij}(h), C_{iji}(h)) + \Psi_i + \Phi_i,$$

$$\Psi_i = \beta_i \cdot \delta_{ji} \cdot \pi_j(a_i, B_{ij}(h), C_{iji}(h)),$$
$$\Phi_i = \alpha_i \cdot \{L_{ij}(a_i, B_{ij}(h)) - \hat{L}_{ij}(B_{ij}(h))\}^-,$$

where Ψ_i represents player i's utility from the intentional kindness and Φ_i represents player i's utility from the consequential kindness.

The value of α_i and β_i are large (small) for these who care (do not care) about reciprocity. The values $\alpha_i = \beta_i = 0$ suggest that player i does not care about kindness, in which case the problem is reduced to classic game theory.

We have by now fully incorporated the reciprocity concerns into our model of game theory and economics. The introduced two-sided intentional kindness indicates that a player's utility also depends on the origin of outcomes; the introduced one-sided consequential kindness suggests that players care about their relative outcomes. One thing that needs to be clear in the intentional kindness term is that the beliefs do not belong to the strategy space, which is fixed before the game proceeds. According to the above discussion, player i chooses the action that maximizes their utility at the given history h.

Definition 9 (Expected reciprocity equilibrium). *The profile $\{a^*, B^*_{ij}(h), C^*_{iji}(h)\}$ is an expected reciprocity equilibrium (ERE) if for all $i \in N$ and for each history $h \in H$ it holds that*

(i) $a^*_i \in \underset{a_i \in A_{i,h}}{\arg\max} \quad U_i(a_i, B_{ij}(h), C_{iji}(h))$

(ii) $B^*_{ij}(h) = a^*_j \;$ for all $\; j \neq i$

(iii) $C^*_{iji}(h) = a^*_i \;$ for all $\; j \neq i$

The condition (i) indicates that players make the optimal choices at history h given their first-order and second-order beliefs, while (ii) and (iii) guarantee that the updated beliefs are consistent with the players' actions.

Theorem. *An expected reciprocity equilibrium always exists.*

See proof in Appendix *A.1.*

4 Applications

In this section, we apply our model to some famous experimental games: the ultimatum game, the sequential prisoner's dilemma with punishment, and the prisoner's dilemma with asymmetric payoffs. We discuss the predictions of our model in these games. All games we will analyse are sequential games, and we mainly focus on the second mover's behaviour given different histories. For the first two games we discuss in this section, all existing intention-based reciprocity models or consequence-based models mentioned in Section 1 fail to predict the empirical results. We start with four mini-ultimatum games, shown in Figures 1, 2, 3, and 4, proposed by Falk, Fehr, and Fischbacher (2003), who found that the rejection rate of the (8,2) offer decreases from Figure 1 to Figure 4. These four games can help us to understand the role of intentional and consequential kindness in games. Next, we review the sequential prisoner's dilemma with or without punishment, in which the material payoffs and original form are taken from the experimental study of Orhun (2018). The experimental results show that the chance of punishment for the second player decreases their perception of how kind the first player is, which is in agreement with our predictions. The third example is also the sequential prisoner's dilemma but with asymmetric payoffs. The basic step is the same as the second application. So we will not list all of them in detail in this game, but only aim to compare the position of players when their payoffs are symmetric or asymmetric. We take the original form from the study of Ahn et al. (2007). Our results are consistent with their experimental findings.

4.1 Negative reciprocity: four mini-ultimatum games

The ultimatum game is one of the most well-known games that reflects negative reciprocity. As discussed in the last three sections, in the ultimatum game, the proposer as the first mover allocates a fixed amount of money between them and the responder. The responder as the second mover either accepts or rejects the offer. If they accept, the resulting payoffs follow the allocation made by the proposer. In case of rejection, the payoffs are zero for both the proposer and the responder. Rejection always implies that the responder would like to sacrifice their own material payoff to punish the proposer (an example of negative reciprocity). Therefore, When the responder decides to accept or reject the allocation is crucial.

An appropriate experimental environment always sheds light on the underlying concepts. In this section, we analyse the responder's behaviour in four mini-ultimatum games proposed by Falk, Fehr, and Fischbacher (2003): games (a)–(d), illustrated in Figures 1 – 4, respectively. Empirical results show that the rejection rate of the (8,2) offer decreases from (a) to (d): in (a) it is 44.4%, in (b) 26.7%, in (c) 18%, and in (d) 8.9%. However, the difference between game (c) and game (d) is not statistically significant ($p = 0.369$, two-sided). By adopting our model, our predictions can be summarized as follows:

Result 1. *"O_1" is the efficient strategy in game (a), game (b), game (c), and game (d); "O_2" is the efficient strategy in game (a), game (b), and game (c) but not in game (d).*

Result 2. *In game (a) and game (b), if the proposer chooses "O_2", the responder will accept the offer (by choosing "y") in every ERE.*

Result 3. *In game (a) and game (b), if the proposer chooses "O_1", the responder will* *reject the offer (by choosing "n") if $\alpha \geq 1/3$.*

Result 4. *In game (a), suppose $\alpha < 1/3$. If the proposer chooses "O_1", the responder will accept the offer (by choosing "y") if $\beta < \frac{(2-6\alpha)(1+e^5)}{40e^5}$, will reject the offer (by choosing "n") if $\beta > \frac{(2-6\alpha)(e^8+e^5)}{24e^5}$, and will choose randomly if $\frac{(2-6\alpha)(1+e^5)}{40e^5} \leq \beta \leq \frac{(2-6\alpha)(e^8+e^5)}{24e^5}$.*

Result 5. *In game (b), suppose $\alpha < 1/3$. If the proposer chooses "O_1", the responder will accept the offer (by choosing "y") if $\beta < \frac{(2-6\alpha)(1+e^2)}{64e^2}$, will reject the offer (by choosing "n") if $\beta > \frac{(2-6\alpha)(e^8+e^2)}{48e^2}$, and will choose randomly if $\frac{(2-6\alpha)(1+e^2)}{64e^2} \leq \beta \leq \frac{(2-6\alpha)(e^8+e^2)}{48e^2}$.*

Result 6. *More responders are willing to accept (by choosing "y") the (8,2) offer (when the proposer chose "O_1") in game (b) than in game (a) given a uniform distribution over the population.*

Result 7. *In game (c) and (d), if the proposer chooses "O_1", the responder will accept the offer (by choosing "y") if $\alpha < 1/3$, will reject the offer (by choosing "n") if $\alpha > 1/3$, and will choose randomly if $\alpha = 1/3$.*

Result 8. *More responders are willing to accept (by choosing "y") the (8,2) offer (when the proposer chose "O_1") in games (c) and (d) than in games (a) and (b).*

See proof in Appendix *A.2.1.*

These 8 results provide us with a very plausible and empirically supported explanation. In Result 1, the responder prefers to punish the proposer without cost if the proposer makes an unkind offer (in game (d)). Our Definitions 2 and 3 suggest that only "O_1" is efficient in game (d). In the view of the responder, the proposer will never choose "O_2", which is also why the difference between game (c) and game (d) is not statistically significant in the experimental findings. Moving to Result 7, we predict that the responder in games (c) and (d) will behave indifferently, which is in line with the findings in the experiment.

Result 2 shows the responder would like to accept a sufficiently kind and fair offer. In games (a) and (b), we notice that if the proposer chooses "O_2", the responder will never experience inequity by our definition of the consequence part Φ_R plus the kind intention by our definition of the intention part Ψ_R. The responder must accept the offer (by choosing "y") in every ERE.

Results 3 and 7 reveal that the responder would reject the unkind offer as long as the responder is extremely inequity-averse ($\alpha > 1/3$). Results 4 and 5 show that as long as the responder is not too inequity-averse ($\alpha < 1/3$) and does not care too much about intentions, they would accept the unfair (8,2) offer. But if the responders care about the intentions, they will reject this unkind offer.

Results 4 and 5 have different thresholds, so we have different rejection rates between game (a) and game (b). To see this, we prove in Appendix *A.2.1.* that more proposers would like to choose "y" in game (b) than in game (a). This prediction is consistent with the experimental findings.

For games (c) and (d), we observe that no intention exists in these two games by Result 7. Compare with the negative intention in games (a) and (b) induced by choosing "O_1'', the proposer is more likely to accept the unfair offer in games (c) and (d). Thus we have Result 8, for the same distribution over the population, more proposers would like to choose "y" in games (c) and (d) than in games (a) and (b). Again, this prediction is in line with the experimental findings.

4.2 Positive reciprocity: the sequential prisoner's dilemma

In the sequential prisoner's dilemma, the first mover can either cooperate or defect. After observing the first mover's choice, the second mover faces the same choice. The standard subgame-perfect solution is that both parties defect. However, many experimental studies reveal that mutual cooperation might be another possible solution (Ahn et al., 2001; Dhaene and Bouckaert, 2010; Charness, Rigotti, and Rustichini, 2016; Engel and Zhurakhovska, 2016; Gächter et al., 2021; Gächter, Lee, and Sefton, 2022; Schneider and Shields, 2022). Our model predicts a similar finding.

In this section we analyse the sequential prisoner's dilemma with and without punishment, shown in Figure 5, proposed by Orhun (2018). Their experimental results reflect that the existing models (Rabin, D&K, F&F, Fehr and Schmidt (1999)) fail to predict the behaviour of players in the sequential prisoner's dilemma with punishment[17]. In both cases, $m = 1.5$ means that player 2 has the option to punish player 1 (this is costly punishment) if player 1 treats them unkindly by choosing "D". With $m = 6.5$, this option is not available. Instead, player 2 has a chance to reward player 1's unkind behaviour "D" at a cost, but no one is willing to reward the unkind action either in the experimental results or the prediction of our model. We will derive the ERE for different values of m, and observe whether our analysis is consistent with the experimental results as shown in Table 1.

RESULTS				
Version	1 choice (C)	2 FOE (C)	2 choice ($c/p \| D$)	2 choice ($c \| C$)
$m = 1.5$	93.22%	72.00%	25.00%	34.55%
$m = 6.5$	65.71%	41.00%	04.17%	56.52%

Table 1: Results of the game shown in Figure 5

[17]The D&K and Rabin models predict a more positive reciprocity environment in the case with $m = 1.5$ than when $m = 6.5$. Hence, player 1 is more kind when $m = 1.5$ than when $m = 6.5$. The F&F model predicts that player 2 will always choose "c" after player 1 chooses "C" when $m = 1.5$ and when $m = 6.5$ (in both experiments). A proof is given in the Appendix of Orhun (2018).

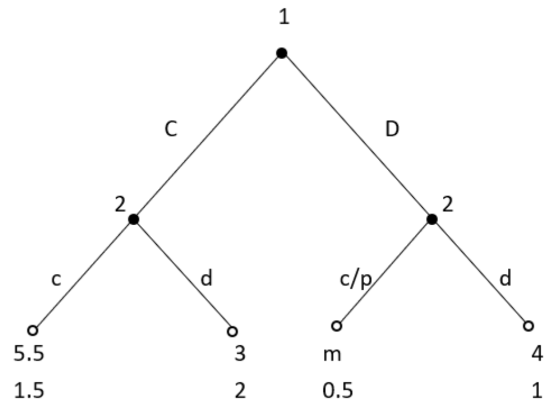

Figure 5: The sequential prisoner's dilemma with/without punishment

Table 2 shows the players' behaviour in experiments on the game from Figure 5 with different values of m. Reading the data in the table, we find that player 2 prefers to choose "c" conditional on player 1 choosing "C" when $m = 6.5$ more than when $m = 1.5$. One reasonable explanation is that when $m = 6.5$, the better choice for player 1 is playing "D" (since they can at least obtain a payoff of 4) instead of playing "C" (for which the worst payoff is 3). In contrast, when $m = 1.5$, the better choice for player 1 is "C" (since at least they can receive 3) instead of playing "D" (for which the worst payoff is 1.5). Therefore, when $m = 6.5$ and player 1 chooses "C", they are perceived as more kind than in the case in which they choose "C" when $m = 1.5$. The experiments with $m = 6.5$ reveal that player 1 would like to sacrifice their material payoff to help player 2 gain more, this effect is not fully apparent when $m = 1.5$ as player 2 does not know the exact intentions of player 1. There are two potential intentions for player 1 to cooperate: one is that player 1 wants to be a kind person and bring their opponent a higher material payoff like when $m = 6.5$. Another is that player 1 wants to avoid the punishment since this may lead to the worst material payoff ($1.5 < 4$, 3, or 5.5). As a result, more player 2 is more likely to cooperate (56.52%) after cooperation when $m = 6.5$ than when $m = 1.5$ (34.55%). This finding contradicts the existing models (see more detail in the Appendix of Orhun (2018)). We will derive the ERE by adopting our model and see whether our model can successfully predict player 2's behaviour[18].

[18] *Note* : in the following study, for the sake of simplicity, we denote player 2's belief about player 1's belief about their choice "c" after "C" as p and player 2's belief about player 1's belief about their choice "p" after "D" as q.

4.2.1 General sequential prisoner's dilemma ($m = 6.5$)

We discuss the general sequential prisoner's dilemma with $m = 6.5$ first, calculate the ERE of player 2 conditional on player 1's behaviour, compare our predictions with the next section in which player 2 has the chance to punish unkind behaviour, and see whether our prediction is consistent with the experimental results shown in Table 2. Here we mainly analyse player 2's behaviour, and our predictions can be summarized as follows:

Result 9. *"C" and "D" are both efficient strategies. If player 1 defects (by choosing "D"), player 2 will always defect (by choosing "d") as a response in every ERE.*

First, "C" and "D" are both efficient strategies according to our Definitions 2 and 3. To explain player 2's behaviour in this result, if player 1 chooses "D", then for any possible outcomes for player 2, they will always gain less than player 1 playing "C". This means that whatever player 1 thinks player 2 will choose, player 1 cannot be kind if player 1 plays "D". In other words, $\delta_{12} < 0$. Moreover, by comparing two consequence terms $\Phi_2(d|D) = \alpha \cdot \{[1-4]^- - [0.5-6.5]^-\}^- = 0$ and $\Phi_2(c|D) = \alpha \cdot \{[0.5-6.5]^- - [1-4]^-\}^- = -3\alpha$, we say that if player 2 plays "c" they will experience inequity but if they play "d" they will not experience it. Overall, player 2 must defect (by choosing "d") in every ERE.

Result 10. *If player 1 cooperates (by choosing "C"), the following holds in ERE:*

(i) *If* $\beta > \frac{(2+12\alpha)(e^{5.5}+e^4))}{5e^4}$, *player 2 will cooperate (by choosing "c").*

(ii) *If* $\beta < \frac{(2+12\alpha)(e^3+e^4)}{10e^4}$, *player 2 will defect (by choosing "d").*

(iii) *If* $\frac{(2+12\alpha)(e^3+e^4)}{10e^4} \leq \beta \leq \frac{(2+12\alpha)(e^{5.5}+e^4))}{5e^4}$, *player 2 will cooperate (by choosing "c") with probability* p *that satisfies* $2 + 12\alpha = 5\beta(2-p)\frac{e^4}{e^{5.5p+3(1-p)}+e^4}$.

See proof in Appendix *A.2.2.1*.

In the case in which player 1 cooperates, our result is very intuitive. When player 1 puts the reciprocity in a very important position, then they would like to cooperate to appreciate player 1's kindness (Result 10(i)). On the other hand, if player 2 is a sufficiently selfish person who does not care about the kindness of others, then they would like to defect (Result 10(ii)). For an intermediate value, player 2 will cooperate with a given probability p (Result 10(iii)).

4.2.2 Sequential prisoner's dilemma with punishment ($m = 1.5$)

Next, we move to the sequential prisoner's dilemma with punishment where $m = 1.5$. The same reason as when $m = 6.5$, player 1 is viewed as unkind if they play "D".

The difference here is that player 2 might choose "p" after "D" to punish the unkind behaviour at a cost to themselves. Therefore, we may have $q \geq 0$ when $m = 1.5$. In this section, we still focus on player 2's behaviour, and our predictions can be summarized as follows:

Result 11. *"C" and "D" are both efficient strategies. If player 1 defects (by choosing "D'), defection (by choosing "d") for player 2 is not the unique ERE.*

Result 12. *If player 1 cooperates (by choosing "C"), the following holds in ERE:*

(i) *If* $\beta > \frac{(2+12\alpha)(e^{5.5}+e^{1.5q+4(1-q)})}{5(1+q)e^{1.5q+4(1-q)}}$*, player 2 will cooperate (by choosing "c").*

(ii) *If* $\beta < \frac{(2+12\alpha)(e^{3}+e^{1.5q+4(1-q)})}{5(2+q)e^{1.5q+4(1-q)}}$*, player 2 will defect (by choosing "d").*

(iii) *If* $\frac{(2+12\alpha)(e^{3}+e^{1.5q+4(1-q)})}{5(2+q)e^{1.5q+4(1-q)}} \leq \beta \leq \frac{(2+12\alpha)(e^{5.5}+e^{1.5q+4(1-q)})}{5(1+q)e^{1.5q+4(1-q)}}$*, player 2 will cooperate (by choosing "c") with a probability* p *that satisfies* $2+12\alpha = 5\beta(2-p+q)\frac{e^{1.5q+4(1-q)}}{e^{5.5p+3(1-p)}+e^{1.5q+4(1-q)}}$.

See proof in Appendix *A.2.2.2*.

Result 11 is straightforward: it shows that as long as the first player shows their unkindness to player 2, player 2 prefers to pay a cost to punish player 1. This result also implies that q is greater than 0, so $q \in (0, 1]$. The explanation for Result 2 is similar to the general sequential prisoner's dilemma. When player 2 cares sufficiently about reciprocity, they will cooperate to reward player 1's kindness (Result 12(i)). On the other hand, if player 2 is a sufficiently selfish person who does not care about the kindness of others, then they will defect (Result 12(ii)). For an intermediate value, player 2 will cooperate with a given probability p (Result 12(iii)).

4.2.3 Comparison: when player 2 will cooperate given that player 1 cooperates

Result 13. *Player 2 is more likely to cooperate (by choosing "c") given that player 1's cooperation (player 1 chose "C") when* $m = 6.5$ *than when* $m = 1.5$.

First, let us see when player 2 cooperates (by choosing "c") conditional on player 1 cooperating (by choosing "C"). In the last two subsections, we have shown that with different values of m, we have two different thresholds:

$2 + 12\alpha - 5\beta\frac{e^4}{e^{5.5}+e^4} < 0$, thus $\beta > \frac{(e^{5.5}+e^4)(2+12\alpha)}{5e^4}$ (in Section 4.2.1)

$2 + 12\alpha - 5\beta(1+q)\frac{e^{1.5q+4(1-q)}}{e^{5.5}+e^{1.5q+4(1-q)}} < 0$, thus $\beta > \frac{(e^{5.5}+e^{4-2.5q})(2+12\alpha)}{5(1+q)e^{4-2.5q}}$ (in Section 4.2.2)

The above two thresholds suggest when player 2 would like to cooperate for $m = 6.5$ and $m = 1.5$. In the latter threshold, we have the non-negative probability q. To see the difference, we can compare the threshold of each case and gain more intuition: $\frac{(e^{5.5}+e^4)(2+12\alpha)}{5e^4} - \frac{(e^{5.5}+e^{4-2.5q})(2+12\alpha)}{5(1+q)e^{4-2.5q}}$, notice α is a non-negative value. Hence, we can set $\Delta = (2+12\alpha)[\frac{(e^{5.5}+e^4)}{5e^4} - \frac{(e^{5.5}+e^{4-2.5q})}{5(1+q)e^{4-2.5q}}]$ where $q \in (0,1]$, because the experimental results indicate a non-positive value of Δ. Let us derive the smallest value of Δ. First, $\frac{\partial\Delta}{\partial q} = -(2+12\alpha)\frac{-2.5e^{4-2.5q}\cdot 5(1+q)e^{4-2.5q}-[5e^{4-2.5q}+5(1+q)(-2.5)e^{4-2.5q}](e^{5.5}+e^{4-2.5q})}{[5(1+q)e^{4-2.5q}]^2} = (2+12\alpha)\frac{e^{8-5q}\cdot(5-7.5e^{1.5+2.5q})-12.5\cdot q\cdot e^{9.5-2.5q})}{[5(1+q)e^{4-2.5q}]^2} < 0$ since $q \in (0,1]$ and $\alpha \geq 0$. Then let us look at when $q = 0$, then $\Delta = (2+12\alpha)[\frac{(e^{5.5}+e^4)}{5e^4} - \frac{(e^{5.5}+e^4)}{5e^4}] = 0$. As a result, $\Delta \leq 0$. This means that the threshold value of β with $m = 6.5$ is lower than for $m = 1.5$. As a result, for the same distribution over the population, more player 2 will choose "c" with $m = 6.5$ than with $m = 1.5$.

Next, let us see when player 2 chooses "d" conditional on player 1 choosing "C". In the last two subsections, we also have shown that with different values of m, we obtain two different thresholds:

$2+12\alpha-10\beta\frac{e^4}{e^3+e^4} > 0$, thus $\beta < \frac{(2+12\alpha)(e^3+e^4)}{10e^4}$ (in Section 4.2.1)

$1+6\alpha-5\beta(1+0.5q)\frac{e^{1.5q+4(1-q)}}{e^3+e^{1.5q+4(1-q)}} > 0$, thus $\beta < \frac{(2+12\alpha)(e^3+e^{1.5q+4(1-q)})}{5(2+q)e^{1.5q+4(1-q)}}$ (in Section 4.2.2)

Similarly, we obtain different thresholds for when player 2 defects given different m. We find that $\Delta = \frac{(2+12\alpha)(e^3+e^4)}{10e^4} - \frac{(2+12\alpha)(e^3+e^{1.5q+4(1-q)})}{5(2+q)e^{1.5q+4(1-q)}} = \frac{2+12\alpha}{5}\cdot(\frac{e^{-1}+1}{2} - \frac{e^{2.5q-1}+1}{2+q})$ where $\alpha \geq 0$ and $q \in (0,1]$, because the experimental results indicate a non-positive value of Δ. Let us derive the smallest value of Δ. First, $\frac{\partial\Delta}{\partial q} = -\frac{2+12\alpha}{5}\cdot(\frac{2.5e^{2.5q-1}(2+q)-(e^{2.5q-1}+1)}{(2+q)^2}) = -\frac{2+12\alpha}{5}\cdot(\frac{4e^{2.5q-1}+2.5qe^{2.5q-1}-1}{(2+q)^2})$. Set $X = (4e^{2.5q-1}+2.5qe^{2.5q-1}-1)$, then $\frac{\partial X}{\partial q} = (10e^{2.5q-1} + 2.5e^{2.5q-1}+6.25qe^{2.5q-1}) > 0$ must be satisfied. Therefore, the smallest value of X is $4e^{-1}-1$, which is greater than 0. Furthermore, $\frac{\partial\Delta}{\partial q} < 0$ must be satisfied. The largest value of $\Delta = 0$ is obtained when $q = 0$. As a result, $\Delta \leq 0$. This means that the threshold value of β with $m = 6.5$ is less than for $m = 1.5$. As a result, for the same distribution over the population, fewer player 2 will choose "d" with $m = 6.5$ than $m = 1.5$.

Finally, let us move to a more complicated case when player 2 chooses a mixed strategy. We denote by p the probability that player 2 will cooperate given that player 1 cooperates. In the last two sections, we have shown that with different values of m, we have the different cooperation probability p:

$2+12\alpha = 5\beta(2-p)\frac{e^4}{e^{5.5p+3(1-p)}+e^4}$ (in Section 4.2.1)

$2+12\alpha = 5\beta(2-p+q)\frac{e^{1.5q+4(1-q)}}{e^{5.5p+3(1-p)}+e^{1.5q+4(1-q)}}$ (in Section 4.2.2)

We have that the cooperation rate p satisfies the above two equations given a fixed β. For the same β and α, we set $M = \frac{2+12\alpha}{5\beta} = \frac{2-p+q}{e^{2.5p+2.5q-1}+1}$. The equation in Section 4.2.1 is the case where $q = 0$ and that in Section 4.2.2 is the case where $q \in (0,1]$. Now, if $q \in (0,1]$, will p decrease or increase compared to $q = 0$? Notice that M is a fixed value, $\frac{\partial M}{\partial p} = \frac{-(e^{2.5p+2.5q-1}+1)-2.5e^{2.5p+2.5q-1}(2-p+q)}{(e^{2.5p+2.5q-1}+1)^2} < 0$ means that p decreases will lead to M increasing. On the other hand, $\frac{\partial M}{\partial q} = \frac{(e^{2.5p+2.5q-1}+1)-2.5e^{2.5p+2.5q-1}(2-p+q)}{(e^{2.5p+2.5q-1}+1)^2} = \frac{1-2.5e^{2.5p+2.5q-1}(1.6-p+q)}{(e^{2.5p+2.5q-1}+1)^2}$. Set $L = e^{2.5p+2.5q-1}(1.6-p+q)$, then $\frac{\partial L}{\partial q} = 2.5e^{2.5p+2.5q-1}(1.6-p+q)+e^{2.5p+2.5q-1} > 0$. So $\max(\frac{\partial M}{\partial q}) = \frac{1-2.5e^{2.5p-1}(1.6-p)}{(e^{2.5p-1}+1)^2}$. Set $J = e^{2.5p-1}(1.6-p)$, then $\frac{\partial J}{\partial p} = 2.5e^{2.5p-1}(1.6-p) - e^{2.5p-1} = e^{2.5p-1}(3-2.5p) > 0$, thus we have $\max(\frac{\partial M}{\partial q}) = \frac{1-4e^{-1}}{(e^{-1}+1)^2} < 0$, so that q decreasing will lead to M increasing. Therefore, when q is nonzero, to keep M the same as when $q = 0$, p must decrease. This suggests that for the same person (M is fixed), they are more willing to cooperate when $m = 6.5$ than when $m = 1.5$.

To summarize these three cases, more player 2 will cooperate given that player 1 cooperates when $m = 6.5$ than when $m = 1.5$. This result is consistent with the experimental results.

4.3 Different roles: prisoner's dilemma with asymmetric payoffs

In the third application, we discuss the impact of a player's role. We revisit the issue of an asymmetric payoff structure in the sequential prisoner's dilemma. The payoff structure of the game that we will analyse is proposed by Ahn et al. (2007) and shown in Figure 7. There are three cases; The first, second, and third rows correspond to treatment 1, treatment 2, and treatment 3, respectively.

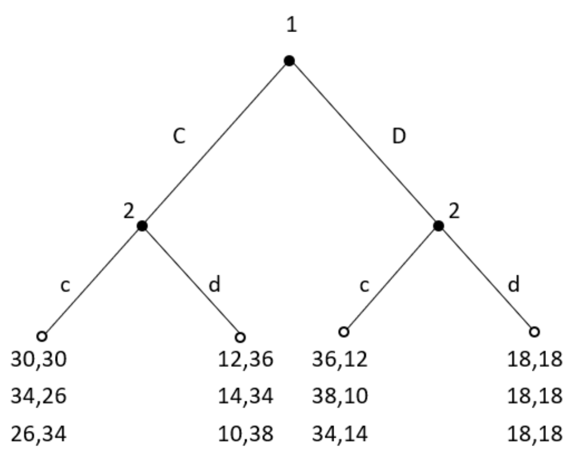

Figure 6: The sequential prisoner's dilemma with asymmetric payoffs

As Ahn et al. (2007) noted, the advantaged player receives a higher material payoff from mutual cooperation than the disadvantaged player and loses less when cooperating with the player who defects. In Figure 6, player 1 is the advantaged player in treatment 2 and the disadvantaged player in treatment 3. Throughout this part, our purpose is to observe whether the players' actions will change when they face different roles.

The analysis is similar to that in Section 4.2.1. In this subsection, our purpose is to theoretically predict player 2's response given player 1's cooperative behaviour and different material payoffs. The predictions can be summarized as follows:

Result 14. *In all treatments, "C" and "D" are both efficient strategies. If player 1 defects (by choosing "D"), player 2 will always defect (by choosing "d") as a response in every ERE.*

Result 15. *In treatment 1, if player 1 cooperates (by choosing "C"), the following hold in ERE:*

(i) *If* $\beta > \frac{e^{30}+e^{18}}{36e^{18}}$*, player 2 will cooperate (by choosing "c").*

(ii) *If* $\beta < \frac{e^{12}+e^{18}}{54e^{18}}$*, player 2 will defect (by choosing "d").*

(iii) *If* $\frac{e^{12}+e^{18}}{54e^{18}} \leq \beta \leq \frac{e^{30}+e^{18}}{36e^{18}}$*, player 2 will cooperate (by choosing "c") with probability* p *that satisfies* $3\beta(18-6p) \cdot \frac{e^{18}}{e^{18p+12}+e^{18}} = 1$.

Result 16. *In treatment 2. If player 1 cooperates (by choosing "C"), the following holds in ERE:*

(i) *If* $\beta > \frac{(e^{34}+e^{18})(\alpha+1)}{20e^{18}}$*, player 2 will cooperate (by choosing "c").*

(ii) *If* $\beta < \frac{(e^{14}+e^{18})(1+\alpha)}{40e^{18}}$*, player 2 will defect (by choosing "d").*

(iii) *If* $\frac{(e^{14}+e^{18})(1+\alpha)}{40e^{18}} \leq \beta \leq \frac{(e^{34}+e^{18})(\alpha+1)}{20e^{18}}$*, player 2 will cooperate (by choosing "c") with probability* p *that satisfies* $20\beta(2-p) \cdot \frac{e^{18}}{e^{20p+14}+e^{18}} = 1+\alpha$.

Result 17. *In treatment 3. If player 1 cooperates (by choosing "C"), the following holds in ERE:*

(i) *If* $\beta > \frac{e^{26}+e^{18}}{64e^{18}}$*, player 2 will cooperate (by choosing "c").*

(ii) *If* $\beta < \frac{e^{10}+e^{18}}{80e^{18}}$*, player 2 will defect (by choosing "d").*

(iii) *If* $\frac{e^{10}+e^{18}}{80e^{18}} \leq \beta \leq \frac{e^{26}+e^{18}}{64e^{18}}$*, player 2 will cooperate (by choosing "c") with probability* p *that satisfies* $4\beta(20-4p) \cdot \frac{e^{18}}{e^{16p+10}+e^{18}} = 1$.

Result 18. *More player 2s are willing to cooperate (by choosing "c") given player 1's cooperation (player 1 chose "C") in treatment 3 than treatment 1 and treatment 2; and* *more player 2s are willing to cooperate (by choosing "c") given player 1's cooperation (player 1 chose "C") in treatment 1 than treatment 2.*

See proof in Appendix *A.2.3.*

Results 14, 15, 16, and 17 have the same logic as in Section 4.2.1. These results suggest that if player 2 cares about reciprocity, they would like to sacrifice some material payoffs to help player 1, and if they do not care about reciprocity, they will behave selfishly. Moreover, the proof of Result 18 is similar to that of Result 13. More detail is provided in Appendix *A.2.3.*.

Our result is consistent with the experimental findings of Ahn et al. (2007), in which 43% of players cooperate after cooperation in treatment 3, 35% cooperate after cooperation in treatment 1, and only 21% will cooperate after cooperation in treatment 2.

5 Discussion

5.1 Comparison of efficient strategy

5.1.1 Three definitions of efficient strategy

The definition of efficient strategy in Rabin is a set that depends on player i's beliefs; it can be expressed as

$$E_i^{Rabin}(B_{ij}) = \left\{ a_i \in A_i \;\middle|\; \begin{array}{l} \text{if } \nexists\ a_i' \in A_i \text{ such that:} \\ (i)\pi_k(a_i', B_{ij}) \geq \pi_k(a_i, B_{ij}) \text{ for all } k \in \{i,j\} \\ (ii)\pi_k(a_i', B_{ij}) > \pi_k(a_i, B_{ij}) \text{ for some } k \in \{i,j\} \end{array} \right\}.$$

This definition indicates that player i's action is efficient if and only if it is Pareto efficient given the first-order belief "B_{ij}". In Figure 7 for $a = 0$, for example, if $B_{12}(c|C) > 1/2$, then only "C" is efficient strategy. If $B_{12}(c|C) \leq 1/2$, then both "C" and "D" are efficient strategies. Since Rabin cares about normal-form games, thus they do not have belief updating rules like in D&K and our model, it retains some limitations when we study the extensive-form games (see discussion in D&K).

In order to extend the study to a more general analysis, D&K have proposed a new definition of efficient strategy that does not depend on player's beliefs; they use the following definition:

$$E_i^{DK} = \left\{ a_i \in A_i \;\middle|\; \begin{array}{l} \text{if } \nexists\ a_i' \in A_i \text{ such that:} \\ (i)\pi_k(a_i'(h), a_j(h)) \geq \pi_k(a_i(h), a_j(h)) \text{ for all } h, a_j, k \in \{i,j\} \\ (ii)\pi_k(a_i'(h), a_j(h)) > \pi_k(a_i(h), a_j(h)) \text{ for some } h, a_j, k \in \{i,j\} \end{array} \right\}$$

In this definition, they point out that player i's strategy should be efficient if it is Pareto efficient for at least one strategy for player j. That is, a strategy is not efficient if there is a strategy that leads higher material payoff for any player in the game. This definition performs better than the Rabin definition in many situations, but we will see that it retains a series of limitations and paradoxes in our discussion in Section 5.1.2.

In our definition, to understand the truly intentional kindness of the player (if they aim to be kind or avoid punishment), we propose a special second-order belief $C_{iji}^{pwo}(a_{j,h})$ in Definition 2, which is based on the intuition of the potential worst outcome. The definition can help us to select the efficient strategies:

$$E_j^{pwo} = \left\{ \quad a_j \in A_j \quad \middle| \quad \begin{array}{l} \text{if } \nexists\, a_j' \in A_j \text{ such that:} \\ (i)\pi_k(a_j', C_{iji}^{pwo}(a_{j,h})) \geq \pi_k(a_j, C_{iji}^{pwo}(a_{j,h}))\ \forall k \in \{i,j\} \text{ and} \\ (ii)\ \pi_k(a_j', C_{iji}^{pwo}(a_{j,h})) > \pi_k(a_j, C_{iji}^{pwo}(a_{j,h}))\ \text{ for some } k \in \{i,j\} \end{array} \right\}$$

In words, E_j^{pwo} suggests that player j's strategy is efficient if it is Pareto efficient based on a special second-order belief $C_{iji}^{pwo}(a_{j,h}))$ that is defined by Definition 2. If we apply the potential worst outcomes to define the second-order belief, then for all potential worst outcomes of player i, there does not exist any strategy that can benefit all players, then we say player j's strategy is efficient. Otherwise, it is not efficient.

5.1.2 Paradox with the definition of efficient strategy

When thinking about players' intentions, there always exist some wasteful strategies, as discussed in Section 3. Before evaluating the kindness of other players, it is necessary to rule out such strategies. So we have three definitions, as discussed in the previous subsection, for how to select the efficient strategy set. We find that only our definition can meet the purpose.

Let us take the game in Figure 7, with $a = 0$ and $a = -2$, as an example. Player 1's cooperative behaviour "C" can always benefit player 2 compared with behaviour "D" since player 2 can receive more material payoffs no matter what their response is $(1, 2 > -1, 0)$, but with different intentions. In the case of $a = 0$, player 2 can receive at least 0 (where player 2 plays "d") when player 1 chose "D". On the other hand, if player 1 plays "C", they might receive -1 (if player 2 plays "d") although they might get a positive payoff of 1 (if player 2 plays "c"). Here we might infer that since player 1 puts themselves in a fragile position to help player 2 obtain a larger payoff, player 1 is kind. When $a = 0$, we can see the limitation of Rabin's definition $E_i^{Rabin}(B_{ij})$. Our definition and D&K's definition both indicate that "C" and "D" are efficient strategies. Under Rabin's definition, however, the action "D" is efficient only when player 1's first-order belief $B_{12}(c, C)$ is greater than 1/2, otherwise only "C" is efficient. Therefore, this definition rules out some strategies that are efficient. Moreover, Rabin's definition excessively depends on the player's beliefs, so is not suitable for the study of sequential games as discussed in D&K's paper.

On the other hand, in the case of $a = -2$, the worst outcome of playing "D" is -2 (when player 2 plays "d")[19], which is strictly worse than playing "C" regardless of player 2's response. Therefore, we cannot infer any intentional kindness of player 1 even though they play the cooperative behaviour "C". Furthermore, player 1 choosing "C" cannot make them more kind in the view of player 1. Under this scenario, we can see drawbacks of D&K's definition of efficient strategy. Both our definition and Rabin's definition suggest that only "C" is an efficient strategy, which is consistent with our inference. In D&K, however, both "C" and "D" are efficient. D&K's definition predicts the same value of kindness for player 1 no matter what the value of a is. This is a contradiction. The reason why their definition fails to explain is that it is independent of the player's beliefs, and it does not show us any psychological concerns when we study reciprocity. Moreover, the definition classifies too many efficient strategies and their model predicts a situation that is too kind, as it gives rise to a reference point that is relatively low.

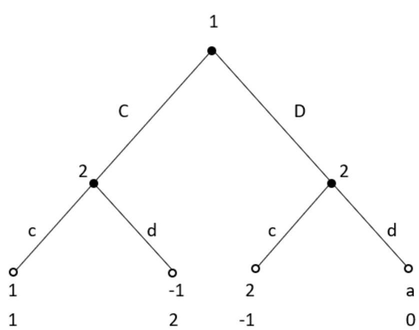

Figure 7: The sequential prisoner's dilemma

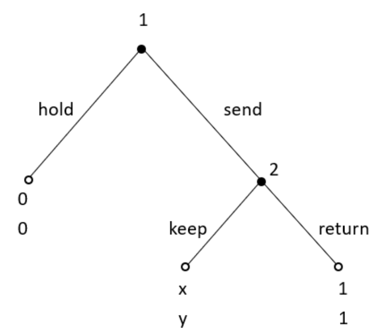

Figure 8: The trust game

Isoni and Sugden (2019)also propose a paradox of trust when studying reciprocity with existing models. They argue that although D&K's model is compatible with the properties of the trust world, it does not provide a psychologically convincing explanation for why reciprocal kindness can clarify the trust world[20]. In Figure 8, for example, having G_1: $(x, y) = (1/2, -1/2)$ and G_2: $(x, y) = (1/2, 1/2)$. The paradox appears once we apply their definition. Intuitively, if "send" is perceived as a kind behaviour in G_1, then "send" should also be perceived as kind in G_2, as the only distinction between the two games is that material payoff for "send" or "keep" for player 2 is smaller in G_1 than in G_2. But in D&K, "hold" is not efficient in G_2 (suggesting that "send" should result in 0 kindness) and "hold" is efficient in G_1 (suggesting that "send" should result in positive kindness). Again, this contradicts intuition and experimental results.

[19]This is in fact the only realistic result, since no one is willing to hurt themselves to help an unkind person.

[20]They further argue that D&K's new definition in 2019 still cannot explain the trust world.

This evidence encourages us to propose a more appropriate methodology to avoid the contradiction, especially when one is able to punish others. Definitions 2 and 3 serve our purpose and provide us with a plausible explanation. If we apply our definition of efficient strategy, in the game in Figure 7, "D" is efficient when $a = 0$ and is not efficient when $a = -2$ in player 2's consideration. In Figure 8, for G_1: $(x, y) = (1/2, -1/2)$ and G_2: $(x, y) = (1/2, 1/2)$, only "send" is efficient strategy. Thus the paradox is resolved.

5.1.3 Connections between the three definitions

As we have mentioned, Rabin's definition depends on the players' beliefs[21]. It might reject some actions that truly matter, and might cause contradictions in sequential games (see the D&K paper for details). D&K's definition shows us a more kind and gentle person, but it does not indicate too much psychological concerns and it leads to paradoxes in many games (such as the trust game and the ultimatum game).

Proposition 1. *Consider the efficient minimum material payoff of player $i \in N$. Then we must have:*

$$E_i^{pwo} \subseteq E_i^{DK} \text{ and } \min_{a_i \in E_i^{PWO}} \pi_i(a_i, B_{ij}) \geq \min_{a_i \in E_i^{DK}} \pi_i(a_i, B_{ij}).$$

Our definition and D&K's definition are both independent of the players' initial beliefs. But our definition of efficient strategy is dependent on a special second-order belief. This second-order belief selects a set of strategies that must be included in the behavioural strategy in D&K. In other words, we rule out some strategies in D&K. Therefore, if one action is efficient in our definition, it must be efficient in D&K's definition. Hence, the minimum material payoff in our definition should be less than or equal to D&K's definition. This proposition also suggests that one player's behaviour would be perceived as less kind and predicts less positive reciprocity in our model than in D&K.

Proposition 2. *If there is a game in which $E_i^{Rabin} = E_i^{pwo} = E_i^{DK}$ after ruling out a set of strategies, then these strategies must be strictly worse than all other strategies for both players.*

To see this, according to three definitions of efficient strategy. E_i^{Rabin} depends on the player's beliefs, so when one strategy is not strictly better than another strategy for both players, each strategy might be efficient or even both strategies might be efficient strategies (because it depends on the player's beliefs). On the other hand, for E_i^{pwo} and E_i^{DK}, the efficient strategy is fixed because it is independent of the player's initial

[21] In our study, we mainly discuss extensive-form games, thus in this section, we compare different definitions of efficient strategy in the sequential environment.

beliefs. Therefore, if there is a game in which $E_i^{Rabin} = E_i^{pwo} = E_i^{DK}$ after ruling out a set of strategies, then these strategies must lead both players to strictly worse payoffs. In Figure 8, for example, with $G_2 : (x, y) = (1/2, 1/2)$, then "hold" is a wasteful strategy, then either for E_i^{Rabin} or E_i^{pwo} or E_i^{DK}, only "send" is efficient.

5.1.4 Two reasons to redefine the efficient strategy

Reason 1. The theoretical and experimental reason

Rabin's definition of efficient strategy shows us limitations in some games, such as the trust game, which are discussed in D&K paper. D&K propose a new definition and apply it to some well-known games. It seems that their definition is appropriate for understanding reciprocity. However, for games such as in Figure 7 where $a = 0$ and $a = -2$, in Figure 8 where G_1: $(x, y) = (1/2, -1/2)$ and G_2: $(x, y) = (1/2, 1/2)$, and in Figure 3 and Figure 4: the ultimatum games with different alternatives, D&K's definition of efficient strategy cannot provide us with a psychological explanation as discussed above. However, if we have not obtained the correct efficient strategies, when we model kindness, the wasteful strategies will decrease the reference point. Thus, if there is a possibility that players reciprocate with less kind behaviour, then we might get a counter-intuitive result. This fact inspires us to find a more convincing rule to define the efficient strategy.

In our definition, we consider rational and psychological factors to explain the non-rational behaviour: PWO as a sequential rationality refinement. For example, applying our definition to G_1 and G_2, "hold" is not efficient for both games. The logic here is straightforward: no one is motivated to reward an unkind behaviour at a cost to themselves, and such a strategy should not influence our perception of kindness.

Reason 2. The psychological reason

Our main task in modelling kindness is to understand the decision maker's psychology and behaviour. For example, when one player behaves in a kind manner but only to avoid worse outcomes for themselves, how do we judge their kindness? In our view, we can first rule out psychologically implausible strategies (wasteful strategies). No one would like to reward unkind behaviour. Instead, they will respond selfishly, or punish if there is no cost, when facing unkind behaviour. Recall the games in Figure 3 and Figure 4 discussed in Section 3 and Section 4. We cannot find any intentions about the proposer in Figure 3 since the (8,2) offer is the only choice available to them. In Figure 4, it is easy to see that the best choice for the proposer is "O_1" because the responder must reject the alternative "O_2". Therefore, the proposer's choice "O_1" will still not include any intention. So these two games should be the same for the responder.

A player will not be perceived as kind when they simply do something that they are required to do. For example, employers paying their workers' salaries on time does not make them kind, and a child refraining from throwing stones at a neighbour's window does not make them kind, because in both cases they would suffer severe negative consequences if they took a different course of action.

5.1.5 The influence on equilibria

We propose a different theory of reciprocity compared with D&K's model. Clearly, their model and ours will produce different results. In this section, we explore the influence on equilibria. Because D&K's model excludes the concern of consequence part, we assume a special case where $\alpha = 0$ in the following discussion in order to compare the two models.

Proposition 3. *Suppose $\alpha = 0$ and $\vartheta(a_{j,h}) = 1/2$ in our model. Then the Dufwenberg and Kirchsteiger (2004) definition of efficient strategy always predicts a more positive reciprocity environment than ours.*

To see this, according to Proposition 1, we find that the minimum material payoff in D&K is less than or equal to ours. Moreover, notice that the definition of efficient strategy only influences the strategies that bring players a lower material payoff, but has no effect on the maximum material payoff. Thus, when we model intentional kindness, our model excludes some strategies that bring players a lower material payoff that are retained in D&K's model. The minimum material payoff further distinguishes the reference point between the two models. The reference point π_i^r is larger in our model than in D&K's model, and hence the intentional kindness term Ψ_i is smaller in our model than in D&K. Therefore, their definition perceives a player as more kind and predicts a more positive reciprocity environment.

This proposition also sheds light on the fundamental problem with D&K's definition. As we have claimed before, the definition of an efficient strategy in their model is not compatible with the intuitive understanding of reciprocity and trust, since they have not ruled out all wasteful strategies and have not provided a sufficiently psychologically convincing foundation.

5.2 Revisiting the sequential prisoner's dilemma and its variant

In the previous subsection, we have discussed the differences between the three definitions of efficient strategies. In this section, we also want to compare our model of the intentional kindness with the reciprocity models of Rabin and D&K. To better understand the following discussion, we assume that we have the same efficient strategies that satisfy Proposition 2. Since Rabin and D&K do not consider the consequential kindness of players, we assume a special case $\alpha = 0$ for the rest of this section.

In Rabin and D&K, players derive direct utility from their own material payoffs, and psychological utility from their and their opponent's kindness. In the most direct statement of modelling kindness, one's kindness is denoted by some rules (it always refers to the reference point). In their models, they pick out a specific form of the equitable payoff rule as the reference point (the equal division of the highest and the lowest payoffs). If one player chooses the action which implies the final material payoff of the decision maker is higher than the reference point, they are deemed kind to the decision maker. Otherwise, they are not kind (and are either neutral or unkind). In their settings, although many scholars have viewed their models as psychological game models, the decision maker only considers their own material payoffs. In other words, a player's kindness or unkindness is determined by the effect of their decisions on the decision maker, without taking into consideration how much they give up their own material payoff to help or hurt the decision maker. Our model does take this into account by introducing the intentional parameter ϑ. As we have introduced in Sections 1 and 3, one contribution of our model is that we incorporate the other player's status into the decision maker's utility and use it to influence their preference. As a result, our model is the only model that can predict the case when one player has the option to punish others at a cost to themselves (see Section 4.2).

As has been shown in many experimental studies, intentional kindness concerns can give rise to mutual cooperation in the prisoner's dilemma. When a player cares enough about the underlying intentions of other players, they prefer to cooperate if the first mover cooperated. In the rest of the paper, we study how the cooperation rate is influenced in the sequential prisoner's dilemma and its variant where the second mover is endowed with a chance to punish the first mover's defection.

		Second mover	
		Cooperate	Defect
First	Cooperate	a, b	c, d
mover	Defect	e, f	g, h

Table 2: General sequential prisoner's dilemma

The payoffs in the general sequential prisoner's dilemma are shown above: d and e denote temptation to defect and are the payoffs received by an individual defecting on a cooperator; a and b are for cooperation and stand for the material payoffs received by one of a pair of cooperators; g and h are for defection and stand for the payoffs received by a pair of defectors; and c and e are the sucker's payoffs received by a cooperator paired with a defector. The sequential prisoner's dilemma games also require that $e > g$, $a > c$, $d > b > h$ ("Cooperate" is kind and "Defect" is unkind), $g > c$ (efficient strategy), $f < h$, and $g + h < a + b$ (mutual cooperation can achieve a higher social surplus).

Proposition 4. *In the general sequential prisoner's dilemma, suppose that* $\alpha = 0$ *and* *that the first mover has chosen cooperation. Then the second mover is more likely to negatively reciprocate when:*

1. *The second mover will gain more if the first mover chooses defection (i.e. h is large).*
2. *The first mover will gain less by choosing defection (i.e. g is small).*

See proof in Appendix *A.3.1*.

This proposition reflects two tendencies. First, faced with kind behaviour in which the first mover chose cooperation, the second mover will infer the first mover's kindness by evaluating how much they could receive if the second mover chooses the alternative. If they can also gain enough material payoffs (i.e. h is not small), then the intentional kindness of the first mover is not so strong, thus the second mover will not be very reciprocal and will behave more selfishly. This is what Proposition 4(1) states, and the D&K model also reflects this observation.

Second, in addition to the second mover's own possible payoff, they will also infer the first mover's kindness by evaluating the extent to which the first mover gives up their own material payoffs in helping them. Proposition 4(2) tells us that if the first mover gains substantially less by defecting (i.e. g is not very large), then the intentional kindness of the first mover is not so strong, thus the second mover will not be very reciprocal and will behave more selfishly. There is no intention-based reciprocity model, to our knowledge, that can explain Proposition 4(2), which matches experimental findings.

Another case is that the second mover has the chance to punish the first mover's defection at a cost to themselves. We have already seen that punishment plays a significant role in influencing the second mover's cooperative behaviour (see Section 4.2). More generally, in addition to the above conditions, we add the punishment choice where the payoffs for defection of first mover and punishment of second mover are i and j, respectively. The punishment is considered costly when $i < g$ and $j < h$.

		Second mover		
		Cooperate	Defect	Punish
First	Cooperate	a, b	c, d	$-, -$
mover	Defect	e, f	g, h	i, j

Table 3: General sequential prisoner's dilemma with punishment

Proposition 5. *In the sequential prisoner's dilemma with punishment choice, suppose $\alpha = 0$, and conditional on the first mover's cooperation. Then the second mover is more likely to negatively reciprocate when the cost of punishment is reduced (i.e. j is increased) if:*

$$(h - j)(d - j) < \frac{e^c(d-b)(g-i)(d-h)(e^c+e^i)}{e^g(a-c)(e^c+e^g)}.$$

See proof in Appendix *A.3.2*.

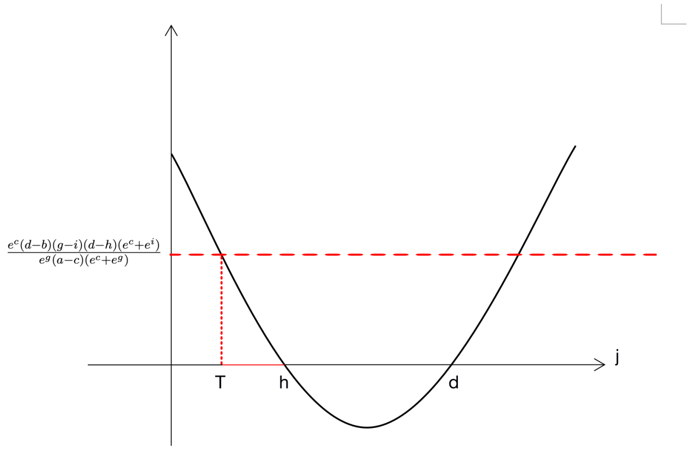

Figure 9: The effect of punishment

Another contribution of this paper is that our model successfully explains the power of punishment that many behavioural models fail to explain. Again, this proposition reflects two aspects. When the first mover chooses "Defect", the second mover has the chance to punish this unkind behaviour. On the one hand, if we have $j \in (T, h)$ (i.e. the cost is low), then when the cost is reduced (i.e. j decreases), the reference point will increase, since $\vartheta(C)$ increases and $\vartheta(D)$ decreases. Hence, the kindness of the first mover in the view of the second mover will decrease by playing "Cooperate" and the second mover is more likely to behave selfishly (by choosing "Defect"). For the possible explanation, because the cost of the punishment is very low, the second mover is highly likely to choose punishment as they do not need to sacrifice a large amount of material payoffs. In addition, if the cost reduces, then they will become more likely to choose punishment. Overall, from the perspective of the second mover, even though the first mover decides to cooperate, the intentional kindness is not very strong, and they may think that the purpose of first mover's cooperation is to avoid punishment rather than to be kind. The lower the cost is, the more the selfish response will appear.

On the other hand, if we have $j \in (0, T)$ (i.e. the cost of punishment is relatively large), then when the cost $h - j$ decreases, the reference point might decrease since $\vartheta(C)$ decreases and $\vartheta(D)$ increases. Hence, the kindness of the first mover in the view of the second mover will increase by playing "Cooperate" and the second mover is more likely to behave reciprocally. This inference might seem counterintuitive because it tells us that the power of punishment has a positive effect on cooperation[22]. To explain this,

[22]We cannot also provide the exact threshold of when they act more reciprocally if the cost of punishment is decreasing, but according to the proof in Appendix *A.3.2*, it shows us a probability of this inference.

we first note that the cost of punishment is now relatively large. On the one hand, the second mover is less likely to punish the first mover as this comes at a greater cost. On the other hand, if the cost decreases, then the second mover will be more likely to punish the first mover, but this still hurts the second mover. In the mind of the second mover, they know the first mover knows that they are more likely to choose punishment if the cost is decreasing, so in order to discourage the second mover from choosing the worse result (i, j), the second mover decides to cooperate. Therefore, even though the cost of punishment is decreasing, the intentional kindness is stronger when the cost is relatively large. The second mover may think the purpose of the first mover's cooperation is to avoid the costly punishment. This inference, for policy making, is useful to yield improved mutual cooperation rates if the policymakers know why people cooperate, that is, policymakers can find a balanced cost of punishment to achieve mutual cooperation, which suggests the first mover might cooperate because of the threat from the punishment and the second mover might also cooperate because punishment is costly.

This proposition is very intuitive. The costly punishment has two sides. First, it has negative effects, which reduce the strength of the intentional kindness. When others choose cooperative behaviour, the decision maker may have an excuse: their cooperation only aims to avoid punishment, and they do not show any intentional kindness to the other player. In another interesting case, costly punishment has positive effects as well. The decision maker may consider that the chance of punishment is a way to hurt both parties, by not giving the decision maker the chance to hurt both might be more kind.

6 Conclusion

In this paper, we propose an economic model of reciprocity that aims to explain a decision maker's reciprocal behaviour by introducing new definitions of efficient strategy, intentional kindness, and consequential kindness.

Our model's definition of efficient strategy successfully solves the paradox in the trust game (Isoni and Sugden, 2019) and successfully explains the results of a range of experimental studies (Falk, Fehr, and Fischbacher, 2003). We also compare our definition with two well-known definitions of efficient strategy and list two reasons why our definition is superior to the belief-dependent efficient strategy of Rabin and the belief-independent efficient strategy of D&K. Then, we split kindness into two parts: intentional kindness and consequential kindness. Our definition of intentional kindness (two-sided) is the first model that takes into account the status of both the decision maker and other player. We argue that when a decision maker evaluates the intentional kindness of other players, in addition to their own payoffs, they also consider the extent to which other players sacrifice their payoffs in help or hurt them. This is also what experimental studies suggest (Falk, Fehr, and Fischbacher, 2003; Falk, Fehr, and Fischbacher, 2008; Orhun, 2018; Ahn et al., 2007). We incorporate this idea in the form of the intention function $\vartheta(a_{j,h})$. With respect to consequential kindness, it differs from the consequence-based models like Fehr and Schmidt (1999) and Bolton and Ockenfels (2000). We assume that players only care about their own inequity and not inequity experienced by others. In particular, they do not experience disutility when receiving more than others. Then we finish our model and develop the concept of ERE and prove its existence.

We also apply our model to some famous experimental games such as the ultimatum game, the sequential prisoner's dilemma with/without punishment, and the sequential prisoner's dilemma with asymmetric payoffs. Our model is in line with experimental findings on these experimental games. Moreover, demonstrate that our model performs better than those of Rabin and D&K. Finally, our model explains why the cooperation rate can fall even though another player chooses cooperative behaviour; the decision maker may have an excuse: their cooperation only aims to avoid a worse result, and they do not show any intentional kindness to others (Proposition 4(1)) or their cooperation only helps others by a small amount (Proposition 4(2)). More importantly, our model also explains why sometimes the decision maker's positive reciprocity falls when they have the opportunity to punish other players at a cost to themselves: the chance of punishment can be viewed as the power for the decision maker (Proposition 5). It also explains why sometimes the decision maker's negative reciprocity falls when they have the chance to punish other players: the cost of punishment is very large, the player may think others choose cooperation aiming to help them instead of letting them get less since punishment is very costly.

According to our theoretical framework, there are several possible further directions can be studied. First, we could focus on a more general case, with more players and more stages in the game. The consideration of the player must be more complicated if they need to interact with more people and play more stages. For example, we could consider whether the updating rules still work in this case (Jiang and Wu, 2019). Moreover, reciprocity under uncertainty is also important since players often have limited information about their opponents (Sohn and Wu, 2022). Finally, we could test whether our model produces correct predictions in additional experimental settings.

Appendix

A.1. proof of the theorem

Let us define the local best response correspondence and the best response correspondence $f_{i,h} : S \rightarrow S_{i,h}$ and $f : S \rightarrow \prod_{(i,h)\in N\times H} S_{i,h}$ by:

$$f_{i,h}(a) = \underset{a_i \in A_{i,h}}{\arg\max} \quad U_i(a_i, B_{ij}(h), C_{iji}(h))$$
$$f(a) = \prod_{(i,h)\in N\times H} f_{i,h}(a)$$

Because $\prod_{(i,h)\in N\times H} S_{i,h}$ and S are topologically equivalent, so f is equivalent to a correspondence $\gamma : S \rightarrow S$, then we see whether there exist a fixed point (a fixed point under γ satisfies the ERE conditions (Definition 9)). To find this, notice that $f_{i,h}$ satisfies condition (i) of Definition 9, and the first and second order beliefs are correct (condition (ii) and (iii) of Definition 9). Thus $f_{i,h}$ finds the optimal strategies of player i at h (consistent with (i) of Definition 9). f and γ are combined best-response correspondences. As γ is a correspondence from S to S, it can be applied to the Kakutani's fixed point theorem.

Let's check if our conditions satisfy all conditions of Kakutani's fixed point theorem. Firstly, notice that $f_{i,h}$ is non-empty, compact, upper hemi-continuous and convex (because $A_{i,h}$ is non-empty, compact, U_i is continuous: as $\pi_i, \pi_j, \Psi_i, \Phi_i$ are all continuous and U_i is linear in i's own choice). Since these properties extend from $f_{i,h}$ to f and γ, all conditions of Kakutani's fixed point theorem are satisfied, thus admitting a fixed point. Therefore, the ERE must exist.

A.2. Applications

A.2.1. Four mini-ultimatum games

First, by Definition 2 and Definition 3, we conclude that both actions "O_1" and "O_2" are efficient strategies in the first three games, and only "O_2" is the efficient strategy in game (d) (Result 1).

We then move to the consequential kindness Φ_i, according to the Definition 7, for these four games, if the proposer chooses "O_1" and the responder chooses "y", then $\Phi_R(y, O_1) = \alpha \cdot \{[2-8]^- - 0\}^- = -6\alpha$. If the proposer chooses "O_1" and the responder chooses "n", then $\Phi_R(n, O_1) = \alpha \cdot \{[0-0]^- - [2-8]^-\}^- = 0$. This suggests that the responder rejects the (8,2) offer is that they get inequity for accepting it.

We next consider the intentional kindness Φ_R when the proposer chooses "O_2" in (a) and (b). We find that the responder won't experience the inequity no matter what they choose "y" or "n" as they can earn equal or more material payoffs than the proposer, so $\Phi_R(O_2) = 0$. Turning to the intentional kindness, the proposer chooses "y" after "O_2" can earn the highest material payoff in the whole game. Overall, the responder selects "y" after "O_2" is the best response (Result 2).

Then we go back to our central discussion of the responder's response conditional on "O_1". For simplicity, in these four games, we set the responder's belief about the proposer's belief about his choice "y" after "O_1" is q (responder's second-order belief).

In the ultimatum game (a), we obtain the responder's utility given proposer chooses "O_1":

$$U_R(y, O_1) = 2 - 6\alpha + 8\beta \cdot \delta_{PR}$$
$$U_R(n, O_1) = 0 + 0\alpha + 0\beta \cdot \delta_{PR}$$

where $\delta_{PR} = 2q - [\frac{e^{8q+0(1-q)}}{e^{8q+0(1-q)}+e^5} \cdot 2q + \frac{e^5}{e^{8q+0(1-q)}+e^5} \cdot 5] = \frac{e^5}{e^{8q+0(1-q)}+e^5} \cdot (2q-5)$. Note that here δ_{PR} is less than 0 since $q \in [0,1]$.

Reading two equations, we find that the responder will never accept the offer if $\alpha \geq 1/3$ as $U_R(y, O_1) < U_R(n, O_1)$ (Result 3).

When having $\alpha < 1/3$, the responder would like to choose "y" if and only if $U_R(y, O_1) > U_R(n, O_1)$. Because all beliefs are correct in every ERE (Definition 9 (ii) and (iii)), q=1 must hold when $U_R(y, O_1) > U_R(n, O_1)$. Therefore:

$$2 - 6\alpha + 8\beta \cdot \delta_{PR} > 0 + 0\alpha + 0\beta \cdot \delta_{PR}$$
$$2 - 3\beta \cdot \frac{e^5}{e^8 + e^5} \cdot 8 - 6\alpha > 0$$
$$\beta < \frac{(2-6\alpha)(e^8+e^5)}{24e^5} \tag{1}$$

Similarly, the responder would like to choose "n" if and only if $U_R(y, O_1) < U_R(n, O_1)$ and q=0, Therefore:

$$2 - 6\alpha + 8\beta \cdot \delta_{PR} < 0 + 0\alpha + 0\beta \cdot \delta_{PR}$$
$$2 - 6\alpha + 8\beta \cdot \frac{-5e^5}{1+e^5} < 0$$
$$\beta > \frac{(2-6\alpha)(1+e^5)}{40e^5} \tag{2}$$

According to above two thresholds, we find that $\frac{(2-6\alpha)(e^8+e^5)}{24e^5} > \frac{(2-6\alpha)(1+e^5)}{40e^5}$. It means that when $\beta \in [\frac{(2-6\alpha)(1+e^5)}{40e^5}, \frac{(2-6\alpha)(e^8+e^5)}{24e^5}]$, there is no difference for the proposer to choose either "y" or "n". Therefore, the proposer will choose randomly. For $\beta \notin [\frac{(2-6\alpha)(1+e^5)}{40e^5}, \frac{(2-6\alpha)(e^8+e^5)}{24e^5}]$, we revise our (1) and (2) as follows (Result 4):

The responder would like to choose "y" when:

$$\beta < \frac{(2-6\alpha)(1+e^5)}{40e^5} \tag{3}$$

The responder would like to choose "n" when:

$$\beta > \frac{(2-6\alpha)(e^8+e^5)}{24e^5} \tag{4}$$

Similarly, in the ultimatum game (b), we can obtain the responder's utility given proposer chooses "O_1'':

$$U_R(y, O_1) = 2 - 6\alpha + 8\beta \cdot \delta_{PR}$$

$$U_R(n, O_1) = 0 + 0\alpha + 0\beta \cdot \delta_{PR}$$

where $\delta_{PR} = 2q - [\frac{e^{8q+0(1-q)}}{e^{8q+0(1-q)}+e^2} \cdot 2q + \frac{e^2}{e^{8q+0(1-q)}+e^2} \cdot 8] = \frac{e^2}{e^{8q+0(1-q)}+e^2} \cdot (2q-8)$. We find that δ_{PR} is less than 0 since $q \in [0, 1]$.

We discuss α first, and we find that the responder will never accept the offer if $\alpha \geq 1/3$ as $U_R(y, O_1) < U_R(n, O_1)$ (Result 3).

What if $\alpha < 1/3$? Then the responder would like to choose "y" if $U_R(y, O_1) > U_R(n, O_1)$. We derive q=1 by Definition 9, and we get:

$$2 - 6\alpha + 8\beta \cdot \delta_{PR} > 0 + 0\alpha + 0\beta \cdot \delta_{PR}$$

$$2 - 6\beta \cdot \frac{e^2}{e^8+e^2} \cdot 8 - 6\alpha > 0$$

$$\beta < \frac{(2-6\alpha)(e^8+e^2)}{48e^2} \tag{5}$$

The responder would like to choose "n" if $U_R(y, O_1) < U_R(n, O_1)$. Now q=0, and we then have the following:

$$2 - 6\alpha + 8\beta \cdot \delta_{PR} < 0 + 0\alpha + 0\beta \cdot \delta_{PR}$$

$$2 - 6\alpha + 8\beta \cdot \frac{-8e^5}{1+e^5} < 0$$

$$\beta > \frac{(2-6\alpha)(1+e^2)}{64e^2} \tag{6}$$

Looking at (5) and (6), we find that $\frac{(2-6\alpha)(e^8+e^2)}{48e^2} > \frac{(2-6\alpha)(1+e^2)}{64e^2}$. It means that when $\beta \in [\frac{(2-6\alpha)(1+e^2)}{64e^2}, \frac{(2-6\alpha)(e^8+e^2)}{48e^2}]$, there is no difference for the proposer to choose either "y" or "n". Therefore, the proposer will choose randomly. For $\beta \notin [\frac{(2-6\alpha)(1+e^2)}{64e^2}, \frac{(2-6\alpha)(e^8+e^2)}{48e^2}]$, we revise (5) and (6) as follows (Result 5):

The responder would like to choose "y" when:

$$\beta < \frac{(2-6\alpha)(1+e^2)}{64e^2} \tag{7}$$

The responder would like to choose "n" when:

$$\beta > \frac{(2-6\alpha)(e^8+e^2)}{48e^2} \tag{8}$$

We have obtained the different thresholds of choosing "y" in games (a) and (b) when $\alpha < 1/3$. To compare the "acccept" rate and "reject" rate in game (a) and game (b), we can compare (3) and (7), (4) and (8).

Comparing two different thresholds in (3) and (7), $\frac{(2-6\alpha)(1+e^2)}{64e^2} - \frac{(2-6\alpha)(1+e^5)}{40e^5} = (2-6\alpha)(\frac{1}{64} - \frac{1}{40} + \frac{1}{64e^2} - \frac{1}{60e^5}) < 0$. Comparing two different thresholds in (4) and (8) $\frac{(2-6\alpha)(e^8+e^2)}{48e^2} - \frac{(2-6\alpha)(e^8+e^5)}{24e^5} = (2-6\alpha)\frac{(e^6-2e^3-1)}{48} > 0$. It tells that if we don't consider the middle interval, then for the same distribution over the population, more proposers would like to choose "n" in game (a) than in game (b) by (4) and (8), and more proposers would like to choose "y" in game (a) than in game (b) by (3) and (7). Therefore, we should also consider whether more responders accept the offer than reject the offer in game (b) in middle interval compared with game (a). So we have $(\frac{(2-6\alpha)(e^8+e^2)}{48e^2} - \frac{(2-6\alpha)(e^8+e^5)}{24e^5}) - (\frac{(2-6\alpha)(1+e^5)}{40e^5} - \frac{(2-6\alpha)(1+e^2)}{64e^2}) = \frac{(2-6\alpha)}{960e^2}(20e^8 - 40e^5 - 29e^2 + 15 - 24e^{-3})$ which is greater than 0. It means that for the uniform distribution over the population, less proposers would like to choose "n" in game (b) than in game (a) (Result 6).

In the ultimatum game (c), there does not exist any intention for the proposer, because he can not choose any intentional action, thus the intention term $\Psi_R = 0$.

In the ultimatum game (d), because "O_1" is the only efficient strategy, there does not exist any intention for the proposer, so the intention term $\Psi_R = 0$.

Hence, in games (c) and (d), we only consider the consequence term $\Phi_R = -6\alpha$ for choosing "y" and $\Phi_R = 0$ for choosing "n", the utility of proposer:

$$U_R(y, O_1) = 2 - 6\alpha$$

$$U_R(n, O_1) = 0 + 0$$

Therefore, the responder chooses "y" after "O_1" if $U_R(y, O_1) > U_R(n, O_1)$, that is, $\alpha < 1/3$; and the responder chooses "n" after "O_1" if $U_R(y, O_1) < U_R(n, O_1)$, that is, $\alpha > 1/3$; and roposer would like to choose randomly after "O_1" if $U_R(y, O_1) = U_R(n, O_1)$, that is, $\alpha = 1/3$ (Result 7).

Moreover, since the intention terms in games (a) and (b) are always negative (unkind "O_1"), thus for the same distribution over the population, more proposers would like to choose "y" in game (c) and game (d) than in game (a) and (b) (Result 8).

A.2.2. The sequential prisoner's dilemma with/without punishment

A.2.2.1. General sequential prisoner's dilemma

For simplicity, in this section, we first set player 2's belief about player 1's belief about his choice "c" after "C" is p. According to our model, we obtain player 2's utility given player 1 chose "C":

$$U_2(c,C) = 1.5 + \alpha\{[1.5-5.5]^- - [2-3]^-\}^- + 5.5\beta \cdot \delta_{12}$$
$$U_2(d,C) = 2 + \alpha\{[2-3]^- - [1.5-5.5]^-\}^- + 3\beta \cdot \delta_{12}$$

where $\delta_{12} = [p \cdot 1.5 + (1-p) \cdot 2] - [\frac{e^{5.5p+3(1-p)}}{e^{5.5p+3(1-p)}+e^4}(p \cdot 1.5 + (1-p) \cdot 2) + \frac{e^4}{e^{5.5p+3(1-p)}+e^4} \cdot 1] = (p \cdot 1.5 + (1-p) \cdot 2 - 1)\frac{e^4}{e^{5.5p+3(1-p)}+e^4}$.

The player 2 would like to choose "c" if $U_2(c,C) > U_2(d,C)$. According to Definition 9, we get p=1, and we then have:

$$1.5 - 3\alpha + 5.5\beta \cdot \delta_{12} > 2 + 3\beta \cdot \delta_{12}$$
$$2 + 12\alpha - 5\beta\frac{e^4}{e^{5.5}+e^4} < 0$$
$$\beta > \frac{(2+12\alpha)(e^{5.5}+e^4))}{5e^4} \tag{9}$$

The player 2 will choose "d" if $U_2(d,C) > U_2(c,C)$. We have p=0 and:

$$1.5 - 3\alpha + 5.5\beta \cdot \delta_{12} < 2 + 3\beta \cdot \delta_{12}$$
$$2 + 12\alpha - 10\beta\frac{e^4}{e^3+e^4} > 0$$
$$\beta < \frac{(2+12\alpha)(e^3+e^4)}{10e^4} \tag{10}$$

When $\frac{(2+12\alpha)(e^3+e^4)}{10e^4} \leq \beta \leq \frac{(2+12\alpha)(e^{5.5}+e^4)}{5e^4}$. Player 2 would like to choose "c" with probability p if $U_2(c,C) = U_2(d,C)$. Because the second-order belief must be correct by Definition 9, so the probability that satisfies:

$$1.5 - 3\alpha + 5.5\beta \cdot \delta_{12} = 2 + 3\beta \cdot \delta_{12}$$
$$2 + 12\alpha = 5\beta(2-p)\frac{e^4}{e^{5.5p+3(1-p)}+e^4} \tag{11}$$

A.2.2.2. Sequential prisoner's dilemma with punishment

When the player 2 has the chance to punish player 1's defection, the process would be slightly different as the defection is not the unique ERE given player defects. So in this section, we set the player 2's belief about player 1's belief about his choice "c" after "C" is p. In addition, we set the player 2's belief about player 1's belief about his choice "p" after "D" is q.

First, consider that if player 1 decides "D", we can obtain player 2's utility as follows:

$$U_2(p,D) = 0.5 + \alpha\{[0.5-1.5]^- - [1-4]^-\}^- + 1.5\beta \cdot \delta_{12}$$

$$U_2(d, D) = 1 + \alpha\{[1-4]^- - [0.5-1.5]^-\}^- + 4\beta \cdot \delta_{12}$$

where $\delta_{12} = [q \cdot 0.5 + (1-q) \cdot 1] - [\frac{e^{5.5p+3(1-p)}}{e^{5.5p+3(1-p)}+e^{1.5q+4(1-q)}}(p \cdot 1.5 + (1-p) \cdot 2) + \frac{e^{1.5q+4(1-q)}}{e^{5.5p+3(1-p)}+e^{1.5q+4(1-q)}}(q \cdot 0.5 + (1-q) \cdot 1)] = (q \cdot 0.5 + (1-q) \cdot 1 - p \cdot 1.5 - (1-p) \cdot 2)\frac{e^{5.5p+3(1-p)}}{e^{5.5p+3(1-p)}+e^{1.5q+4(1-q)}}$. Because $q, p \in [0, 1]$, $\delta_{12} < 0$ is satisfied.

So now player 2 would like to choose "d" if $U_2(p, D) < U_2(d, D)$. To verify Result 11, we need to prove that $U_2(p, D) < U_2(d, D)$ is not always held. When choosing "d", we must have p=0, so:

$$0.5 + 1.5\beta \cdot \delta_{12} < 1 - 2\alpha + 4\beta \cdot \delta_{12}$$

$$-2.5\beta \cdot \delta_{12} < 0.5 - 2\alpha \tag{12}$$

Since δ_{12} is less than 0 and β is no less than 0, so we can get $-2.5\beta \cdot \delta_{12} \geq 0$. Furthermore, we know $\alpha \geq 0$. It suggests that if $\alpha > 0.25$, the equation (1) will never be satisfied. Therefore, if player 1 defects (by choosing "D'), defection (by choosing "d") for player 2 is not the unique equilibrium (Result 11).

Now, move to our main discussion of the player 2's behaviour given player cooperates. We then obtain player 2's utility given player 1 chooses "C":

$$U_2(c, C) = 1.5 + \alpha\{[1.5-5.5]^- - [2-3]^-\}^- + 5.5\beta \cdot \delta_{12}$$

$$U_2(d, C) = 2 + \alpha\{[2-3]^- - [1.5-5.5]^-\}^- + 3\beta \cdot \delta_{12}$$

where $\delta_{12} = [p \cdot 1.5 + (1-p) \cdot 2] - [\frac{e^{5.5p+3(1-p)}}{e^{5.5p+3(1-p)}+e^{1.5q+4(1-q)}}(p \cdot 1.5 + (1-p) \cdot 2) + (q \cdot 0.5 + (1-q))\frac{e^{1.5q+4(1-q)}}{e^{5.5p+3(1-p)}+e^{1.5q+4(1-q)}}] = (p \cdot 1.5 + (1-p) \cdot 2 - q \cdot 0.5 - (1-q))\frac{e^{1.5q+4(1-q)}}{e^{5.5p+3(1-p)}+e^{1.5q+4(1-q)}}$.

The player 2 would like to choose "c" if $U_2(c, C) > U_2(d, C)$. We get p=1 by Definition 9, and we have:

$$1.5 - 3\alpha + 5.5\beta \cdot \delta_{12} > 2 + 3\beta \cdot \delta_{12}$$

$$2 + 12\alpha - 5\beta(1+q)\frac{e^{1.5q+4(1-q)}}{e^{5.5} + e^{1.5q+4(1-q)}} < 0$$

$$\beta > \frac{(2+12\alpha)(e^{5.5} + e^{1.5q+4(1-q)})}{5(1+q)e^{1.5q+4(1-q)}} \tag{13}$$

The player 2 would like to choose "d" if $U_2(d, C) > U_2(c, C)$. Now p=0, and we have:

$$1.5 - 3\alpha + 5.5\beta \cdot \delta_{12} < 2 + 3\beta \cdot \delta_{12}$$

$$1 + 6\alpha - 5\beta(1+0.5q)\frac{e^{1.5q+4(1-q)}}{e^{3} + e^{1.5q+4(1-q)}} > 0$$

$$\beta < \frac{(2+12\alpha)(e^{3} + e^{1.5q+4(1-q)})}{5(2+q)e^{1.5q+4(1-q)}} \tag{14}$$

When $\frac{(2+12\alpha)(e^3+e^{1.5q+4(1-q)})}{5(2+q)e^{1.5q+4(1-q)}} \leq \beta \leq \frac{(2+12\alpha)(e^{5.5}+e^{1.5q+4(1-q)})}{5(1+q)e^{1.5q+4(1-q)}}$. The player 2 would like to choose "c" with probability p if $U_2(c, C) = U_2(d, C)$. Because the second-order belief must be correct by Definition 9, we have:

$$1.5 - 3\alpha + 5.5\beta \cdot \delta_{12} = 2 + 3\beta \cdot \delta_{12}$$

$$2 + 12\alpha = 5\beta(2 - p + q)\frac{e^{1.5q+4(1-q)}}{e^{5.5p+3(1-p)} + e^{1.5q+4(1-q)}} \tag{15}$$

A.2.3. Prisoner's dilemma with asymmetric payoffs

A.2.3.1. Treatment 1

In this section, the analysis is very similar compared with last section. For the sake of simplicity, in this section, we still set the player 2's belief about player 1's belief about his choice "c" after "C" is p. The proof of Result 14 is the same as the sequential prisoner's dilemma without punishment, and both "C" and "D" are efficient strategies and player 2 will always defect (by choosing "d") as a response in every ERE.

We obtain player 2's utility given player 1 chooses "C":

$$U_2(c, C) = 30 + \alpha\{[30 - 30]^- - [36 - 12]^-\}^- + 30\beta \cdot \delta_{12}$$

$$U_2(d, C) = 36 + \alpha\{[36 - 12]^- - [30 - 30]^-\}^- + 12\beta \cdot \delta_{12}$$

where $\delta_{12} = [p\cdot 30+(1-p)\cdot 36]-[\frac{e^{30p+12(1-p)}}{e^{30p+12(1-p)}+e^{18}}(p\cdot 30+(1-p)\cdot 36)+\frac{e^{18}}{e^{30p+12(1-p)}+e^{18}}\cdot 18] = (p \cdot 30 + (1-p) \cdot 36 - 18)\frac{e^{18}}{e^{30p+12(1-p)}+e^{18}}$.

The player 2 would like to choose "c" if $U_2(c, C) > U_2(d, C)$. We have p=1 by Definition 9, and we can derive that:

$$\beta > \frac{e^{30} + e^{18}}{36e^{18}} \tag{16}$$

The player 2 would like to choose "d" if $U_2(d, C) > U_2(c, C)$. We have p=0 by Definition 9, and we can derive that:

$$\beta < \frac{e^{12} + e^{18}}{54e^{18}} \tag{17}$$

When $\frac{e^{12}+e^{18}}{54e^{18}} \leq \beta \leq \frac{e^{30}+e^{18}}{36e^{18}}$. The player 2 would like to choose "c" with probability p if $U_2(c, C) = U_2(d, C)$. Because the second-order belief must be correct by Definition 9, now we have:

$$3\beta(18 - 6p) \cdot \frac{e^{18}}{e^{18p+12} + e^{18}} = 1 \tag{18}$$

A.2.3.2. Treatment 2

Similarly, in treatment 2, we also get player 2's utility given player 1 chooses "C":

$$U_2(c, C) = 26 + \alpha\{[26 - 34]^- - [34 - 14]^-\}^- + 34\beta \cdot \delta_{12}$$

$$U_2(d, C) = 34 + \alpha\{[34 - 14]^- - [26 - 34]^-\}^- + 14\beta \cdot \delta_{12}$$

where $\delta_{12} = [p \cdot 26 + (1-p) \cdot 34] - [\frac{e^{34p+14(1-p)}}{e^{34p+14(1-p)}+e^{18}}(p \cdot 26 + (1-p) \cdot 34) + \frac{e^{18}}{e^{34p+14(1-p)}+e^{18}} \cdot 18] = (p \cdot 26 + (1-p) \cdot 34 - 18)\frac{e^{18}}{e^{34p+14(1-p)}+e^{18}}$.

The player 2 would like to choose "c" if $U_2(c, C) > U_2(d, C)$. We have p=1 by Definition 9, and we can derive that:

$$\beta > \frac{(e^{34} + e^{18})(\alpha + 1)}{20e^{18}} \tag{19}$$

The player 2 would like to choose "d" if $U_2(d, C) > U_2(c, C)$. Now p=0, and we have:

$$\beta < \frac{(e^{14} + e^{18})(1 + \alpha)}{40e^{18}} \tag{20}$$

When $\frac{(e^{14}+e^{18})(1+\alpha)}{40e^{18}} \leq \beta \leq \frac{(e^{34}+e^{18})(\alpha+1)}{20e^{18}}$. The player 2 would like to choose "c" with probability p if $U_2(c, C) = U_2(d, C)$. Because the second-order belief must be correct by Definition 9, now we have:

$$20\beta(2 - p) \cdot \frac{e^{18}}{e^{20p+14} + e^{18}} = 1 + \alpha \tag{21}$$

A.2.3.3. Treatment 3

Similarly, in treatment 3, we also get player 2's utility given player 1 chooses "C":

$$U_2(c, C) = 34 + \alpha\{[34 - 26]^- - [38 - 10]^-\}^- + 26\beta \cdot \delta_{12}$$

$$U_2(d, C) = 38 + \alpha\{[38 - 10]^- - [34 - 26]^-\}^- + 10\beta \cdot \delta_{12}$$

where $\delta_{12} = [p \cdot 34 + (1-p) \cdot 38] - [\frac{e^{26p+10(1-p)}}{e^{26p+10(1-p)}+e^{18}}(p \cdot 34 + (1-p) \cdot 38) + \frac{e^{18}}{e^{26p+10(1-p)}+e^{18}} \cdot 18] = (p \cdot 34 + (1-p) \cdot 38 - 18)\frac{e^{18}}{e^{26p+10(1-p)}+e^{18}}$.

The player 2 would like to choose "c" if $U_2(c, C) > U_2(d, C)$. We have p=1 by Definition 9, and we can derive that:

$$\beta > \frac{e^{26} + e^{18}}{64e^{18}} \tag{22}$$

The player 2 would like to choose "d" if $U_2(d,C) > U_2(c,C)$. Now p=0, and we have:

$$\beta < \frac{e^{10}+e^{18}}{80e^{18}} \tag{23}$$

When $\frac{e^{10}+e^{18}}{80e^{18}} \leq \beta \leq \frac{e^{26}+e^{18}}{64e^{18}}$. The player 2 would like to choose "c" with probability p if $U_2(c,C) = U_2(d,C)$. Because the second-order belief must be correct by Definition 9, now we have:

$$4\beta(20-4p)\cdot\frac{e^{18}}{e^{16p+10}+e^{18}} = 1 \tag{24}$$

A.2.3.4. Comparison

First, comparing when player 2 cooperates (by choosing "c") conditional on player 1 cooperates (by choosing "C"). We have three different thresholds in three treatments:

$$\beta > \frac{e^{30}+e^{18}}{36e^{18}} \ (treatment\ 1)$$
$$\beta > \frac{(e^{34}+e^{18})(\alpha+1)}{20e^{18}} \ (treatment\ 2)$$
$$\beta > \frac{e^{26}+e^{18}}{64e^{18}} \ (treatment\ 3)$$

In these three treatments, we find that $\frac{e^{26}+e^{18}}{64e^{18}} < \frac{e^{30}+e^{18}}{36e^{18}}$. Therefore, for the same distribution over the population, a higher fraction of player 2 would like to choose "c" in treatment 3 than in treatment 1. We also notice that $\frac{e^{30}+e^{18}}{36e^{18}} < \frac{(e^{34}+e^{18})(\alpha+1)}{20e^{18}}$, since α is the non-negative value, even we set $\alpha = 0$, $\frac{e^{30}+e^{18}}{36e^{18}} < \frac{e^{34}+e^{18}}{20e^{18}}$ is satisfied. Again, for the same distribution over the population, a higher fraction of player 2 would like to choose "c" in treatment 1 than in treatment 2.

Then comparing when player 2 chooses "d" conditional on player 1 chooses "C". We also have three different thresholds in three treatments:

$$\beta < \frac{e^{12}+e^{18}}{54e^{18}} \ (treatment\ 1)$$
$$\beta < \frac{(e^{14}+e^{18})(1+\alpha)}{40e^{18}} \ (treatment\ 2)$$
$$\beta < \frac{e^{10}+e^{18}}{80e^{18}} \ (treatment\ 3)$$

So we obtain different thresholds of when player 2 defects given different different treatments. We find that $\frac{(e^{14}+e^{18})(1+\alpha)}{40e^{18}} > \frac{e^{12}+e^{18}}{54e^{18}} > \frac{e^{10}+e^{18}}{80e^{18}}$. It means that for the same distribution over the population, a higher fraction of player 2 would like to choose "d" in treatment 2 than in treatment 1 and 3, and a higher fraction of player 2 would like to choose "d" in treatment 1 than in treatment 3.

Finally, moving to a more complicated case when player 2 chooses the mixed strategy. We have set that player 2 will choose to cooperate with probability p given player 1 cooperates. We have shown that with different m, we have different cooperation rates p:

$$3\beta(18-6p)\cdot\frac{e^{18}}{e^{18p+12}+e^{18}} = 1\ (treatment\ 1)$$
$$20\beta(2-p)\cdot\frac{e^{18}}{e^{20p+14}+e^{18}} = 1+\alpha\ (treatment\ 2)$$
$$4\beta(20-4p)\cdot\frac{e^{18}}{e^{16p+10}+e^{18}} = 1\ (treatment\ 3)$$

We have cooperation rate p satisfies the above three equations given fixed β. For the same β and α, what we want to see is how the value of p in three treatments changes. Looking at treatment 3 and treatment 1 first, we set $M = \frac{1}{2\beta} = \frac{9(3-p_1)}{e^{18p_1-6}+1} = \frac{8(5-p_3)}{e^{16p_3-8}+1}$. It is not hard find that M will decrease if p_1 and p_2 increase. Besides, if $p_1 = p_3$, we find that $\frac{8(5-p_3)}{e^{16p_3-8}+1} - \frac{9(3-p_1)}{e^{18p_1-6}+1} > 0$ since $8(5-p_3) - 9(3-p_1) = 13 + p_1 > 0$ and $(e^{16p_3-8}+1) - (e^{18p_1-6}+1) = e^{16p_3-8} - e^{18p_1-6} < 0$ (because $(16p_3-8) - (18p_1-6) = -2p_1 - 2 < 0$) and $p \in [0, 1]$. Therefore, p_3 is greater the p_1 for fixed β. This suggests that for the same person (M is fixed), he is more willing to cooperate in treatment 3 than in treatment 1.

Similarly, turning to treatment 1 and treatment 2. Notice we have α in treatment 2, the value of p is smaller when $\alpha > 0$ compared to $\alpha = 0$. Our purpose in this part is to prove for the same β p in treatment 1 is larger than in treatment 2. So we can set $\alpha = 0$ and $M' = \frac{1}{2\beta} = \frac{9(3-p_1)}{e^{18p_1-6}+1} = \frac{10(2-p_2)}{e^{20p_2-4}+1}$. It is not hard find that M' will decrease if p_1 and p_3 increase. If $p_1 = p_2$, we find that $\frac{9(3-p_1)}{e^{18p_1-6}+1} - \frac{10(2-p_2)}{e^{20p_2-4}+1} > 0$ since $9(3-p_1) - 10(2-p_2) = 7 + p_1 > 0$ and $(e^{18p_1-6}+1) - (e^{20p_2-4}+1) = e^{18p_1-6} - e^{20p_2-4} < 0$ (because $(18p_1-6) - (20p_2-4) = -2p_1 - 2 < 0$) and $p \in [0, 1]$. Therefore, p_1 is greater the p_2 for fixed β. This suggests that for the same person (M is fixed), he is more willing to cooperate in treatment 1 than treatment 3.

To conclude these three cases, more player 2s would like to cooperate given player 1 cooperates in treatment 3 than treatment 1, and more player 2s would like to cooperate given player 1 cooperates in treatment 1 than treatment 2. Our prediction is also consistent with the experimental results.

A.3. Discussion

A.3.1 proof of Proposition 4

First of all, since $\{b, d\} > \{h\}$ and $f < h$, the second mover must choose "d" after "D" ("D" is unkind and our model suggests that no one is motivated to reward the unkind behaviour).

We set p=p(c,C). We obtain the second mover's utility given the first mover chooses "C":

$$U_2(c, C) = b + a \cdot \beta \cdot \delta_{12}$$
$$U_2(d, C) = d + c \cdot \beta \cdot \delta_{12}$$

where $\delta_{12} = bp+d(1-p)-\{\frac{e^{ap+c(1-p)}}{e^{ap+c(1-p)}+e^g}[bp+d(1-p)]+\frac{e^g}{e^{ap+c(1-p)}+e^g}(h)\} = \frac{e^g}{e^{ap+c(1-p)}+e^g}[bp+d(1-p)-h]$.

So now the second mover would like to choose "c" if $U_2(c, C) > U_2(d, C)$ and we have p=1. So:

$$\beta > \frac{(d-b)(e^a + e^g)}{(a-c) \cdot e^g \cdot (b-h)} \tag{25}$$

The second mover would like to choose "d" if $U_2(d, C) > U_2(c, C)$. Now p=0, and we have:

$$\beta < \frac{(d-b)(e^c + e^g)}{(a-c) \cdot e^g \cdot (d-h)} \tag{26}$$

When $\frac{(d-b)(e^c+e^g)}{(a-c)\cdot e^g\cdot(d-h)} \leq \beta \leq \frac{(d-b)(e^a+e^g)}{(a-c)\cdot e^g\cdot(b-h)}$. The second mover would like to choose "c" with probability p if $U_2(c, C) = U_2(d, C)$. Now we have p satisfies:

$$\frac{e^g}{e^{ap+c(1-p)} + e^g}[bp + d(1-p) - h] = \frac{d-b}{(a-c)\beta} \tag{27}$$

First, look at (25) the threshold of $\frac{(d-b)(e^a+e^g)}{(a-c)\cdot e^g\cdot(b-h)}$, take first order condition with respect to g, we get $\frac{-(d-b)\cdot e^{a-g}}{(a-c)\cdot(b-h)} < 0$. This means that the threshold will decrease if g increases. So for the same distribution over the population, more second movers would like to choose "c" after "C" if g increases. This part is consistent with our Proposition 4(1).

Next, see (26) the threshold of $\frac{(d-b)(e^c+e^g)}{(a-c)\cdot e^g\cdot(d-h)}$, take first order condition with respect to g, we get $\frac{-(d-b)\cdot e^{c-g}}{(a-c)\cdot(b-h)} < 0$. This means that the threshold will decrease if g increases. So for the same distribution over the population, less second movers would like to choose "d" after "C" if g increases. This part is also consistent with our Proposition 4(1).

Finally, see (27) when $\frac{(d-b)(e^c+e^g)}{(a-c)\cdot e^g\cdot(d-h)} \leq \beta \leq \frac{(d-b)(e^a+e^g)}{(a-c)\cdot e^g\cdot(b-h)}$. The second mover would like to choose "y" with probability p which satisfies $\frac{e^g}{e^{ap+c(1-p)}+e^g}[bp+d(1-p)-h] = \frac{d-b}{(a-c)\beta}$. We can find that, for a given β, p will increase if g increases since $\beta = \frac{(d-b)[e^{ap+c(1-p)-g}+1]}{(a-c)\cdot[bp+d(1-p)-h]}$

and $a > c$ and $d > b$. This means that when $\beta \in [\frac{(d-b)(e^c+e^g)}{(a-c)\cdot e^g\cdot(d-h)}, \frac{(d-b)(e^a+e^g)}{(a-c)\cdot e^g\cdot(b-h)}]$, more second movers would like to choose "c" after "C" if g increases. Again, this part is also consistent with our Proposition 4(1).

Similarly, we find that player is more likely to respond selfishly when the value of h increases.

Having the threshold of $\frac{(d-b)(e^a+e^g)}{(a-c)\cdot e^g\cdot(b-h)}$ from (25), take first order condition with respect to h, we get $\frac{(d-b)(e^a+e^g)}{(a-c)\cdot e^g\cdot(b-h)^2} > 0$. This means that the threshold will increase if h increases. So for the same distribution over the population, less second movers would like to choose "c" after "C" if h increases. This part is consistent with our Proposition 4(2).

Then, look at the threshold of $\frac{(d-b)(e^c+e^g)}{(a-c)\cdot e^g\cdot(d-h)}$ from (26), take first order condition with respect to h, we get $\frac{(d-b)(e^c+e^g)}{(a-c)\cdot e^g\cdot(d-h)^2} > 0$. This means that the threshold will increase if h increases. So for the same distribution over the population, more second movers would like to choose "d" after "C" if h increases. This part is also consistent with our Proposition 4(2).

Last, the threshold of $\frac{(d-b)(e^c+e^g)}{(a-c)\cdot e^g\cdot(d-h)} \leq \beta \leq \frac{(d-b)(e^a+e^g)}{(a-c)\cdot e^g\cdot(b-h)}$ from (27). The player 2 would like to choose "y" with probability p which satisfies $\frac{e^g}{e^{ap+c(1-p)}+e^g}[bp+d(1-p)-h] = \frac{d-b}{(a-c)\beta}$. We can find that, for a given β, p will decrease if h increases since $\beta = \frac{(d-b)[e^{ap+c(1-p)-g}+1]}{(a-c)\cdot[bp+d(1-p)-h]}$ and $a > c$ and $d > b$. This means that when $\beta \in [\frac{(d-b)(e^c+e^g)}{(a-c)\cdot e^g\cdot(d-h)}, \frac{(d-b)(e^a+e^g)}{(a-c)\cdot e^g\cdot(b-h)}]$, less second movers would like to choose "c" after "C" if h increases. Again, this part is also consistent with our Proposition 4(2).

To conclude, the decision maker will also evaluate his opponent's status and full underlying intentions. he will consider how much his opponent potentially gives him and why his opponent gives him more (avoid a worse result or purely kind to him).

A.3.2 proof of Proposition 5

Another more complicated case is that the second mover has the chance to punish the first mover. First of all, set p=p(c,C) and q=p(p,D). Then we derive the second mover's utility given the first mover chooses "C" as follows:

$$U_2(c, C) = b + a \cdot \beta \cdot \delta_{12}$$

$$U_2(d, C) = d + c \cdot \beta \cdot \delta_{12}$$

where $\delta_{12} = bp+d(1-p)-\{\frac{e^{ap+c(1-p)}}{e^{ap+c(1-p)}+e^{iq+g(1-q)}}[bp+d(1-p)]+\frac{e^{iq+g(1-q)}}{e^{ap+c(1-p)}+e^{iq+g(1-q)}}[jq+h(1-q)]\} = \frac{e^{iq+g(1-q)}}{e^{ap+c(1-p)}+e^{iq+g(1-q)}}[bp+d(1-p)-jq-h(1-q)]$.

So the second mover would like to choose "c" if $U_2(c, C) > U_2(d, C)$. Now p=1, and we have:

$$\beta > \frac{(d-b)(e^a + e^{iq+g(1-q)})}{(a-c)\cdot e^{iq+g(1-q)}\cdot[b - jq - h(1-q)]} \quad (28)$$

The second mover would like to choose "n" if $U_2(d,C) > U_2(c,C)$. Now p=0, and we have:

$$\beta < \frac{(d-b)(e^c + e^{iq+g(1-q)})}{(a-c)\cdot e^{iq+g(1-q)}\cdot[d-jq-h(1-q)]} \tag{29}$$

When $\frac{(d-b)(e^c+e^{iq+g(1-q)})}{(a-c)\cdot e^{iq+g(1-q)}\cdot[d-jq-h(1-q)]} \leq \beta \leq \frac{(d-b)(e^a+e^{iq+g(1-q)})}{(a-c)\cdot e^{iq+g(1-q)}\cdot[b-jq-h(1-q)]}$. The second mover would like to choose "c" with probability p if $U_2(c,C) = U_2(d,C)$. Now we have p satisfies:

$$\frac{e^{iq+g(1-q)}}{e^{ap+c(1-p)} + e^{iq+g(1-q)}}[bp + d(1-p) - jq - h(1-q)] = \frac{d-b}{(a-c)\beta} \tag{30}$$

Secondly, let's discuss the second mover's utility given the first mover chooses "D"[23]:

$$U_2(p,D) = j + i\cdot\beta\cdot\delta_{12}$$
$$U_2(d,D) = h + g\cdot\beta\cdot\delta_{12}$$

where $\delta_{12} = jq+h(1-q) - \{\frac{e^{iq+g(1-q)}}{e^{ap+c(1-p)}+e^{iq+g(1-q)}}[jq+h(1-q)] + \frac{e^{ap+c(1-p)}}{e^{ap+c(1-p)}+e^{iq+g(1-q)}}[bp+d(1-p)]\} = \frac{e^{ap+c(1-p)}}{e^{ap+c(1-p)}+e^{iq+g(1-q)}}[jq+h(1-q)-bp-d(1-p)]$.

So now the second mover would like to choose "p" if $U_2(p,D) > U_2(d,D)$. Now q=1, and we have:

$$\beta > \frac{(h-j)(e^{ap+c(1-p)} + e^i)}{(g-i)\cdot e^{ap+c(1-p)}\cdot[bp+d(1-p)-j]} \tag{31}$$

The second mover would like to choose "d" if $U_2(p,D) < U_2(d,D)$. Now q=0, and we have:

$$\beta < \frac{(h-j)(e^{ap+c(1-p)} + e^g)}{(g-i)\cdot e^{ap+c(1-p)}\cdot[bp+d(1-p)-h]} \tag{32}$$

We have obtained (28) to (32) now. If we want to explore player 2's cooperation rate after the first mover's "C", we should consider the different values of p and q. Therefore, we should have six combinations:(28) and (31), (28) and (32), (29) and (31), (29) and (32), (30) and (31), (30) and (32).

Looking at (28)'s combinations. First, (28) and (31) where p=1 and q=1. Therefore, we should have $\beta > \frac{(d-b)(e^a+e^i)}{(a-c)\cdot e^i\cdot[b-j]} \cap \beta > \frac{(h-j)(e^a+e^i)}{(g-i)\cdot e^a\cdot[b-j]}$. Second, (28) and (32) where p=1 and q=0, again, we have $\beta > \frac{(d-b)(e^a+e^g)}{(a-c)\cdot e^g\cdot[b-h]} \cap \beta < \frac{(h-j)(e^a+e^g)}{(g-i)\cdot e^a\cdot[b-h]}$. Overall, the second mover will cooperate when $\{\beta > \frac{(d-b)(e^a+e^i)}{(a-c)\cdot e^i\cdot[b-j]} \cap \beta > \frac{(h-j)(e^a+e^i)}{(g-i)\cdot e^a\cdot[b-j]}\} \cup \{\beta > \frac{(d-b)(e^a+e^g)}{(a-c)\cdot e^g\cdot[b-h]} \cap \beta < \frac{(h-j)(e^a+e^g)}{(g-i)\cdot e^a\cdot[b-h]}\}$.

Proposition 5 assumes a low cost of punishment $(h-j)(d-j) < \frac{e^c(d-b)(g-i)(d-h)(e^c+e^i)}{e^g(a-c)(e^c+e^g)}$ that implies $\frac{(d-b)(e^c+e^i)}{(a-c)e^g(d-j)} > \frac{(h-j)(e^c+e^g)}{(g-i)e^c(d-h)}$. Because $g > i$ and $h > j$, we must have $\frac{h-j}{(g-i)e^c} <$

[23]The same proof as the sequential prisoner's dilemma in section 4.2.1, the second mover never chooses "c" to reward the unkind behaviour. So we only consider "d" and "p" after "D'.

$\frac{d-b}{(a-c)e^g}$. Then we obtain $\beta > \frac{(d-b)(e^a+e^g)}{(a-c)\cdot e^g\cdot[b-h]} \cap \beta < \frac{(h-j)(e^a+e^g)}{(g-i)\cdot e^a\cdot[b-h]} = \varnothing$ as a ¿ c. For (28) and (31), we should have $\frac{h-j}{(g-i)e^a} < \frac{d-b}{(a-c)e^g} < \frac{d-b}{(a-c)e^i}$ as $i < g$. Then we get $\beta > \frac{(d-b)(e^a+e^i)}{(a-c)\cdot e^i\cdot[b-j]} \cap \beta > \frac{(h-j)(e^a+e^i)}{(g-i)\cdot e^a\cdot[b-j]} = \beta > \frac{(d-b)(e^a+e^i)}{(a-c)\cdot e^i\cdot[b-j]}$. Totally, the first mover would like to cooperate if $\beta > \frac{(d-b)(e^a+e^i)}{(a-c)\cdot e^i\cdot[b-j]}$, and it is not hard to find that the threshold $\frac{(d-b)(e^a+e^i)}{(a-c)\cdot e^i\cdot[b-j]}$ will increase if the value of j increases. So for the same distribution over the population, less second movers will cooperate if j increases.

Then look at (29)'s combination. First, (29) and (31) where p=0 and q=1. Therefore, we should have $\beta < \frac{(d-b)(e^c+e^i)}{(a-c)\cdot e^i\cdot[d-j]} \cap \beta > \frac{(h-j)(e^c+e^i)}{(g-i)\cdot e^c\cdot[d-j]}$. Second, (29) and (32) where p=0 and q=0, again, we have $\beta < \frac{(d-b)(e^c+e^g)}{(a-c)\cdot e^g\cdot[d-h]} \cap \beta < \frac{(h-j)(e^c+e^g)}{(g-i)\cdot e^c\cdot[d-h]}$. Overall, the second mover will defect when $\{\beta < \frac{(d-b)(e^c+e^i)}{(a-c)\cdot e^i\cdot[d-j]} \cap \beta > \frac{(h-j)(e^c+e^i)}{(g-i)\cdot e^c\cdot[d-j]}\} \cup \{\beta < \frac{(d-b)(e^c+e^g)}{(a-c)\cdot e^g\cdot[d-h]} \cap \beta < \frac{(h-j)(e^c+e^g)}{(g-i)\cdot e^c\cdot[d-h]}\}$.

Proposition 5 assumes a low cost of punishment $(h-j)(d-j) < \frac{e^c(d-b)(g-i)(d-h)(e^c+e^i)}{e^g(a-c)(e^c+e^g)}$ that implies $\frac{h-j}{(g-i)e^c} < \frac{d-b}{(a-c)e^g}$. Then for (29) and (32), we should have $\beta < \frac{(d-b)(e^c+e^g)}{(a-c)\cdot e^g\cdot[d-h]} \cap \beta < \frac{(h-j)(e^c+e^g)}{(g-i)\cdot e^c\cdot[d-h]} = \beta < \frac{(h-j)(e^c+e^g)}{(g-i)\cdot e^c\cdot[d-h]}$. For (29) and (31), we should have $\frac{h-j}{(g-i)e^a} < \frac{h-j}{(g-i)e^c} < \frac{d-b}{(a-c)e^g} < \frac{d-b}{(a-c)e^i}$ as $i < g$ and $c < a$. Then we get $\beta < \frac{(d-b)(e^c+e^i)}{(a-c)\cdot e^i\cdot[d-j]} \cap \beta > r\frac{(h-j)(e^c+e^i)}{(g-i)\cdot e^c\cdot[d-j]} = \frac{(h-j)(e^c+e^i)}{(g-i)\cdot e^c\cdot[d-j]} < \beta < \frac{(d-b)(e^c+e^i)}{(a-c)\cdot e^i\cdot[d-j]}$. Therefore, the second mover would like to defect if $\beta \in (\frac{(h-j)(e^c+e^i)}{(g-i)\cdot e^c\cdot[d-j]}, \frac{(d-b)(e^c+e^i)}{(a-c)\cdot e^i\cdot[d-j]}) \cup (0, \frac{(h-j)(e^c+e^g)}{(g-i)\cdot e^c\cdot[d-h]})$. But here we notice that $\frac{(h-j)(e^c+e^i)}{(g-i)\cdot e^c\cdot[d-j]} < \frac{(h-j)(e^c+e^g)}{(g-i)\cdot e^c\cdot[d-h]}$ since g¿ i and h ¿ j, and $\frac{(d-b)(e^c+e^i)}{(a-c)e^i(d-j)} > \frac{(h-j)(e^c+e^g)}{(g-i)e^c(d-h)}$ since $\frac{(d-b)(e^c+e^i)}{(a-c)e^i(d-j)} > \frac{(d-b)(e^c+e^i)}{(a-c)e^g(d-j)}$ as g¿ i. So we further have the second mover would like to defect if $\beta \in (0, \frac{(d-b)(e^c+e^i)}{(a-c)\cdot e^i\cdot[d-j]})$. It is not hard to find that the threshold $\frac{(d-b)(e^c+e^i)}{(a-c)\cdot e^i\cdot[d-j]}$ will increase if the value of j increases. It suggests that for the same distribution over the population, more second movers will defect if j increases.

Next, looking at (30)'s combination. According to the above analysis, the second mover will cooperate with p satisfies $\frac{e^{iq+g(1-q)}}{e^{ap+c(1-p)}+e^{iq+g(1-q)}}[bp + d(1-p) - jq - h(1-q)] = \frac{d-b}{(a-c)\beta}$ when $\beta \in (\frac{(d-b)(e^c+e^i)}{(a-c)\cdot e^i\cdot[d-j]}, \frac{(d-b)(e^a+e^i)}{(a-c)\cdot e^i\cdot[b-j]})$. Under this case, we have two possibilities based on (31) and (32) where q=1 or q=0. If q=0, we have $\frac{e^g}{e^{ap+c(1-p)}+e^g}[bp+d(1-p)-h] = \frac{d-b}{(a-c)\beta}$. It is easy to find that j will not influence the probability of cooperation given the first mover cooperates. if q=1, we have $\frac{e^i}{e^{ap+c(1-p)}+e^i}[bp + d(1-p) - j] = \frac{d-b}{(a-c)\beta}$. Under this case, given unchanged β, if j increases, if we want a unchanged β, the cooperation p must decrease as d>b and a>c. So for (30)'s combination, we still observe that more second movers prefer defect if j increases.

To conclude this six combinations, we find that more second movers would like to negatively reciprocal when the cost of punishment is relatively low: $(h-j)(d-j) < \frac{e^c(d-b)(g-i)(d-h)(e^c+e^i)}{e^g(a-c)(e^c+e^g)}$ and the cost is smaller (j increases).

References

Ahn, Toh-Kyeong et al. (2001). "Cooperation in PD games: Fear, greed, and history of play". In: *Public Choice* 106.1, pp. 137–155.

Ahn, Toh-Kyeong et al. (2007). "Asymmetric payoffs in simultaneous and sequential prisoner's dilemma games". In: *Public Choice* 132.3, pp. 353–366.

Andreoni, James and B Douglas Bernheim (2009). "Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects". In: *Econometrica* 77.5, pp. 1607–1636.

Bardsley, Nicholas (2008). "Dictator game giving: altruism or artefact?" In: *Experimental economics* 11.2, pp. 122–133.

Battigalli, Pierpaolo and Martin Dufwenberg (2009). "Dynamic psychological games". In: *Journal of Economic Theory* 144.1, pp. 1–35.

— (2022). "Belief-dependent motivations and psychological game theory". In: *Journal of Economic Literature* 60.3, pp. 833–82.

Berg, Joyce, John Dickhaut, and Kevin McCabe (1995). "Trust, reciprocity, and social history". In: *Games and economic behavior* 10.1, pp. 122–142.

Blount, Sally (1995). "When social outcomes aren t fair: The effect of causal attributions on preferences". In: *Organizational behavior and human decision processes* 63.2, pp. 131–144.

Bolton, Gary E and Axel Ockenfels (2000). "ERC: A theory of equity, reciprocity, and competition". In: *American economic review* 90.1, pp. 166–193.

Camerer, Colin F (1997). "Progress in behavioral game theory". In: *Journal of economic perspectives* 11.4, pp. 167–188.

Camerer, Colin F, Marc J Knez, and Roberto A Weber (1996). "Timing and Virtual Observability in Ultimatum Bargaining and Weak Link Coordination Games". In.

Çelen, Boğaçhan, Andrew Schotter, and Mariana Blanco (2017). "On blame and reciprocity: Theory and experiments". In: *Journal of Economic Theory* 169, pp. 62–92.

Charness, Gary and Matthew Rabin (2002). "Understanding social preferences with simple tests". In: *The quarterly journal of economics* 117.3, pp. 817–869.

Charness, Gary, Luca Rigotti, and Aldo Rustichini (2016). "Social surplus determines cooperation rates in the one-shot Prisoner's Dilemma". In: *Games and Economic Behavior* 100, pp. 113–124.

Clark, Kenneth and Martin Sefton (2001). "The sequential prisoner's dilemma: evidence on reciprocation". In: *The economic journal* 111.468, pp. 51–68.

Cooper, Russell et al. (1993). "Forward induction in the battle-of-the-sexes games". In: *The American Economic Review*, pp. 1303–1316.

Dana, Jason, Roberto A Weber, and Jason Xi Kuang (2007). "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness". In: *Economic Theory* 33.1, pp. 67–80.

Dawes, Robyn M and Richard H Thaler (1988). "Anomalies: cooperation". In: *Journal of economic perspectives* 2.3, pp. 187–197.

Dhaene, Geert and Jan Bouckaert (2010). “Sequential reciprocity in two-player, two-stage games: An experimental analysis”. In: *Games and Economic Behavior* 70.2, pp. 289–303.

Dufwenberg, Martin and Georg Kirchsteiger (2004). “A theory of sequential reciprocity”. In: *Games and economic behavior* 47.2, pp. 268–298.

— (2019). “Modelling kindness”. In: *Journal of Economic Behavior & Organization* 167, pp. 228–234.

Dufwenberg, Martin, Alec Smith, and Matt Van Essen (2013). “Hold-up: with a vengeance”. In: *Economic Inquiry* 51.1, pp. 896–908.

Engel, Christoph and Lilia Zhurakhovska (2016). “When is the risk of cooperation worth taking? The prisoner’s dilemma as a game of multiple motives”. In: *Applied Economics Letters* 23.16, pp. 1157–1161.

Falk, Armin (2007). “Gift exchange in the field”. In: *Econometrica* 75.5, pp. 1501–1511.

Falk, Armin, Ernst Fehr, and Urs Fischbacher (2003). “On the nature of fair behavior”. In: *Economic inquiry* 41.1, pp. 20–26.

— (2008). “Testing theories of fairness—Intentions matter”. In: *Games and Economic Behavior* 62.1, pp. 287–303.

Falk, Armin and Urs Fischbacher (2006). “A theory of reciprocity”. In: *Games and economic behavior* 54.2, pp. 293–315.

Fehr, Ernst and Klaus M Schmidt (1999). “A theory of fairness, competition, and cooperation”. In: *The quarterly journal of economics* 114.3, pp. 817–868.

Franzen, Axel and Sonja Pointner (2012). “Anonymity in the dictator game revisited”. In: *Journal of Economic Behavior & Organization* 81.1, pp. 74–81.

Gächter, Simon, Kyeongtae Lee, and Martin Sefton (2022). *The variability of conditional cooperation in sequential prisoner’s dilemmas*. Tech. rep. CeDEx Discussion Paper Series.

Gächter, Simon et al. (2021). “Risk, Temptation, and Efficiency in the One-Shot Prisoner’s Dilemma”. In: *Available at SSRN 3979354*.

Geanakoplos, John, David Pearce, and Ennio Stacchetti (1989). “Psychological games and sequential rationality”. In: *Games and economic Behavior* 1.1, pp. 60–79.

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze (1982). “An experimental analysis of ultimatum bargaining”. In: *Journal of economic behavior & organization* 3.4, pp. 367–388.

Isoni, Andrea and Robert Sugden (2019). “Reciprocity and the Paradox of Trust in psychological game theory”. In: *Journal of Economic Behavior & Organization* 167, pp. 219–227.

Jiang, Lianjie and Jiabin Wu (2019). “Belief-updating rule and sequential reciprocity”. In: *Games and Economic Behavior* 113, pp. 770–780.

Kahneman, Daniel, Jack L Knetsch, and Richard H Thaler (1986). “Fairness and the assumptions of economics”. In: *Journal of business*, S285–S300.

Klempt, Charlotte (2012). “Fairness, spite, and intentions: Testing different motives behind punishment in a prisoners’ dilemma game”. In: *Economics Letters* 116.3, pp. 429–431.

List, John A (2007). "On the interpretation of giving in dictator games". In: *Journal of Political economy* 115.3, pp. 482–493.

Messick, David M (1995). "Equality, fairness, and social conflict". In: *Social Justice Research* 8.2, pp. 153–173.

Orhun, A Yeşim (2018). "Perceived motives and reciprocity". In: *Games and Economic Behavior* 109, pp. 436–451.

Rabin, Matthew (1993). "Incorporating fairness into game theory and economics". In: *The American economic review*, pp. 1281–1302.

Schneider, Mark and Timothy W Shields (2022). "Motives for Cooperation in the One-Shot Prisoner's Dilemma". In: *Journal of Behavioral Finance, forthcoming.*

Sohn, Jin-yeong and Wenhao Wu (2022). "Reciprocity with uncertainty about others". In: *Games and Economic Behavior* 136, pp. 299–324.

Thaler, Richard H (1988). "Anomalies: The ultimatum game". In: *Journal of economic perspectives* 2.4, pp. 195–206.

Warner, Stanley L (1965). "Randomized response: A survey technique for eliminating evasive answer bias". In: *Journal of the American Statistical Association* 60.309, pp. 63–69.